8-17-2016 12:00 AM

Anisotropic Waveform Tomography: Application to Crosshole data for Transversely Isotropic Media

Shaun Michael Hadden
The University of Western Ontario

Supervisor
Dr. R. 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

Abstract

Anisotropic Traveltime Tomography and Full Waveform Inversion were applied first to synthetic and then to real data following the development of a transversely isotropic model for handling anisotropy. Best-fitting models of seismic velocity and Thomsen's anisotropy parameters were initially obtained from traveltime tomography, and then used as the starting models for Full Waveform Inversion. The use of a Laplace transform approach effectively damps late arriving S-wave artifacts that introduce errors into the modelling process. The results of the synthetic study highlights the tradeoffs in resolution between the two parameter classes, but verify anisotropic traveltime tomography as a valid method for generating starting models for Full Waveform Inversion. The joint technique was then applied to field gathers from Western Canada and compared to a similar analyses that used a simpler anisotropy model. The transversely isotropic approach yielded a Full Waveform Inversion model with superior resolution that better predicted the true data.

Keywords: Tilted Transverse Isotropy, Traveltime Tomography, Full Waveform Inversion, Seismic Wave Modelling, Crosswell Seismic Survey, Anisotropy Modelling, Inverse Methods

Acknowledgements

First and foremost, I would like to thank Dr. Gerhard Pratt for agreeing to supervise my MSc. project. Without his continued guidance and support, as well as his constant communication through email as well as in person, I would not have been able to complete this work in such a timely matter. Furthermore, Dr. Pratt's expertise in the field of FWI contributed significantly to the impact of my research, and I would also like to thank him for granting me full access to his lab and computer workstations, as well as his traveltime tomography and Full Waveform Inversion software packages. I would also like to extend special thanks to Dr. Brendan Smithyman, who went above and beyond his role as a Post Doctoral Fellow within our research group to help me in completing my thesis project. Dr. Smithyman was the main author behind the Python-Fullwv integration that allowed me to develop the FWI workflow described in this Thesis, and he was always willing to aid me in troubleshooting my coding bugs, and served as a crucial resource throughout the various stages of my project. I would also like to thank Dr. Robert Shcherbakov for agreeing to be a part of my Supervisory Committee and for providing me with technical feedback on my interim progress report, as well as my final submission.

I think it is also important for me to acknowledge and thank the various members of my research group over the past two years: Bernie Dunn, Akbar Zuberi, German Rubino, Sjors Mol, Felix Hlousek and Benjamin Consolvo. I would like to thank you all for enlightening discussions, providing feedback at our weekly group meetings, for always offering a second pair of eyes if I was having trouble debugging or interpreting my results, as well as for your personal support through many long days and nights in the waveform inversion lab. I thoroughly enjoyed your company, and I hope to stay in touch will all of you. Outside of our research group, I would like to thank Dr. Stephane Operto for his correspondence regarding the finite-difference frequency-domain implementation discussed in Chapter 2, as well as Dr. Hansruedi Maurer for his correspondence and for connecting me to various other researches within the acoustic Full Waveform Inversion field.

Next, I must thank my family for supporting me in my decision to remain abroad and pursue my Master's Degree, as well as for encouraging me and providing positive reinforcement through challenging periods. I know that it is difficult to be separated from your loved ones so I

sincerely appreciate your understanding in both my academic and career decisions. I would like to thank my partner Kendra Davis who has stood directly by my side over the past two years and whose unconditional love and support helped me achieve the level of success demonstrated by this work. I must also thank Brenda Davis, Jim McNair and Keith Davis who together have been by 'Canadian Parents' over the past two years. Your emotional and physical support has provided me with confidence and stability through some of the busier moments of these past few years.

Last, but certainly not least, I would like to thank all of my fellow graduate and undergraduate students whom I have met and spent time with over the last two years: Derek Kouhi, Sidhanth Kothari, Krysta Kaski, Mark Novakovic, Emma-Dawn-Furguson, Simon Thorpe, Freddie Jackson, Joe Farrugia, Andrew Thomson, Joelle Langford, Ronan Drysdale and Joanna Holmgren. Thank you all for supporting me, whether it be through attending my presentations and providing me with valued feedback, or by simply spending time with me on weekends and weeknights. In conclusion, I would like to once again acknowledge everyone mentioned above as you have all played a critical role in my two-year journey through the Master's programme here at Western; thank you all for your support, and I wish you all the best in your future endeavours.

Statement of Co-Authorship

The codebases for traveltime tomography and full waveform inversion were written and provided to me by Dr. Gerhard Pratt. The anisotropic forward operator and finite-difference implementation was developed entirely by myself in *Python*, however they are integrated into the seismic wavefield modelling framework named *Zephyr*, which was primarily developed by Dr. Brendan Smithyman. Dr. Smithyman also wrote the extension module that allowed my *Python* scripts to interface with the full waveform inversion code, which was written in *FORTRAN*. All of the traveltime tomography and full waveform inversion strategies, discussions and results presented in Chapters 3 and 4 were performed and developed entirely by myself, with suggestions for improvement from Dr. Pratt, Dr. Smithyman and Akbar Zuberi. The field data presented in Chapter 4 were provided to us in strict confidentiality and anonymity, and the data preprocessing was performed by Dr. Pratt. Finally, the conclusions on the success of anisotropic FWI and the recommendations for future studies were developed entirely by myself.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Geophysics can be broadly defined as the study of Earth and Space's natural processes using well established physical phenomena. More specifically, Exploration Geophysics is an applied branch of Geophysics in which these physical methods are used to locate, delineate or quantify important features within the Earth's subsurface. A simple example of this is the study of Earth's varying magnetic field to develop models of magnetic susceptibility, which can then be used to locate economic metals that are included in, or are associated with, magnetic minerals (see Nabighian et al. (2005) for a recent review of this exploration technique).

The study of vibrations in the form of seismic waves travelling through the earth's subsurface has emerged as the most widely-used technique in modern geophysical applications due to its extreme relevance to the hydrocarbon exploration industry. Over the last five decades, seismic surveys have been conducted for a variety of industrial applications, including the imaging of layered sediments for hydrocarbon exploration (Bois et al., 1972; Shin and Min, 2006), fault delineation (Pratt and Shipp, 1999), gas hydrate imaging (Pratt et al., 2005), ore grade estimation in mines (Perozzi et al., 2012) and geotechnical site investigation (Pinches and Thompson, 1990). The primary objective of such surveys is to develop a model of the distribution of seismic velocities from the seismic arrivals measured between each source and receiver pair. These velocities are then used to image structures such as fracture networks, identify map lithologies and changes in porosity, locate stratigraphic differences as well as to characterize the in-situ stress regime, as these features all typically correlate with contrasts in seismic velocity.

In present day studies, seismic velocity models can be determined in high resolution from the joint analysis of the arrival times, as well as the recorded waveforms themselves (see Brenders and Pratt (2007) and Symes (2008) for successful applications of these joint analyses). Traveltime Tomography models and inverts traveltimes along raypaths, using simplified physics. Full Waveform Inversion (FWI) is concerned with fitting the amplitude and phase of waveform data. Traveltime analyses are often limited by over-simplified assumptions, whereas

the multifrequency approach of FWI produces superior resolution and yields more information about the propagation medium (Brenders and Pratt, 2007). A limitation of FWI, however, is a stringent requirement that the starting velocity models be significantly closer to the true velocities when compared to the traveltime tomography case. Consequently, best-fitting velocities obtained from traveltime tomography are often used as starting models for FWI. In this work, I will refer to the joint analyses of arrival times followed by waveform phase and amplitude as Waveform Tomography (WT), after Brenders (2011).

Two major limitations that must be considered during WT are the presence of significant velocity variations across the media (inhomogeneity), as well as dependencies on the direction of propagation (anisotropy). Figure 1.1 illustrates the differences between these two variational properties. Although large velocity variations introduce significant complexities into the equations, this is well accounted for by several complete modelling techniques, such as the finite difference schemes proposed by Virieux (1986) and Pratt (1990) for the time- and frequency-domain modelling of elastic waves respectively. Properly accounting for anisotropy, however, is nontrivial as it requires a more complex forward modelling algorithm, and typically requires *a priori* information about the symmetric properties of the rockmass. Chapman and Pratt (1992) proposed techniques for traveltime tomography in weakly anisotropic, heterogeneous media. This approach has been used in tandem with the frequency-domain waveform inversion method of Pratt et al. (1998) for the past two and a half decades to resolve velocity models for several case studies (see Pratt et al., 2005; Afanasiev et al., 2014).

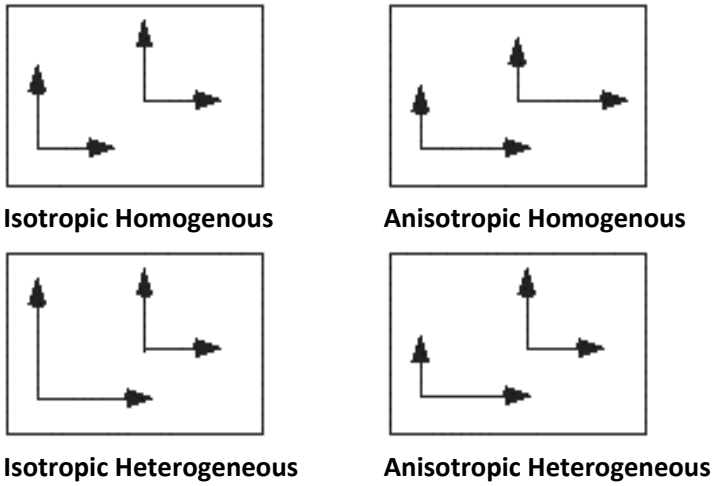


Figure 1.1: Illustration showing the differences between heterogeneity and anisotropy. In this context, the length of the arrows represents the magnitude of seismic velocity and the boxes represent the geological medium. Original image modified from http://www.waterrights.utah.gov/wellinfo/theis/darcy5.gif.

Results of this joint technique, however, have been restricted to date by the fact that the method of Pratt et al. (1998) assumes isotropic or simplified one-dimensional anisotropic waveform modelling. Presently, the two methods are only compatible if the results produced by the traveltime method are simplified significantly. This can have severe implications in structurally complex environments where significant anisotropy exists, as they require substantial anisotropy models for the waveform inversion solution to converge. For example, Afanasiev et al. (2014) required an alternative approach for generating the starting models for FWI after the ones obtained from traveltime tomography produced unsatisfactory results. The focus of this thesis, therefore, is to improve the accuracy of Anisotropic Waveform Tomography (AWT) by recasting the frequency-domain method of Pratt et al. (1998) to follow a more general isotropy system.

1.1 Crosswell Seismic Imaging

At the exploration scale, early seismic experiments were limited to the traditional 'surface seismic' acquisition geometry, whereby all source and receiver locations are placed at the surface of the survey region. This setup relies on waves travelling directly into the subsurface from the source location and then returning to the surface where the arriving energy is measured at the receiver locations. More specifically, the seismic data are analyzed for the presence of reflected waves, which are reflected back to the surface when a velocity contrast is encountered, as well as transmitted and diving waves, which are refracted waves bent back towards the surface in strong velocity-gradient zones. In present day surface seismic experiments are still the most widely used geophysical survey due to the expansive coverage that acquisition geometries can provide, as well as their ability to successfully image subsea geological structures through the use of ocean bottom streamers (OBS) and other marine acquisition technology. These features coupled with the growing investment in seismic migration analyses (see Yilmaz, 2001) make surface seismic surveys the ideal candidate when surveying large regions to image complex hydrocarbon fields or to understand the local tectonic regime (Claerbout et al., 1985; Sava and Fomel, 2003).

Despite the numerous successes surrounding the surface seismic technique, there are also several limitations which called for and gave rise to alternative acquisition methods. Transmission of higher frequencies is limited by the intrinsic attenuation of the earth, as these frequencies attenuate faster and are therefore restricted in the temporal bandwidth of the data when the signal-to-noise ratio (SNR) is considered. Furthermore, survey aperture is limited by the spatial position of receivers which can be restricted by surface topography, the expected quality of the source signal, as well physical constraints on the geophone arrays themselves. The latter two

are less significant in modern studies, as the development of source and receiver technology has made the collection and analysis of wide-aperture surface data more feasible (Hill et al., 2006; Savazzi and Spagnolini, 2008).

To overcome the bandwidth limits imposed by surface acquisition geometries, seismic researchers began deploying source and receiver arrays down boreholes, which led to the simultaneous development of two seismic acquisition methods: well-to-surface/surface-to-well and well-to-well measurements. The former is commonly referred to as a vertical seismic profile (VSP) survey, with applications dating back to the late 1940s (Gardner, 1949). The latter is also referred to as a crosswell or crosshole survey. One of the first known applications of crosswell seismic imaging was presented by Ricker (1953) who studied wave propagation through the Pierre Shales in eastern Colorado by deploying dynamite charges and receiver into boreholes that were a few hundred meters deep. Crosswell surveys are ideal for smaller scale projects where the target's location is known to within a certain degree of accuracy. This allows the borehole locations to be chosen to cover specific areas within the subsurface. Due to the intrinsic downscaling of crosshole applications, the observed propagation distances for seismic waves are greatly reduced, which relaxes the high frequency limit on the bandwidth imposed by surface acquisitions. As crosswell techniques saw increasing popularity, several seismic source types were developed, such as downhole airguns (Lee et al., 1984), vibratory sources (Airhart, 1989), piezoelectric transducers (Wong et al., 1983) and electric discharge sources (Owen et al., 1988). Crosswell imaging techniques can be broadly classified into direct arrival methods (Lines and LaFehr, 1989; Chapman and Pratt, 1992) and reflection methods (Harris et al., 1995). The former of these will be employed within this thesis. Figure 1.2 illustrates the acquisition geometry for a conventional crosswell survey.

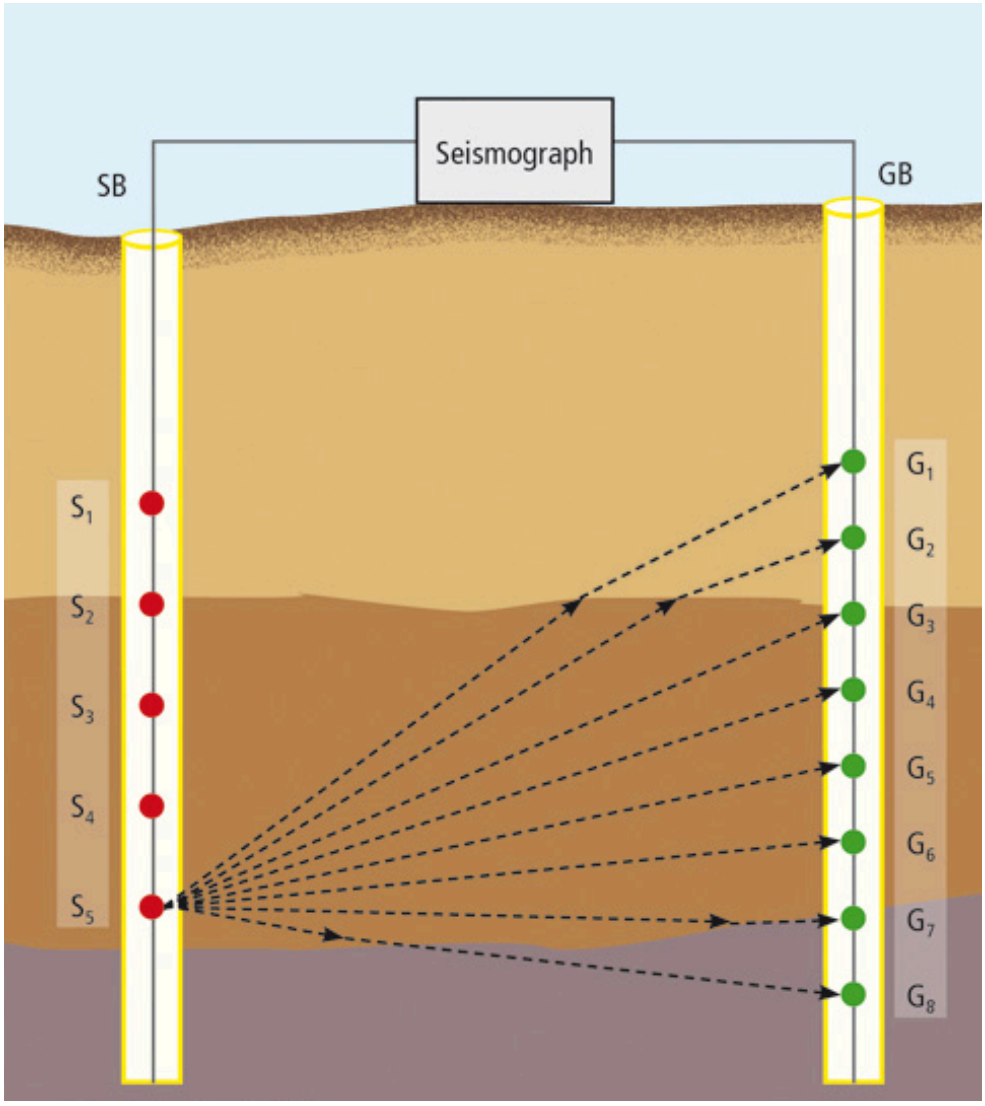


Figure 1.2: Typical acquisition geometry for crosswell seismic surveys. The sources are denoted by 'S' and the receivers are denoted by 'G'. Original image modified from http://asstgroup.com/images/crosshole_seismic_profiling_s.jpg.

1.2 Seismic Anisotropy

A substance is classified as 'anisotropic' if its physical properties vary as a function of the direction from which they are observed. If this is not the case, these properties are the same in all orientations, and the medium is said to be isotropic. In the context of seismic wave propagation, significant advancements in modelling anisotropy were not realized until the early 1970s. Accounting for anisotropy greatly increases the number of parameters required when solving the elastic wave equation, increasing the computational power required, as well as the overall complexity of the system. One of the major developments related to seismic anisotropy studies is attributed to Thomsen (1986) who proposed the anisotropy notation necessary to describe one of the possible symmetry systems in modern anisotropy studies (the one used in this thesis). These parameters, known as 'Thomsen's Parameters', relate seismic velocity to an 'isotropic' velocity along a predefined axis of symmetry (V_{P_0}). The first parameter, ϵ, is commonly referred to as *the* 'P-wave anisotropy' as it relates the fractional difference between horizontal (V_H) and vertical (V_V) P-wave velocities. The second parameter, δ, can be thought of as 'angular anisotropy', as it describes the rate of change of velocity as a function of the angle away from the vertical axis (Tsvankin, 2012). Figure 2.3 in Chapter 2 shows an anisotropic wavefront for several combinations of ϵ and δ compared to the isotropic counterpart.

Significant anisotropy in rock sequences can be caused by preferential crack orientation and alignment of anisotropic mineral grains (e.g. clay), or the presence of thin inter-bedded isotropic layers (Winterstein, 1990). Thomsen (1986) showed that *elliptical anisotropy* ($\epsilon = \delta$), a common approach for modelling anisotropy, is usually inappropriate for cases where significant anisotropy might exist, such as thin-interbedded rock sequences within hydrocarbon-bearing sedimentary basins. This provides motivation for the further study and refinement of more general symmetry systems within the seismic modelling context. This thesis will focus primarily on Transverse Isotropy (TI), a common anisotropic model encountered in the subsurface due to the intrinsic properties of sedimentary formations. In the TI case, anisotropy varies in one symmetry direction while the other two directional variations remain equal (Winterstein, 1990). The most straightforward case of TI symmetry is one in which the symmetry axis is equal to the vertical (VTI) or horizontal (HTI) axis, the former of which will be used for the data cases in this thesis. In structurally complex regions, however, anisotropic layers might dip, such that the symmetry axis may be tilted at an angle to the vertical; this dip can be accounted for using the Tilted Transverse Isotropy (TTI) approach to the TI modelling technique. Grechka et al. (2001) showed that TTI layers can cause significant errors in current imaging techniques, but demonstrated the feasibility of this anisotropy model by performing anisotropic velocity analysis on reflection data from an overthrusted area of the Canadian Foothills.

1.3 Waveform Tomography

WT is the joint approach of traveltime tomography followed by FWI to develop high resolution models of seismic velocity (Brenders, 2011). Here I will describe the major elements pertaining to each of these techniques.

1.3.1 Traveltime Tomography

Models of seismic velocity are commonly derived from a non-linear data processing and inversion technique known as Traveltime Tomography, whereby seismic raypaths are forward modeled and the resulting arrival times are computed and compared to the observed data. The differences between the expected and the true arrivals are then used to iteratively improve the velocity model (Pratt and Chapman, 1992). Traveltime tomography is a widely accepted technique for imaging complex earth structure as it is both a robust and inexpensive method, as demonstrated by the work of Hole (1992), Pratt and Chapman (1992) and Zelt and Barton (1998) who successfully applied traveltime tomography to both real and synthetic data.

Conventional raytracing methods are based on the high-frequency wave approximation, whereby seismic energy of infinitely high frequency can be modeled as rays and will propagate in the same direction until a velocity contrast is encountered (Červený et al., 1977). The total traveltime of the wave, T, can thus be computed efficiently as the line integral measured along the infinitesimally-thin raypath from the source to the receiver, such that:

$$T = \int_{ray} \frac{1}{c(\mathbf{x})} dl, \tag{1.1}$$

where c($\mathbf{x}$) is the seismic velocity at each location along the raypath and dl is the incremental length (Cervený, 1972; Chapman and Pratt, 1992). The simplest method of raytracing is the straight-ray approximation, describing the shortest path between the source and receiver as a straight line between them. For heterogenous media, this is a poor assumption as it disregards entirely the deflection of the raypath at contrasts in seismic velocity. In modern tomographic studies, geological complexity requires more complex raytracing techniques. The 'bending' method, which was first suggested by Wesson (1971), is one such approach for more accurately computing raypaths in complex media. In this technique, an initial assumption is made of the raypath, which is then discretized into segments of varying length. With the two endpoints remaining fixed, the inner nodes along the raypath are then perturbed to minimize the total traveltime (Um and Thurber, 1987; Davison, 1991).

The resolution of ray tomography (r_{min}) has been characterized on the basis of the assump-

tions made about scattering effects. Williamson (1991) suggested that the lateral resolution of ray-based tomographic methods is restricted to the width of the first Fresnel zone, such that:

$$r_{min} = \sqrt{\frac{\lambda^2}{4} + \lambda L}, \tag{1.2}$$

$$\simeq \sqrt{\lambda L} \text{ (for small values of } \lambda), \tag{1.3}$$

where λ is the wavelength and L is the propagation distance. Nolet (1987) states that the assumption of an infinitely narrow ray only holds where λ approaches zero, which is only valid when the wavelength is small compared to the scale of the heterogeneity. The literature suggests that, when the velocity perturbation is on the scale of the wavelength, the 'ray' approximation breaks down. The width of the first Fresnel zone is often underestimated (the reduction of equation (1.2) to equation (1.3) is difficult to justify for finite frequency data), such that the claimed lateral resolution for ray-based methods must be questioned. This holds significance on the exploration scale, where broadband seismic surveys rely on the lower frequency (long wavelength) components of the data to attain maximum coverage of the surveyed region. Nevertheless, traveltime tomography is a versatile and inexpensive method for resolving velocity structure in the subsurface. As we will see further on, the models obtained from traveltime tomography are excellent candidates as starting models for Full Waveform Inversion as they describe the low wavenumber (spatial wave frequency) features of the models accurately; a characteristic which greatly aids in waveform inversion model convergence.

Chapman and Pratt (1992) proposed a traveltime perturbation method for accounting for anisotropy in weakly anisotropic, inhomogeneous media. They applied the linearized traveltime perturbation method (an application of the aforementioned 'bending method') developed by Cervený (1972, 1982); Cervený and Jech (1982) to both real and synthetic tomographic experiments, allowing for the simultaneous recovery of 2-D models of velocity, ϵ and δ (Pratt and Chapman, 1992). This technique has since been used to obtain adequate starting models of both velocity and anisotropy for several real data cases (Pratt et al., 2005; Afanasiev et al., 2014) without *a priori* information of which anisotropy system is suitable.

1.3.2 Full Waveform Inversion

In an attempt to overcome the resolution limits imposed by traditional ray tomography, mathematicians and geophysicists looked to more thorough modelling techniques that utilized the full range of information contained within the seismic signal. This gave birth to FWI; a technique whereby two-way wave equations are implemented within forward modelling and inversion

techniques to simulate waveform amplitude and phase for direct comparison against observed seismic gathers. Younger than its traveltime counterpart, FWI is based on a more complete modelling technique and is thus able to overcome some of the resolution limits imposed by traveltime tomography, with Wu and Toksöz (1987) finding the resolution to be governed by half of the seismic wavelength of the data. Until the early 1980s, seismic waveform inversion was disregarded due to its high computing and memory costs, as estimating the synthetic wavefield at each measured location was expensive. The feasibility of seismic waveform inversion saw significant improvement following the work of Lailly (1983) and Tarantola (1984), who improved the efficiency of the inversion algorithm with an adaptation of traditional pre-stack migration algorithms whereby the data residuals are back propagated and directly correlated with the forward propagated wavefields in order to obtain a model update. Despite this, the problem of model convergence (i.e. the failure of the algorithm to find a solution) was still an important consideration.

As forward modelling techniques were refined and technological advancements allowed for greater computing power and storage of larger sparse matrices (in which most of the elements are zero), Pratt (1990) and Pratt et al. (1998) designed and implemented a viscoacoustic, finite-difference method for solving the seismic waveform inversion problem in the frequency-domain, applying these 'borrowed' computational improvements. For this method, the wavefield is modelled isotropically, or a layer-by-layer coordinate stretch is applied to the isotropic grid to simulate 1-D elliptical isotropy (an approximation first proposed by Pratt et al. (2005) that is based on the earlier work of Dellinger (1991)). For 'weak anisotropy', the elliptical assumption is able to kinematically account for the effects of anisotropy; for more complex anisotropy systems, this approximation is poor as it restricts variations to one dimension and enforces ϵ to be equal to δ which is not representative of typical hydrocarbon-bearing sedimentary formations (Thomsen, 1986). Nevertheless, this method of accounting for anisotropy has been employed successfully for several real data cases (Pratt et al., 2005; Afanasiev et al., 2014).

Accounting for anisotropy within the wave equation operator has attracted great interest from FWI researchers in recent years. Operto et al. (2009) developed a 2-D finite-difference frequency-domain (FDFD) approach for modelling viscoacoustic wave propagation in transversely isotropic media with an arbitrary tilted axis; this work was based on the 2-D acoustic wave equation for Tilted Transverse Isotropy (TTI) media proposed by Zhou et al. (2006b): an extension of the intial equation proposed by Zhou et al. (2006a) to incorporate the angle of the symmetry axis from the vertical (θ_0). In this approach, the mixed grid method of Jo et al. (1996) is modified to follow the parsimonious staggered-grid method of Hustedt et al. (2004); the resulting system is discretized on two distinct coordinate systems and combined linearly

to form the parsimonious mixed-grid method. Operto et al. (2009) show good kinematic and dynamic accuracy for several synthetic anisotropic examples when compared to a time-domain elastic method, promoting their technique as a suitable candidate for handling anisotropy in FWI.

Still, the question of a suitable inversion parameterization remains due to the ill-posed nature of waveform inversion. Plessix and Cao (2011) highlighted the tradeoffs between the velocity and anisotropy parameter classes during simultaneous multiparameter FWI for VTI symmetry. Gholami et al. (2013b) performed sensitivity analyses for several possible combinations of parameters (their Table 1 on page R84) that showed wave speed as the most dominant influence on the data, followed by Thomsen's anisotropy parameters. For the crosshole case, they recommended monoparameter FWI of vertical velocity (V_{P_0}) while keeping the ϵ and δ models fixed. With continued computational advancements in the form of distributed memory platforms and massively parallelized solvers, multiparamter FWI for 2-D and 3-D surveys is now plausible for both the acoustic case (Plessix and Cao, 2011; Gholami et al., 2013a) and the elastic case (Brossier, 2011; Kamath and Tsvankin, 2013), however this thesis will focus entirely on the acoustic case[1].

[1]The 2-D method of Operto et al. (2009) has been extended to the 3D case by Operto et al. (2014), but this is also outside the scope of this project.

1.4 Objective of Thesis

The primary goal of this thesis is to develop and validate an anisotropic waveform tomography (AWT) workflow for obtaining high resolution models of seismic velocity and anisotropy for 2-D crosshole field gathers where significant anisotropy is known to exist. First, I apply the anisotropic traveltime tomography method of Chapman and Pratt (1992) to obtain best-fitting models of V_{P_0}, ϵ and δ. These models are then used as starting models for VTI FWI using the finite-difference frequency-domain (FDFD) method of Operto et al. (2009) and the parameterization proposed by Gholami et al. (2013b). I adapt the inversion and optimization scheme of Pratt et al. (1998) to allow for these more general anisotropy models. To do this, I implement the FDFD stencil into a seismic waveform modelling framework developed in *Python* (Smithyman et al., 2015), which is part of an open-source framework for waveform inversion called *Zephyr* (https://zephyr.space). Tables 1.1 and 1.2 summarize the mathematical notation used in this thesis.

In Chapter 2, I review in further detail the traveltime theory of Chapman and Pratt (1992) and the inversion methods of Pratt et al. (1998), as well as the FDFD method of Operto et al. (2009). In Chapter 3, I validate my proposed AWT workflow using a synthetic example in which the starting models of V_{PO}, ϵ and δ are known. In Chapter 4 I apply this technique to crosshole seismic gathers from Western Canada. First presented by Pratt et al. (2008), this crosswell survey was conducted across finely-layered sediments to identify the structures present within local sandstone reservoirs. As these sedimentary layers appear to have a near horizontal dip, I propose a VTI isotropy model derived from anisotropic traveltime tomography. Frequency-domain wave modelling and waveform inversion have previously been conducted on these data by Pratt et al. (2008), however they used the aforementioned 1-D elliptical model. I compare my inversion results to those obtained by Pratt et al. (2008) in order to directly quantify the benefits of properly accounting for velocity anisotropy. Finally, in Chapter 5 I give my concluding thoughts on the success of the proposed AWT workflow and make suggestions for further studies that will advance the work of this thesis.

The implementation of VTI during this study improves the accuracy of the FWI solution for the anisotropic case by removing a significant source of error in the results where the elliptical assumption is invalid, which serves as a technical advancement for modelling anisotropy in FWI studies. Successful results potentially contribute to the development of reservoir models at the Western Canada crosshole site, as well as to the overall understanding of seismic anisotropy in sedimentary basin environments. Furthermore, the application of VTI to this problem promotes the implementation of this anisotropy approach for crosshole seismic studies, and validates FWI as a state-of-the-art technique for solving seismic inversion problems.

Symbol	Description
V_V	Vertical velocity
V_H	Horizontal velocity
ϵ	Thomsen's parameter relating vertical to horizontal velocity
δ	Thomsen's parameter describing rate of change of velocity as a function of angle from the vertical
θ	Angle of the symmetry axis with respect to the vertical
V_{P0}	'Isotropic ' P-wave velocity along the symmetry axis
T	Total traveltime
c	Seismic velocity
dl	Incremental ray length
r	Lateral resolution
λ	Seismic wavelength/'Lamé' parameter
L	propagation distance
S_H	Horizontal shear wave component
S_V	Vertical shear wave component
κ	Bulk modulus
μ	Shear modulus
$\mathbf{u}$	Displacement
ρ	Density
t	Time
τ	Elastic stress tensor
$\mathbf{w}$	Externally applied force per unit volume
c	Fourth-order stiffness tensor
e	Strain tensor
∇	Del operator
P	Hydrostatic pressure field
A	Wave amplitude
s	Source function
G	Forward modelling operator
$\mathbf{m}$	Model parameter vector
$\mathbf{d}$	Data vector
$\delta\mathbf{d}$	Data residuals
E($\mathbf{m}$)	Objective function
$\mathcal{M}$	Model space
$\mathcal{D}$	Data space
$\nabla_m E(\mathbf{m})$	Gradient of the misfit function
$\mathbf{H}$	Hessian matrix
α	Gradient step length
τ	Time-domain damping function
$\mathbf{p}$	Slowness vector
$\mathbf{q}$	Traveltime model parameter matrix
a	Density-normalized elastic tensor
$\mathbf{F}$	Partial derivative matrix of the arrival times

Table 1.1: Summary of all mathematical notation used in this thesis [1 of 2].

Symbol	Description
$\Delta\mathbf{l}$	Ray lengths for each model cell
$\mathbf{C}_d$	Data covariance matrix
$\mathbf{C}_p$	Model covariance matrix
σ_d	Data variance
σ_p	Model variance
ν	Regularization parameter controlling deviation from starting model
β	Regularization parameter controlling anisotropy penalty
ε	Regularization parameter controlling model roughness
$\mathbf{I}$	Identity matrix
$\mathbf{R}$	Roughness matrix
ω	Angular frequency
Ω	Complex-valued frequency
$\mathbf{S}$	Impedance matrix
$\mathbf{u}$	Complex-valued pressure wavefield
$\mathbf{f}$	Source term vector
$\mathbf{J}$	Partial derivative matrix of the wavefield
$\mathbf{v}$	Complex-valued 'back propagated ' pressure wavefield
F_t	Pressure wavefield in the time-domain
p_h	Horizontal slowness
p_z	Vertical slowness
V_{nmo}	NMO velocity
η	Anellipticity parammeter
v	Interval velocity
$\mathbf{k}$	Wavenumber vector
$\mathbf{q}$	Auxilliary wavefield
$V(\theta)$	Phase velocity
$\bar{\theta}$	Phase angle with respect to the symmetry axis
$V_{S_0}^2$	S-wave velocity along the symmetry axis
θ	Phase angle with respect to the z-axis
θ_0	Angle between the symmetry axis and the vertical
$\bar{H}$,$\bar{H}_0$	Differential operators
κ_0	Bulk modulus along the symmetry axis
b	Buoyancy
ξ	1D CPML damping function
γ	Damping coefficient
χ	Damping coefficient
i	Imaginary number
$\mathbf{M}$	Mass term diagonal
$\mathbf{A}$,$\mathbf{B}$,$\mathbf{C}$,$\mathbf{D}$	Stiffness matrices
w_1,w_{m1},w_{m2}	Weighting coefficients

Table 1.2: Summary of all mathematical notation used in this thesis [2 of 2].

Chapter 2

Background Theory and Methodology

In this chapter I first derive the elastic wave equation from fundamental stress-strain relations to show how the equation evolves and changes with the introduction of heterogeneity, source terms, anisotropy etc. Next, I provide a brief overview of the foundational principles of both forward modelling and inversion, which I will then describe in the context of the methods used to perform the works presented in this thesis. Finally, I review in detail the contributions and improvements that I have made to the techniques used by the Seismic Imaging group here at Western University. To do this, I provide a full derivation of the acoustic anisotropic wave equation, as well as a detailed description of the implementation I have adopted for tackling seismic anisotropy in the FWI context.

2.1 The Elastic Wave equation

Whether one is solving for traveltimes or waveforms, a system that accurately describes the behaviour of seismic waves through a given medium is a fundamental requirement. A material is said to display elastic behaviour if its physical properties (shape and structure) are altered by an acting force, but return to their original state once the acting force is removed (for a modern review of linear elasticity, see Sadd, 2009). Elastic materials support both pressure (P-wave) and shear (S_H- and S_V-wave) deformation from propagating seismic waves, which can be related to the bulk (κ) and shear (μ) modulii of the rockmass respectively. The elastic wave equation relates the observed displacement components (u_i, for i= 1,2,3) at a given point within a continuous elastic solid, to the elastic properties of the medium as they are subjected to propagating elastic waves. As u_i is a vector quantity, the equation must be solved for each component of displacement.

Over the past several decades, the elastic wave equation has been implemented in inversion techniques for both real data (Crase et al., 1990; Brossier et al., 2009) and synthetic data (Mora,

1987; Pratt, 1990). Elastic modelling techniques are referred to as 'complete' as they account for all types of seismic waves, as well as wave conversions at kinematic boundaries. As such, solving the system of vector equations requires multiparameter, multicomponent modelling techniques. Furthermore, the model grid size must be set in accordance with the shortest observable wavelength to avoid undersampling and to minimize numerical dispersion. This has significant implications for elastic modelling, as shear wavelengths are smaller than their pressure counterparts ($\lambda_{\text{shear}} \simeq \lambda_{\text{pressure}}/3$), significantly increasing model size and, hence, runtimes as well as memory usage.

Here, I follow the derivation of the elastic wave equation outlined by Aki and Richards (2002) and Tsvankin (2012). The general equation of motion for an elastic, heterogenous medium is

$$\rho \frac{\partial^2 u_i}{\partial t^2} - \frac{\partial \tau_{ij}}{\partial x_j} = w_i, \tag{2.1}$$

where ρ is the density of the medium, $\mathbf{u}$ is the displacement vector (u_1, u_2, u_3), t is the time, τ_{ij} are components of the elastic stress tensor, $\mathbf{x} = (x_1, x_2, x_3)$ is the cartesian coordinate system and $\mathbf{w} = (w_1, w_2, w_3)$ is the externally applied force per unit of volume (Aki and Richards, 2002). Note that in equation (2.1) and from now on, summation over repeated indices is implied (as per the Einstein summation convention). In order to extract information about the elastic properties of a rockmass, we must first relate the displacement, $\mathbf{u}$ to the stresses, τ. In the case of small strain, the generalized Hooke's law states that the stress-strain relationship is approximately linear, such that

$$\tau_{ij} = c_{ijkl} e_{kl}, \tag{2.2}$$

for strains

$$e_{kl} = \frac{1}{2}\left(\frac{\partial u_k}{\partial x_l} + \frac{\partial u_l}{\partial x_k}\right), \tag{2.3}$$

where c_{ijkl} is the fourth-order stiffness tensor for the rockmass and e_{kl} is the strain tensor. We assume that this small strain approximation is appropriate for seismic wave propagation (Tsvankin, 2012). Therefore, by substituting equations (2.2) and (2.3) into equation (2.1), we obtain the anisotropic, heterogenous formulation of the elastic wave equation for general displacement,

$$\rho \frac{\partial^2 u_i}{\partial t^2} - c_{ijkl} \frac{\partial^2 u_k}{\partial x_j \partial x_l} = w_i. \tag{2.4}$$

By setting the force term equal to 0, we then obtain the homogenous form of equation (2.4),

$$\rho \frac{\partial^2 u_i}{\partial t^2} - c_{ijkl} \frac{\partial^2 u_k}{\partial x_j \partial x_l} = 0. \tag{2.5}$$

To resolve the displacement field, we must know the components of c_{ijkl} everywhere in space. This tensor comprises 81 components which vary depending on the symmetric properties of the rockmass. However, inherent symmetries within the tensor ($c_{ijkl} = c_{klij} = c_{jilk}$) reduce the number of individual components to 21 (Aki and Richards, 2002). If we describe the elastic tensor using Voigt's notation for symmetric tensors (Musgrave, 1970), the elastic tensor can now be expressed as a symmetric 6x6 matrix with 21 independent constants:

$$C_{ij} = \begin{pmatrix} C_{11} & C_{12} & C_{13} & C_{14} & C_{15} & C_{16} \\ C_{12} & C_{22} & C_{23} & C_{24} & C_{25} & C_{26} \\ C_{13} & C_{23} & C_{33} & C_{34} & C_{35} & C_{36} \\ C_{14} & C_{24} & C_{34} & C_{44} & C_{45} & C_{46} \\ C_{15} & C_{25} & C_{35} & C_{45} & C_{55} & C_{56} \\ C_{16} & C_{26} & C_{36} & C_{46} & C_{56} & C_{66} \end{pmatrix}. \tag{2.6}$$

For TI media, the elastic tensor in equation (2.6) can be reduced to a total of 5 independent parameters, such that

$$C_{ij} = \begin{pmatrix} C_{11} & C_{11} - 2C_{66} & C_{13} & 0 & 0 & 0 \\ C_{11} - 2C_{66} & C_{11} & C_{13} & 0 & 0 & 0 \\ C_{13} & C_{13} & C_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & C_{55} & 0 & 0 \\ 0 & 0 & 0 & 0 & C_{55} & 0 \\ 0 & 0 & 0 & 0 & 0 & C_{66} \end{pmatrix}, \tag{2.7}$$

which must be rotated if the axis of symmetry is not vertical. Thomsen (1986) showed that, for TI media, several alternative parameters could be derived from the stiffness tensor in equation (2.7):

$$V_{P_0} \equiv \sqrt{c_{33}}\rho; \tag{2.8}$$

$$\epsilon \equiv \frac{c_{11} - c_{33}}{2c_{33}}; \tag{2.9}$$

$$\delta \equiv \frac{(c_{13} + c_{55})^2 - (c_{33} - c_{55})^2}{2c_{33}(c_{33} - c_{55})}, \tag{2.10}$$

where the number of independent parameters for P-wave studies in VTI media can be reduced to 3. This notation gives each of the parameters a physical meaning that is relatable to seismic wave propagation, rather than simply solving for the elastic coefficients. For weak anisotropy, Thomsen (1986) showed that the P-wave phase velocity in TI media obeys the following the relationship:

$$V_P(\theta) = \alpha_0 \left(1 + \delta \sin^2\theta \cos^2\theta + \epsilon \sin^4\theta\right), \tag{2.11}$$

for

$$\alpha_0 = \sqrt{\frac{c_{33}}{\rho}}. \tag{2.12}$$

By setting $\epsilon = \delta$, equation (2.11) reduces to

$$V_P(\theta) = \alpha_0 \left(1 + \epsilon \sin^2\theta\right), \tag{2.13}$$

which perfectly describes the shape of an ellipse for any angle, θ.

For the isotropic (simplest) case, the number of independent parameters in equation (2.6) can be further reduced to 2; the bulk modulus, κ, and the shear modulus, μ such that

$$C_{ij} = \begin{pmatrix} \lambda + 2\mu & \lambda & \lambda & 0 & 0 & 0 \\ \lambda & \lambda + 2\mu & \lambda & 0 & 0 & 0 \\ \lambda & \lambda & \lambda + 2\mu & 0 & 0 & 0 \\ 0 & 0 & 0 & \mu & 0 & 0 \\ 0 & 0 & 0 & 0 & \mu & 0 \\ 0 & 0 & 0 & 0 & 0 & \mu \end{pmatrix}, \tag{2.14}$$

where λ is the 'Lamé' parameter that relates to the bulk modulus, κ such that $\lambda = \kappa + \frac{2}{3}\mu$. Note that here I am using the same notation for the 'Lamé' parameter as for seismic wavelength, as

the λ symbol is universally recognized for both of these parameters.

It is well acknowledged in the literature that reservoir units should be characterized by more general symmetry models. For example, Tsvankin (1997) suggests that reservoirs with complex fracture networks may be represented by the orthorhombic symmetry system. In this case, the number of individual parameters can only be reduced to 9, and the elastic tensor becomes:

$$C_{ij} = \begin{pmatrix} C_{11} & C_{12} & C_{13} & 0 & 0 & 0 \\ C_{12} & C_{22} & C_{23} & 0 & 0 & 0 \\ C_{13} & C_{23} & C_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & C_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & C_{55} & 0 \\ 0 & 0 & 0 & 0 & 0 & C_{66} \end{pmatrix}. \tag{2.15}$$

In this thesis, I assume that all geological media observed in this study are well represented by the TI symmetry system. The more general symmetry models, while useful, increase the parameterization of the problem, and we will see later on that this complicates both the forward modelling and inversion processes. As such, I choose to restrict the number of individual parameters to simplify forward modelling algorithms and improve model convergence by exercising greater control over the number of parameters involved.

2.1.1 The Acoustic Approximation

The acoustic wave equation describes the passage of pressure waves through a continuous fluid, such as sound waves emitted from a speaker or sonar waves traveling underwater, and is a good approximation to the kinematics of P-wave propagation through a medium. The solution of the acoustic wave equation is the temporal and spatial behaviour of the pressure field, which is a scalar quantity. This equation relies on simpler model physics than its elastic counterpart, and thus requires less complex modelling techniques. Acoustic wave modelling is an 'engineering' approach to modelling seismic wavefields, as it oversimplifies the physical model in an attempt to reduce model complexity and simulation runtimes. This has significance for multisource, exploration-scale data that are nontrivial to solve with elastic techniques.

Let us now derive the acoustic wave equation from its elastic equivalent. First, we can recast equation (2.5) in terms of the elastic modulii of an isotropic medium. The equation now takes the form

$$\rho \frac{\partial^2 u_i}{\partial t^2} = (\lambda + \mu) \frac{\partial}{\partial x_i} \left(\frac{\partial u_j}{\partial x_j} \right) + \mu \left(\frac{\partial^2 u_i}{\partial x_k \partial x_k} \right), \tag{2.16}$$

which can be described more compactly by introducing vector operators:

$$\rho \frac{\partial^2 u_i}{\partial t^2} = (\lambda + \mu) \nabla (\nabla \cdot \mathbf{u}) + \mu (\nabla^2 \mathbf{u}), \tag{2.17}$$

where

$$\nabla \equiv \frac{\partial}{\partial x_i} \tag{2.18}$$

(remembering again that summation over all indices is implied). Since shear waves cannot propagate within a fluid, we can set μ=0 in equation (2.17), such that λ is reduced to κ in the first term and the second term is negated. Finally, we divide both sides of equation (2.17) by density and introduce the definition of hydrostatic pressure for a fluid from Hooke's Law. After taking the divergence of both sides, the equation reduces to the source-free, acoustic wave equation for heterogenous, isotropic media:

$$\frac{\partial^2 P}{\partial t^2} = \kappa \frac{\partial}{\partial x_i} \frac{1}{\rho} \frac{\partial}{\partial x_i} P, \tag{2.19}$$

for

$$P \equiv -\kappa \nabla \cdot \mathbf{u}. \tag{2.20}$$

In realistic geological media, heterogenous density is expected as this physical property is know to vary as a function of depth, mineralogy and structure. In order to implement equation (2.19) for heterogenous model cases, however, we must utilize the finite difference schemes described later on in this chapter. This is because a wave equation for heterogenous media cannot be solved analytically.

Let us now assume homogenous density and bulk modulus in equation (2.19) and inject the following 'plane' wave solution:

$$P = As \left(t - \frac{x}{c} \right), \tag{2.21}$$

where A is the amplitude of the wave, c is the velocity, x is the 'x coordinate' and s(t) is the time-dependent source function. The solution is labelled as a plane wave as it only varies in one spatial dimension. Based on equation (2.21), we see that the wave travels in the positive x-direction with a velocity, c by an amount, x/c. By taking the derivatives of P with respect to time and x-direction, it can be shown by simple substitution that equation (2.21) satisfies equation (2.19) provided

$$c = \sqrt{\frac{\kappa}{\rho}}, \tag{2.22}$$

which in turn describes the velocity for pressure wave propagation in an isotropic, homogenous medium. It is important to note that the isotropic acoustic wave equation (2.19) for heterogenous media is the one that is currently implemented within the inversion strategies of Pratt et al. (1998), and has been utilized by Dr. Gerhard Pratt and his students and collaborators over the past two decades. The novel works presented in this thesis are not based on equation (2.19), however I will compare my results in Chapter 4 to an elliptically isotropic solution obtained using this equation. Thus, I considered it important to show the reader how this equation was obtained, as well as how it differs from the anisotropic equivalent derived later on.

2.2 Modelling and Inversion Techniques

The term forward modelling refers to the process of generating the response (**u**) to a given physical property model (**m**) based on the governing physical process or processes (**G**), or:

$$\mathbf{G}(\mathbf{m}) \quad = \quad \mathbf{u}, \tag{2.23}$$

where **m** and **u** are column vectors of lengths n and m respectively. In mathematical terms, we say that the governing physics 'acts' on the model parameters in order to produce the predicted data. Forward modelling is fairly intuitive and is prevalent across many disciplines of science. Let us illustrate this with a fairly straightforward example, such as predicting the horizontal distance that a projectile travels after it is launched off of a surface. To keep things simple, we will assume that such projectiles share a common point of origin (x_{s_0}), the gravitational field (g) is uniform across the surveyed region, and the influence of additional forces (such as wind drag) is negligible. The maximum horizontal distance, commonly referred to as the Range (R) of the projectile obeys

$$R = \frac{v_i^2 \sin(2\theta_i)}{g}, \tag{2.24}$$

where v_i is the velocity at the point of origin. g is the gravitational acceleration and θ_i is the launch angle. In this case, G is the physical model that describes the object's range, v_i, g and θ_i are the model parameters ($\mathbf{m} = m_1, m_2, m_3$) and the travelled horizontal distance is the observed response, $\mathbf{u}$. Experiments can thus be conducted for numerous discrete combinations of model parameters, and the predicted ranges can be tabulated and compared against the observed distances to draw conclusions about the correctness of the model parameters, $\mathbf{m}$.

Inversion is the reverse operation of forward modelling, as we are attempting to recover the physical model properties from the observed data. In this case, we are attempting to estimate the model parameters from the data:

$$\mathbf{m} = \mathbf{G}^{-1}\mathbf{u}, \tag{2.25}$$

where equation (2.25) is often referred to as the 'inverse problem'. Inversion is widely used in geophysics, as the physical properties of the subsurface are often unknown or difficult to measure reliably. In many realistic cases, the 'true' model parameters cannot be obtained as simply as in equation (2.25), as the forward operator is often singular and extremely ill-conditioned (small changes or errors in the inputs greatly impact the output), making $\mathbf{G}^{-1}$ impossible to compute directly. Furthermore, equation (2.23) is an oversimplification of the forward modelling process (as we will see later on, $\mathbf{G}$ requires a more general approach). As such, our primary goal becomes obtaining a set of model parameters that best fit the data. We are forced to relax the equality in equation (2.25), and our result for $\mathbf{m}$ can now only be described as an 'estimate' of the true model.

Once sufficient data have been obtained, the important question becomes: how can one obtain the best-fitting model? The most common approach is to make a starting guess, forward model the data with these starting values and compare the predicted data to those obtained when conducting the experiment. Based on the differences between the predicted and observed data, adjustments can be made to the model values and the entire process can then be repeated until the predicted data match the observed data to within an acceptable degree of error. Figure 2.1 outlines the important steps involved in the inversion process. This iterative approach is the basis for inversion studies and has been employed successfully across numerous disciplines over the past five decades (see for example, Aki et al., 1977).

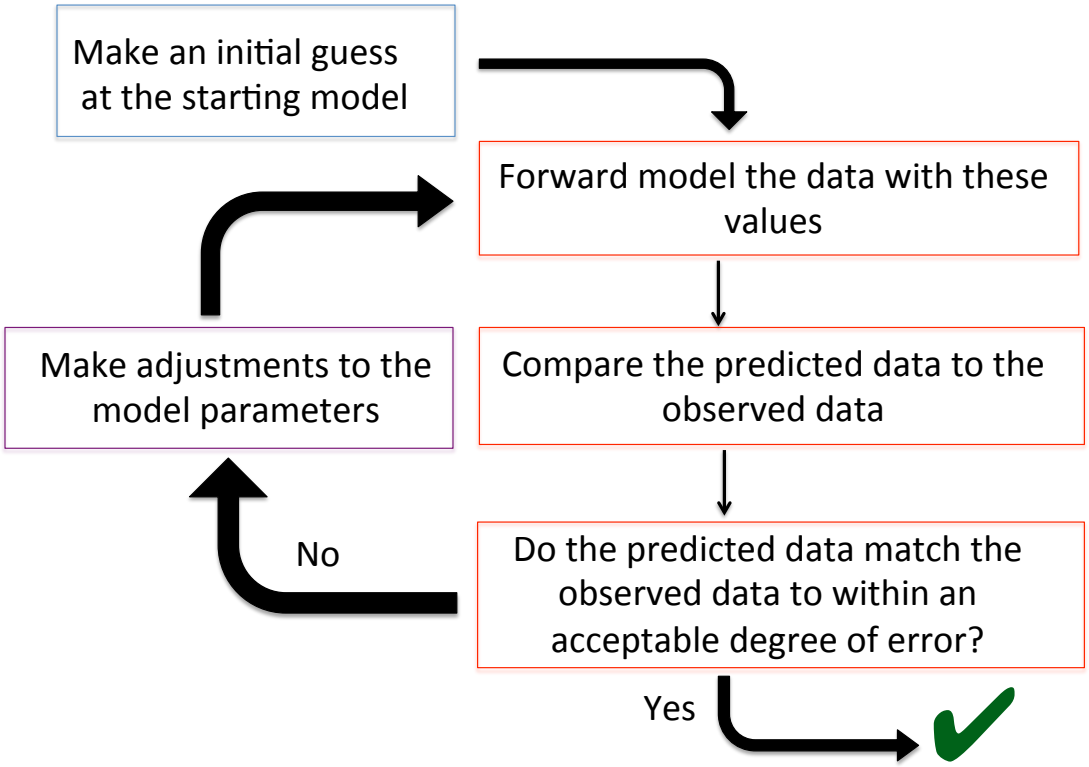


Figure 2.1: A simplified flow chart highlighting each of the important steps involved in the inversion process.

In order to fully comprehend the intricacies of inversion modelling, we must further examine the optimization process. This requires us to think again about the equality in equation (2.25), as well as what exactly it means to 'solve' the inverse problem. Here I follow the derivation and discussion of Pratt (2011). Let us first define the differences between the predicted and observed data. If we conducted our trajectory experiment n times, the experimental data, $\mathbf{d}$, and the predicted data, $\mathbf{u}$, can be expressed as n-dimensional column vectors of observed and predicted ranges, such that

$$\mathbf{d} = \begin{pmatrix} d_1 \\ d_2 \\ d_3 \\ \dots \\ d_n \end{pmatrix}, \ \mathbf{u} = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ \dots \\ u_n \end{pmatrix}. \tag{2.26}$$

The differences, referred to as the data residuals, $\delta\mathbf{d}$, can then be defined as

$$\delta d_i = u_i - d_i \quad (\text{for} \quad i = 1, 2, 3, ..., n). \tag{2.27}$$

The next question one might ask is how exactly can we use the data residuals to alter or improve our model parameters? To answer this, we must first define the 'misfit' or 'objective' function, E(**m**), which is a quantitative measure of how well our current model fits the observed data. Several formulations of this function exist in the literature that consider the role of the data residuals as well as their relative weighting with respect to one another. The simplest formulation, and the one I adopt in the inverse methods described herein, is the least squares or L_2 norm:

$$E(\mathbf{m}) = \frac{1}{2}\delta\mathbf{d}^T\delta\mathbf{d} = \frac{1}{2}\sum_{i=1}^{n}\delta d_i^*\delta d_i, \tag{2.28}$$

where $\delta\mathbf{d}^T$ is the Hermitian transpose and d_i^* is the complex conjugate, introduced to allow a formulation of the misfit functional for complex-valued data. Until this point, our data have been horizontal distances, which are real-valued by design. In FWI, we formulate the problem in the frequency-domain, so the data will be complex-valued. We account for this in equation (2.28), noting that the norm of the data residuals is real-valued, as expected by the definition of E(**m**). We should also note that equation (2.28) neglects the incorporation of any *a priori* statistical information, such as the measure of uncertainty in the data or the model. We can finally answer the philosophical question posed earlier. In order to 'solve' the inverse problem, we seek to find the set of model parameters that minimize the misfit function in equation (2.28) by searching for models that make the data residuals small.

With a clear definition of our objective, we must now discuss the methods by which we plan to minimize the misfit function, E(**m**). To do this, we should first define the model space, $\mathcal{M}$ spanned by the elements of **m**, and the data space, $\mathcal{D}$, spanned by the elements of **d** ($\mathbf{m} \subset \mathcal{M}$, $\mathbf{d} \subset \mathcal{D}$). *Global* inverse methods put little emphasis on the 'topography' of the misfit function, and instead attempt to find the best-fitting model by exploring the full range of models within $\mathcal{M}$. Similarly, *semi-global* inverse methods search a large portion of $\mathcal{M}$, and some emphasis is put on the local topography of the misfit function. Afanasiev et al. (2014) successfully applied the semi-global inverse method of simulated annealing to develop best-fitting 1-D anisotropy models for acoustic frequency-domain waveform tomography, after unsuccessful inversions were obtained using the anisotropy models recovered from traveltime tomography. While robust, true global methods have been largely disregarded in the field of waveform tomography due to the computational costs incurred by multiple forward model runs, as well as due to the enormous size of the model space. We will see later on that the incorporation of anisotropy into the acoustic wave equation greatly increases the size of the model space. Furthermore, for real data applications, FWI relies on the highest frequencies to resolve the finer structure in the

model. This in turn requires a finer grid size, further increasing model size and runtimes. As such, global and semi-global inverse methods will not be considered further in this thesis.

Local inverse methods utilize the topography of the misfit function to develop perturbations about the starting model. These techniques operate under the assumption that the starting model is within an acceptable degree of accuracy, such that the global minimum can be obtained by searching locally about the starting model, $\mathbf{m}$. To do this, we must first apply a Taylor Series expansion to the misfit function by applying an infinitesimal perturbation to the model parameters, $\delta\mathbf{m}$:

$$E(\mathbf{m} + \delta\mathbf{m}) = E(\mathbf{m}) + \delta\mathbf{m}^t \nabla_m E(\mathbf{m}) + \frac{1}{2}\delta\mathbf{m}^t \mathbf{H} \delta\mathbf{m} + O(||\delta\mathbf{m}||^3), \tag{2.29}$$

where $\nabla_m E(\mathbf{m})$ is the gradient of the misfit function and $\mathbf{H}$ is the Hessian, a matrix of second derivatives. If p is the number of model parameters, these terms can be represented mathematically as follows:

$$\nabla_m E = \begin{pmatrix} \frac{\partial E}{\partial m_1} \\ \frac{\partial E}{\partial m_2} \\ \cdots \\ \frac{\partial E}{\partial m_p} \end{pmatrix}, \tag{2.30}$$

and

$$H_{i,j} = \frac{\partial^2 E}{\partial m_i \partial m_j}, \tag{2.31}$$

where $H_{i,j}$ is the element of the Hessian in the ith row and jth column, and $i,j \subset \{1, 2, ..., p\}$ such that the Hessian is an pxp matrix. Since we are trying to find the model perturbation, $\delta\mathbf{m}$, that minimizes the misfit function, we differentiate equation (2.29) with respect to $\delta\mathbf{m}$ and set the result to 0 to give

$$\mathbf{H}\delta\mathbf{m} = -\nabla_m E(\mathbf{m}). \tag{2.32}$$

To solve for the model perturbation, $\delta\mathbf{m}$, one could apply the Newton algorithm, which is to simply invert the Hessian in equation (2.32) to give

$$\delta \mathbf{m} = -\mathbf{H}^{-1} \nabla_m E(\mathbf{m}). \tag{2.33}$$

The Newton method is simple to interpret and easy to implement from a coding standpoint, however there are several structural downfalls which must be identified. First, we should recall the definition of the Hessian from equation (2.31), which shows that the Hessian matrix requires one row and one column for all m model parameters. Even for small, isotropic problems, the size of the Hessian is very large, requiring the allocation of large portions of storage memory. We must remember also that practical cases are constrained by computational costs. Even if we disregard the size of the Hessian altogether, there is yet another potential problem. For most real data cases, a lack of resolution for one or more sets of parameters is common, which introduces zero eigenvalues (interpreted as directions of zero curvature) into the Hessian matrix. The presence of zero curvature in any direction would make it a singular matrix, therefore it cannot be mathematically inverted ($\mathbf{H}^{-1}$ is not possible).

An alternative to the Newton method that is both practical and computationally efficient is to ignore the Hessian altogether and calculate the model update from the gradient directly. This method, termed the 'steepest descent' or gradient algorithm, is based on the following principle: If the gradient, $\nabla_m E(\mathbf{m})$, shows where and in which direction the misfit function is increasing most rapidly, why not simply find the direction that points 'downhill' i.e. the direction in which the objective function is decreasing most rapidly? Thus, equation (2.33) instead becomes:

$$\delta \mathbf{m} = -\alpha \nabla_m E(\mathbf{m}), \tag{2.34}$$

where α is a positive scalar chosen to minimize the least squares norm in the direction opposite to the gradient of E($\mathbf{m}$) (Pratt et al., 1998). It is common to implement the gradient algorithm iteratively, whereby the model perturbation is calculated several times and repeatedly applied to the most recent model:

$$\mathbf{m}^{k+1} = \mathbf{m}^k - \alpha^k \nabla_m E^k, \tag{2.35}$$

where k is the current iteration number. In this approach, the current 'best' model, $\mathbf{m}^k$, is perturbed by the model update, $-\alpha^k \nabla_m E^k$, where α is chosen to ensure that the model descends to the point where the misfit function is optimally reduced. Note that both the step length and

the gradient must be re-calculated at each iteration.

The gradient algorithm is ideal for its computational efficiency, as it avoids computation or inversion of the Hessian and simply uses the same gradient that was required for calculating the model update in the Newton method. One obvious pitfall of this method, however, is that we are assuming that the absolute minimum lies 'downhill' from our starting model. This would not be a concern if E(**m**) had a single definite minimum, whereby any local inverse method would eventually arrive at the global minimum. In most realistic cases, however, E(**m**) is far more complicated and may contain multiple local minima. Therefore, when implementing local inverse methods, one runs the risk of instead converging to a local minimum. This illustrates the stringent requirement that the starting model be 'accurate enough' to avoid converging to a local minimum. In the following subchapters we will see that equation (2.35) has different but specific meanings for both traveltime tomography and FWI, and the question of starting model accuracy will be addressed for each of these cases respectively.

2.2.1 Traveltime Tomography

Let us now return to the traveltime tomography method for weakly anisotropic, inhomogeneous media proposed by Chapman and Pratt (1992) and Pratt and Chapman (1992). Here I will review only the major elements from these works that pertain to P-wave propagation. If we assume that the perturbations to the raypath can be ignored to first-order accuracy, the traveltime perturbation based on equation (1.1) is simply:

$$\delta T = \int_{ray} \frac{1}{\delta c(\mathbf{x})} dl, \tag{2.36}$$

which is linear for small perturbations that do not change the path of the ray (Cervený, 1972, 1982; Cervený and Jech, 1982). For isotropic media, equations (1.1) and (2.36) are relatively straight-forward, but for anisotropic media the application is more complicated as the group slowness is now a function of both the media properties as well as the direction of propagation. Pratt and Chapman (1992) showed that the traveltime perturbation, δT, for a small variation in background elastic properties can be written as:

$$\delta T = -\frac{1}{2} \int_{ray} \frac{1}{V_{P_0}^3} \hat{p}_i \hat{p}_l \hat{p}_j \hat{p}_k \delta a_{ijkl} ds, \tag{2.37}$$

for

$$a_{ijkl} = \frac{c_{ijkl}}{\rho}, \tag{2.38}$$

where $\hat{\mathbf{p}} = (\hat{p_1}, \hat{p_2}, \hat{p_3})$ represents the slowness unit vector in an undisturbed isotropic medium, V_{P_0} is the isotropic velocity and δa_{ijkl} is a small perturbation to the density-normalized elastic tensor. Note that in equation (2.37), the traveltime perturbation is a function of perturbations to the components of the elastic tensor. This gives flexibility in that the anisotropy formulation can be completely general. Since we are only interested in 2-D traveltime data, the out-of-plane terms in equation (2.37) can be ignored ($\hat{p_2} = 0$), and the number of independent terms can be reduced to 5:

$$\delta T = -\frac{1}{2V_{P_0}^3} \int_{ray} \left(\hat{p}_1^4 \delta q_1 - \hat{p}_1^3 \hat{p}_3 \delta q_2 - \hat{p}_1^2 \hat{p}_3^2 \delta q_3 - \hat{p}_1 \hat{p}_3^3 \delta q_4 - \hat{p}_3^4 \delta q_5 \right) ds, \tag{2.39}$$

where

$$\begin{aligned} &\delta q_1 = \delta a_{1111},\ \delta q_2 = 4\delta a_{1131}, \\ &\delta q_3 = 2\delta a_{1133} + 4\delta a_{3131},\ \delta q_4 = 4\delta a_{3331}, \\ &\delta q_5 = \delta a_{3333}. \end{aligned} \tag{2.40}$$

Let us now discuss the concepts of discretization and data coverage in the context of traveltime tomography. We hope to acquire a dense coverage of the survey region through the selection of source and receiver locations. These are ideally chosen such that, when the paths between all source-receiver pairs are traced by rays, there is sufficient ray coverage in each homogenous 'cell' of the tomography model. However, there exists both geometrical and physical limitations to this idealization. Recall that the common geometrical setup for a cross-well experiment employs vertical source and receiver boreholes. Therefore, even in the case of a perfectly homogenous region, there will exist regions near the top and bottom of the model where ray coverage is sparse. Furthermore, the governing physics of rays is that they will bend at contrasts in seismic velocity. Therefore in more realistic heterogenous models, there may exist regions, termed 'shadow zones', in which ray coverage is sparse due to this inherent bending of the rays. Nevertheless, we define our data, $\mathbf{d}$, as a vector of values containing the traveltimes, $\mathbf{T}$, for the path between each source-receiver pair.

After discretization, equation (2.39) becomes:

$$\delta T_k = F_{kij}\delta q_{ij}, \tag{2.41}$$

where δT_k is the traveltime perturbation for the kth ray, q_{ij} is the perturbation of the jth elastic parameter (j=1, ..., 5) in the ith model gridpoint, and F_{kij} is the partial derivative of the arrival time of the kth ray to changes in δq_{ij} (Pratt and Chapman, 1992). In this approach, the integral in equation (2.39) is replaced with a summation over all elements of $\Delta\mathbf{l}$, where Δl_{ki} is the length of the kth raypath that travelled through the ith cell, such that:

$$F_{kij} = \frac{\partial T_k}{\partial q_{ij}} = -\frac{1}{2V_{P_0}^3}\Delta l_{ki}\hat{p}_1^{5-j}\hat{p}_3^{j-1}, \text{for } j = 1, ..., 5, \tag{2.42}$$

and equation (2.42) becomes:

$$\delta\mathbf{T} = \mathbf{F}\delta\mathbf{q}. \tag{2.43}$$

We can now draw some comparisons between equation (2.43) and our generalized equation for forward modelling shown earlier ($\mathbf{G}(\mathbf{m}) = \mathbf{u}$) as equation (2.23). We see that $\mathbf{F}\mapsto\mathbf{G}$ operates on the elastic parameters, $\mathbf{q}\mapsto\mathbf{m}$, to give the traveltime perturbation, $\delta\mathbf{T}\mapsto\mathbf{d}$, where $\mapsto$ denotes a mapping function. Until this point, we have formulated the forward problem under the basis that the isotropic background velocity, V_{P_0}, will remain unchanged throughout. This is a poor assumption as without significant *a priori* knowledge of the survey region, it is unreasonable to assume that our starting guess for the isotropic velocity is exactly correct. To account for this, Chapman and Pratt (1992) added a sixth term, δq_0, such that:

$$\delta q_0 = \delta V_{P_0}^2, \tag{2.44}$$

where δV_{P_0} is the perturbation to the isotropic background velocity. Before this term is introduced, equations (2.41) and (2.42) form a system of linear equations for anisotropic tomography. With the addition of δq_0, however, the system becomes non-linear as the model update $\mathbf{q}$ may now contain an update to the isotropic background velocity, and Δl_{ki} in equation (2.42) is dependant on the raypaths. These non-linearities are handled by iteratively recomputing the distribution of V_{P_0}, followed by re-linearizing equations (2.41) and (2.42) by retracing the new rays (Pratt and Chapman, 1992).

The goal of anisotropic traveltime tomography is to develop best-fitting models of velocity

(in this case, the P-wave velocity along the symmetry axis, V_{P_0}), as well as Thomsen's parameters (i.e., models of ϵ and δ) for a set of observed or computed traveltimes. We are therefore interested in solving the system of equations for the perturbations to the elastic parameters, $\delta\mathbf{q}$. It is important to clarify, however, that the elastic parameters in $\mathbf{q}$ are not the same parameters that affect P-wave velocity in TI media shown in equation (2.7), nor the ones shown in Thomsen's notation (equations (2.8), (2.9) and (2.10)). This is due to the generalized P-wave formulation of the forward problem that considers all possible symmetry systems. Instead, the 'best-fitting' TI parameters (V_{P_0}, ϵ, δ and θ) are estimated from the solution of the tomography problem ($\delta\mathbf{q}$) by a coordinate transform, as described in Appendix E of Chapman and Pratt (1992).

Recalling our earlier discussion of inverse theory, we must first define a 'level of misfit' for the data generated by our starting models through the definition of the objective or misfit function, E($\mathbf{m}$). To do this, we must return to the least-squares formulation shown in equation (2.28). For traveltime tomography, this formulation of the L_2 norm is oversimplified, as it does not take into account any *a priori* information that can be used to improve model convergence and parameter resolution. In fact, anisotropic traveltime tomography is extremely ill-posed, the reasons for which are two-fold. Firstly, each value in $\mathbf{d}$ is calculated from a raypath that will only pass through a very small subset of the model cells. Secondly, for TI media, the parameter-space is large as there are six model parameters for each cell. As such, the Fréchet matrix, $\mathbf{F}$, is both sparse and extremely ill-conditioned (Pratt and Chapman, 1992). Thus, our inverse method must contain some criteria for choosing the components of the null space that materialize in the solution, the incorporation of which serves as a form of matrix regularization for ill-posed inverse problems (Tikhonov and Arsenin, 1977).

As described by Pratt and Chapman (1992), any combination of the following external information can be incorporated to aid the traveltime tomography process:

1. Adequate starting models
2. An estimate of the magnitude and degree of anisotropy
3. The expectation that the elastic properties vary smoothly

In the simplest of cases, a homogenous background velocity can be chosen as a starting model. However, it is common practice to use external information, such as borehole sonic logs for the crosshole case, or Normal Moveout (NMO) Analyses for the surface reflection case, to obtain a starting model that might be closer to the true model. Regarding anisotropy, it is difficult to quantify the exact magnitude of the expected anisotropy parameters for many real data cases. It is instead more common to anticipate the type of anisotropy. For example, if the

region is dominated by horizontal, unfractured sedimentary layers, one could anticipate VTI symmetry. Finally, the assumption that the true elastic models are smooth is difficult to justify for complex geological media, as we may expect sharp velocity (and anisotropy) contrasts along faults, alteration zones, or within finely-laminated sedimentary formations. Nevertheless, the success of traveltime tomography is reliant on the assumption of model smoothness, which may seem counterintuitive as the method itself is based on a high-frequency approximation to seismic wave propagation. Cervený (2005) states that the high frequency approximation governing ray theory is valid for cases where the heterogeneities are significantly larger than the seismic wavelength. For complex media containing strong, rapidly varying heterogeneities, such as the sedimentary formations analyzed in Chapter 4, this assumption breaks down. We will see later on, however, that our methodology relies on FWI to provide the short wavenumber updates necessary to image finer structure. As such, the goal for traveltime tomography is to find a long wavelength model that is accurate enough to serve as a starting point for FWI.

To incorporate these *a priori* information, let us follow the stochastic formulation of the inverse problem presented by Tarantola (1987). Here, the solution of the inverse problem is defined as a probability distribution over all possible models, $\delta\mathbf{q}$, within the model space, $\mathcal{M}$:

$$P\left(\delta\mathbf{q} \subseteq \mathcal{M}\right) = \exp\left(-\frac{1}{2}E(\delta\mathbf{q})\right), \tag{2.45}$$

where the least-squares objective functional, $E(\delta\mathbf{q})$, can be defined as follows:

$$E(\delta\mathbf{q}) = \left(\mathbf{F}\delta\mathbf{q} - \delta\mathbf{T}\right)^T \mathbf{C}_d^{-1}\left(\mathbf{F}\delta\mathbf{q} - \delta\mathbf{T}\right) + \delta\mathbf{q}^T\mathbf{C}_p^{-1}\delta\mathbf{q}. \tag{2.46}$$

Let us now directly compare the formulation above to the simple objective function defined in equation (2.28). The major difference is that, in equation (2.46), we have introduced covariance matrices, $\mathbf{C}_d$ and $\mathbf{C}_p$, to weight the relative contributions of the data residuals (the differences between the predicted and observed data) and the parameter perturbations (the differences between the estimated parameters and those obtained using *a priori* information) within the objective function, $E(\delta\mathbf{q})$ (Pratt and Chapman, 1992). Only if we assume that the data have uniform variance and are uncorrelated ($\mathbf{C}_d = \sigma_d^2\mathbf{I}$, where $\mathbf{I}$ is an identity matrix), and that there is no *a priori* information to constrain our parameter estimates (there are infinite, *a priori* model variances), can equation (2.46) be reduced to equation (2.28).

The stochastic formulation assumes a Gaussian probability distribution over the data and model spaces, such that each solution to the tomography problem, $\delta\mathbf{q}$, can be assigned a unique probability value. In this sense, we are saying that the 'solution' to the inverse problem is to

find the *most probable* model given the data and the prior model. The assumption of Gaussian statistics is difficult to justify, however we accept this assumption on the basis that it is an effective statistical model that describes this problem (Pratt and Chapman, 1992).

As earlier, we now differentiate with respect to the model perturbation, $\delta\mathbf{q}$, and set the solution of equation (2.46) to 0. If we also assume the data are uncorrelated and have uniform variance (as described above), equation (2.46) becomes:

$$\left(\mathbf{F}^T\mathbf{F} + \sigma_d^2\mathbf{C}_p^{-1}\right)\delta\mathbf{q} = \mathbf{F}^T\delta\mathbf{T}, \tag{2.47}$$

$$\therefore \delta\mathbf{q} = \left(\mathbf{F}^T\mathbf{F} + \sigma_d^2\mathbf{C}_p^{-1}\right)^{-1}\mathbf{F}^T\delta\mathbf{T}. \tag{2.48}$$

Let us carefully examine the structure of equation (2.48). If we compare it to our preliminary solution for calculating the model update shown in equation (2.33), we see that the expression inside the brackets is the Hessian ($\mathbf{F}^T\mathbf{F}$), with the addition of the term for model covariance, $\sigma_d^2\mathbf{C}_p^{-1}$. Let us recall that the inverse of the Hessian does not exist in the case of a singular matrix, which we have already addressed to be a common occurrence in the traveltime tomography problem. Therefore, we rely on the finite-valued $\mathbf{C}_p$ to obtain a solution, and it is this term that we can modify in order to incorporate external information.

I will now introduce a suite of regularization parameters, β, ν, ε and explain their role within the regularization process. In Chapters 3 and 4, we will see how the selection of these parameters affects the resulting tomographic models. Returning to our *a priori* assumptions, let us first address the question of starting model accuracy. If we assume that each of the model parameters are uncorrelated and have an uncertain variance ($\mathbf{C}_p = \sigma_p^2\mathbf{I}$), equation (2.47) is reduced to the damped least squares formulation, and

$$\delta\mathbf{q} = \left(\mathbf{F}^T\mathbf{F} + \nu^2\mathbf{I}\right)^{-1}\mathbf{F}^T\delta\mathbf{T}, \tag{2.49}$$

for

$$\nu = \frac{\sigma_d}{\sigma_p}. \tag{2.50}$$

Let us briefly examine the role of ν. For large values of ν, $\sigma_d > \sigma_p$ and the inversion will 'prefer' solutions that are similar to the starting models, at the expense of trying to more accurately fit the data. For small values of ν, $\sigma_d < \sigma_p$ we are effectively trusting that our data are reliable, so the inversion will search for best-fitting models that may stray considerably from

the starting values.

Next, we must answer the question of variance between the isotropic velocity and the anisotropic parameters. Recalling that we have defined q_0 to represent isotropic velocity, and q_j, j={1,...,5} to represent the anisotropic parameters, we can modify $\mathbf{C}_p$ such that:

$$\delta\mathbf{q} = \left(\mathbf{F}^T\mathbf{F} + \nu^2\mathbf{I} + \beta^2\mathbf{I}'\right)^{-1}\mathbf{F}^T\delta\mathbf{T}, \tag{2.51}$$

where

$$\mathbf{C}_p = \sigma_p^2\frac{\nu^2}{\nu^2+\beta^2}\text{for } q_j, j \neq 0, \tag{2.52}$$

and $\mathbf{I}'$ is a modified identity matrix in which the diagonal elements corresponding to q_0 are set to zero (Pratt and Chapman, 1992). Again, we pause briefly to examine the role of β. We see that for large values of β, any perturbations to the anisotropic values will greatly impact the misfit function, $E(\delta\mathbf{q})$, thereby reducing the introduction of anisotropy into the model. For small values of β, anisotropy can be introduced into the model without significant increases in $E(\delta\mathbf{q})$. Thus, β can be thought of as an 'anisotropy penalty', as increasing β forces the inversion algorithm towards a purely isotropic result. Since the geological media of interest to this thesis are known to be anisotropic, we will implement small values of β in our inversion strategies throughout.

The third, and arguably most significant parameter of interest is that which imposes model smoothness, ε (note here the subtle difference in notation when compared to the anisotropy parameter ϵ). Recalling the theoretical limitations on the resolution of traveltime tomography (equation (1.3)), as well as the high frequency approximation that models seismic energy along infinitely thin raypaths, we are limited to models that only contain the long wavelength features of the data. For regions with strong, rapidly varying heterogeneities, we will rely on subsequent FWI to delineate the high wavenumber features. As $\mathbf{C}_p$ effectively serves as a smoothing operator incorporating smoothness constraints, $\mathbf{C}_p^{-1}$ will act as a 'roughness penalty' for the tomographic solution. For the traveltime tomography results presented in Chapters 3 and 4, we will see that smoothing is imposed at the expense of data fit, (E($\delta\mathbf{q}$)). To do this, equation (2.52) is further modified to give:

$$\delta\mathbf{q} = \left(\mathbf{F}^T\mathbf{F} + \nu^2\mathbf{I} + \beta^2\mathbf{I}' + \varepsilon^2\mathbf{R}^T\mathbf{R}\right)^{-1}\mathbf{F}^T\delta\mathbf{T}, \tag{2.53}$$

where $\mathbf{R}$ is a discrete, sparse, difference operator referred to as the roughness operator, and ε is a third regularization parameter introduced to give weighting to the various components of $\mathbf{R}$. Specifically, we introduce derivatives of the solution, $\delta\mathbf{q}$, to the objective function to compute the roughness, as direct incorporation of $\mathbf{C}_p^{-1}$ would reduce the sparseness of the matrix we need to invert (Constable et al., 1987; Pratt and Chapman, 1992). For this thesis, the implementation includes three separate roughness matrices, $\mathbf{R}_{xx}$, $\mathbf{R}_{zz}$, and $\mathbf{R}_{xz}$. The first two are finite difference operators applied to the gradient in the x- and -z directions respectively, and $\mathbf{R}_{xz}$ is a 2-D Laplacian finite difference operator that describes the curvature in the model (Pratt et al., 1993).

Finally, we can combine all of the regularization terms described in this section to form the system of equations:

$$\delta\mathbf{q} = \left(\hat{\mathbf{F}}^T\hat{\mathbf{F}}\right)^{-1}\hat{\mathbf{F}}^T\delta\hat{\mathbf{T}}, \tag{2.54}$$

where we introduce the augmented forms of the Fréchet matrix and data vector:

$$\hat{\mathbf{F}} = \begin{pmatrix} \mathbf{F} \\ \varepsilon\mathbf{R}_{xx} \\ \varepsilon\mathbf{R}_{zz} \\ \varepsilon\mathbf{R}_{xz} \\ \nu^2\mathbf{I} \\ \beta^2\mathbf{I}' \end{pmatrix}, \tag{2.55}$$

and

$$\delta\hat{\mathbf{T}} = \begin{pmatrix} \delta\mathbf{T} \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \tag{2.56}$$

such that the traveltime perturbation vector, $\delta\hat{\mathbf{T}}$, has been augmented with the appropriate number of zeros. Note here that the augmented matrix $\hat{\mathbf{F}}$ is sparse by design. Pratt and Chapman (1992) and Pratt et al. (1993) solved the above system using the LSQR solver first proposed by Paige and Saunders (1982). In this approach the model update, $\delta\mathbf{q}$, is calculated using a

conjugate gradient algorithm that is analogous to the one shown in equation (2.35), as $\mathbf{F}^T\mathbf{F}$ is large, and is therefore quite expensive to compute. We adopt this strategy as well, as the LSQR formulation allows us to perform the inversions for several combinations of the regularization parameters (β, ν and ε) at a relatively modest computational cost.

One final comment must be made regarding the accuracy in the models produced: The results of traveltime tomography are extremely reliant on the choice of regularization parameters, as well as the choice of starting models, so deciding when an acceptable result has been obtained is problematic. Pratt and Chapman (1992) suggest an approach where several initial tests are performed using unique combinations of β, ν and ε. As the testing proceeds, the values of the regularization parameters are 'tightened' and an 'acceptable' result can be defined as one which looks geologically plausible, yet still provides an accurate fit of the data to within the estimated traveltime picking error. If these models will be used as the starting point for FWI, however, we will see in the next section that a more quantifiable level of starting model accuracy is defined in order to overcome the cycle-skipping phenomena (Sirgue and Pratt, 2004). Between inversion passes, Pratt and Chapman (1992) also suggest retracing the rays, as changes to the velocity and anisotropy parameters determine the raypaths themselves.

2.2.2 Full Waveform Inversion

With the methods and goals of traveltime tomography clearly defined, let us now turn our attention to FWI, and how its development over the years has made it an industry frontrunner for developing high-resolution velocity models for complex geological media. Lailly (1983) and Tarantola (1984) showed that the steepest descent direction of the inverse problem could be determined for the acoustic wave equation without explicit calculation of the partial derivatives themselves, an improvement to the computational cost that was critical in times before modern memory storage and matrix solving capacities. In this approach, the gradient was calculated by 'back propagating' the data residuals from the receiver location, and cross-correlating the result with the forward propagated wavefield from the source. Several numerical implementations of this technique exist throughout the literature, such as the time-domain formulation suggested by Gauthier et al. (1986), and the frequency-domain formulation suggested by Pratt (1990).

Based on the aforementioned problems regarding matrix size, the gradient algorithm was the preferred optimization choice in early FWI studies. This localized inversion method was used in conjunction with finite-differencing modelling (FDM) for several crosshole field data cases (see, for example, the time-domain approach of Zhou et al., 1995). Though other numerical modelling techniques have been proposed for wave equation modelling, such as the Spectral Element Method (SEM) described by Komatitsch et al. (1999), I rely solely on FDM

methods in this thesis on the basis of their superior efficiency. Marfurt (1984) showed that, for multisource modelling, frequency-domain FDM methods were most appropriate, which led to further refinements of this approach by Jo et al. (1996) and Stekl and Pratt (1998). With regards to time-domain vs. frequency-domain implementations, let us consider that the wavefield at every point in the model must be computed and stored for every time step (δt_i, for i=1,...,N_t) in the time-domain, versus every modelled frequency (f_i, for i=1,...,N_f) in the frequency-domain. In most cases, especially when modelling in reduced time, or when inverting a limited number of frequencies, $N_f << N_t$.

Pratt (1990) and Pratt et al. (1998) proposed a matrix formulation of the seismic waveform inversion problem that is based on a discrete frequency-domain FDM method, which incorporates the work of Marfurt (1984), Jo et al. (1996) and Stekl and Pratt (1998), and utilizes Lailly (1983)'s computational improvements for calculating the gradient function. In this approach, the resulting wavefield can be computed efficiently for each source in Fourier space, and the resulting numerical system can be represented as follows:

$$\mathbf{S}(\omega, \mathbf{m})\, \mathbf{u}(\omega) \quad = \quad \mathbf{f}(\omega), \tag{2.57}$$

where ω is the angular frequency, $\mathbf{u}$ is the complex-valued pressure wavefield, $\mathbf{f}$ contains the source terms and $\mathbf{S}$ is the impedance matrix that represents the 'wave physics', which in this case contains the Helmholtz operator ($\nabla^2 - \omega^2/v^2$) for the pressure wavefield (Pratt, 1990; Marfurt, 1984). Note that in the Fourier domain, ω itself is real-valued. Shin and Cha (2009) proposed an alternative formulation to equation (2.57) they termed the 'Laplace-Fourier domain' for waveform inversion. In this approach, ω is replaced by the complex-valued frequency, Ω, such that:

$$\Omega \quad = \quad \omega + \frac{i}{\tau}, \tag{2.58}$$

where τ is a real-valued characteristic delay time (Phinney, 1965; Pratt, 1990; Kamei et al., 2014) and $\mathbf{u}(\Omega)$ is called the 'Laplace-Fourier domain wavefield' (Shin and Cha, 2009). The value of τ is typically governed by the total modelled time and acts as a data preconditioner for the arrivals (Brenders and Pratt, 2007), whereby a small value of τ would place more emphasis on the earlier arrivals and surpress the later arrivals. As we will see later on, this will be especially useful in the anisotropic case, where the predicted data are plagued by late-arriving artifacts.

An advantage of the matrix formulation in equation (2.57) is that it is valid for any wave

equation or parameterization, which has extreme significance for the works in this thesis. In Chapter 4, we will use this formulation to solve the 2-D acoustic isotropic wave equation (equation (2.19)) for Horizontal Velocity (V_H), and in Chapters 3 and 4 we will use this formulation to solve the 2-D acoustic anisotropic wave equation for TI media (equations (2.109) and (2.110), derived in the next section of this chapter) for Vertical Velocity (V_V). To adapt this technique to the different wave equations and parameters, one only needs to adjust the physics described within the impedance matrix, $\mathbf{S}$. There are some subtle differences to the other terms of equation (2.57) for the anisotropic case as well, but more on this later.

In FWI, we are attempting to match the waveform amplitude and phase, and as we shall see we are interested in modelling the displacement field at any point in the model. Re-arranging equation (2.57) in the Laplace-Fourier domain:

$$\mathbf{u}(\Omega) \quad = \quad \mathbf{S}^{-1}(\Omega)\,\mathbf{f}(\Omega), \tag{2.59}$$

which one might erroneously relate to the inverse problem (equation 2.25). Equation (2.59) is the solution to the forward problem (equation 2.23), recalling the reformulation posed by Lailly (1983) and Tarantola (1984). Here, $\mathbf{S}(\mathbf{m})^{-1}\mathbf{f}$ is anlagous to $\mathbf{G}(\mathbf{m})$ in equation (2.23) and represents the forward propagated wavefield, and $\mathbf{u}$ is the displacement predicted at the receiver position (Pratt et al., 1998). For computational efficiency, $\mathbf{S}^{-1}$ is never explicitly computed. Rather, equation (2.59) is factorized into upper and lower triangular terms by the well known LU decomposition technique (Press et al., 1992). In this approach, the matrix factors can be re-used for new source terms, $\mathbf{f}$, allowing for efficient calculation of both real and 'virtual' sources throughout the iterative inversion process (Pratt, 1990; Pratt et al., 1998).

With the forward model for FWI clearly defined, we turn our attention once again to the inverse problem, and trying to find the best fitting model that predicts the data. Letting n equal the number of receiver nodes and m equal the number of model parameters, we recall our definitions of the data residuals, $\delta\mathbf{d}$, (equation 2.27), the misfit function, $\mathbf{E}(\mathbf{m})$, (equation 2.28), and the gradient of the misfit function,$\nabla_m E$, (equation 2.29). Here we start with the simplified objective function ($\mathbf{C}_d = \sigma_d^2\mathbf{I}$, and we assume infinite, *a priori* model variances). We purposely omit these to avoid obscuring the simplicity of the algorithm (Tarantola, 1987; Pratt et al., 1998). Let us recast equation (2.28) as follows:

$$\nabla_m E = \Re\left\{\mathbf{J}^t \delta\mathbf{d}^\star\right\}, \tag{2.60}$$

where $\mathbf{J}^t$ is the tranpose of the n x m complex-valued Fréchet derivative matrix such that:

$$\mathbf{J}_{ij} = \frac{\partial u_i}{\partial m_j}, \tag{2.61}$$

noting that the real-part in equation (2.60) is taken to ensure that the gradient itself is real, which is expected by its definition.

Let us add some physical interpretations to equation (2.60) in the context of seismic wave equation modelling. If we define the estimated wavefield that is produced by propagating a source into the current model by $\mathbf{u}_{est}$, and the observed wavefield as measured at each receiver location by $\mathbf{d}$, we see that the data residuals is simply the difference between these two wavefields ($\mathbf{u}_{est}$ - $\mathbf{d}$), and equation (2.60) becomes:

$$\nabla_m E(\mathbf{m}) = \Re \left\{ \left(\frac{\partial \mathbf{u}_{est}}{\partial \mathbf{m}} \right) (\mathbf{u}_{est} - \mathbf{d})^\star \right\}. \tag{2.62}$$

While this equation might appear simple in its interpretation, we must be cautious and recall the steps that were taken to define the gradient function in equation (2.28). For one, it is derived for many small model perturbations, $\delta\mathbf{m}$, therefore its use with an iterative model update approach, such as the one described by equation (2.35), is expensive. This is because the forward modelled wavefield ($\mathbf{u}_{est}$) must be calculated for each set of model parameters (equation (2.59)), therefore the first term in equation (2.62) is prohibitive to the overall efficiency of the inversion problem. As we will see in Chapters 3 and 4, it is not uncommon in FWI to perform multiple iterations of FWI for different combinations of frequencies, offsets, regularization parameters etc. Therefore, the forward modelled wavefield would need to be calculated many times, which is not ideal.

Fortunately, Pratt et al. (1998) provided an alternate, efficient approach for calculating the gradient of the seismic waveform problem that does not require the explicit calculation of $\mathbf{J}$ at each gradient step[1]. First, let us recast equation (2.60) above as follows

$$\nabla_m E = \Re \left\{ \hat{\mathbf{J}}^t \delta \hat{\mathbf{d}}^\star \right\}, \tag{2.63}$$

where we introduce the augmented forms, $\hat{\mathbf{J}}^t$, $\hat{\mathbf{d}}$, recalling that the lengths of the unaltered terms were n x m, and n x 1 respectively. Now, we introduce the dimensional parameter, l, equal to the number of nodes within the model. We now augment $\mathbf{J}^t$ to $\hat{\mathbf{J}}^t$ by adding partial derivatives

[1]Note that this exact step was the 'trick' discovered by the works of Lailly (1983) and Tarantola (1984) which made FWI computationally feasible.

for all node points (u_i, for i= 1, ..., l and $l > n$) such that $\mathbf{J}$ now becomes an l x m matrix. To augment the data residual matrix, we simply add (l - m) zero terms after the nth term in $\mathbf{d}$ to give an l x 1 matrix.

Now, let us return to the general formulation posed in equations (2.57) and (2.59). If we take the derivative of both sides with respect to the ith model parameter, m_i, and introduce the terminology for the estimated and observed wavefields, we get:

$$\mathbf{S}\frac{\partial \mathbf{u}_{est}}{\partial m_i} = \mathbf{f}^{(i)}, \tag{2.64}$$

or

$$\frac{\partial \mathbf{u}_{est}}{\partial m_i} = \mathbf{S}^{-1}\mathbf{f}^{(i)}, \tag{2.65}$$

where we define the ith 'virtual source' as:

$$\mathbf{f}^{(i)} = -\frac{\partial \mathbf{S}}{\partial m_i}, \tag{2.66}$$

which is itself, an l x l matrix (Pratt et al., 1998). A quick comparison between equation (2.65) and the one used to derive equation (2.62) makes it clear that the partial derivative wavefields are now determined locally for each model parameter by the propagation of the virtual sources throughout the model. In this formulation, the virtual source represents the interaction of the predicted wavefield with the nodal parameter or parameters (Pratt et al., 1998).

If we substitute equation (2.65) into equation (2.63), the gradient becomes[2]:

$$\nabla_m E(\mathbf{m}) = \Re\left\{\left(\mathbf{F}^t\left(\mathbf{S}^{-1}\right)\right)\delta\hat{\mathbf{d}}^{\star}\right\}, \tag{2.67}$$

which can be written in the alternative form:

$$\nabla_m E(\mathbf{m}) = \Re\left\{\mathbf{F}^t\mathbf{v}\right\}, \tag{2.68}$$

where $\mathbf{v}$ can be introduced to represent the 'backpropagated wavefield':

$$\mathbf{v} = \left(\mathbf{S}^{-1}\right)\delta\hat{\mathbf{d}}^{\star}. \tag{2.69}$$

[2]Substituting equation (2.65) into equation (2.63) actually gives $(\mathbf{S}^{-1})^t$, however $\mathbf{S}$ is symmetric.

We can now summarize the calculation of the gradient as follows: First, the backpropagated wavefield, $\mathbf{v}$, is computed by propagating the time-reversed data residuals, $\hat{\mathbf{d}}^{\star}$, throughout the model, $\mathbf{m}$, by the forward operator, $\mathbf{S}^{-1}$. The backpropagated wavefield is then multiplied by the virtual source at each nodal point, $\mathbf{F}$, and the real part is taken (Pratt et al., 1998).

The gradient calculation in equations (2.67) and (2.68) can then be used in conjunction with the gradient algorithm described by equation (2.39) by starting from an initial or starting model, $\mathbf{m}_0$. We also require the linear step length estimate,

$$\alpha = \frac{||\nabla_m E(\mathbf{m})||}{||\mathbf{J}\nabla_m E(\mathbf{m})||}. \tag{2.70}$$

As we have alluded to several times up until this point, the starting model $\mathbf{m}_0$ is often derived from traveltime tomography, as described in Section 2.2.1. For synthetic FWI studies, it is not uncommon to use a Gaussian-smoothed version of the true models as the starting model (Gholami et al., 2013b). Note that by adopting the gradient algorithm, we have assumed that the updated model, $\mathbf{m}$, is simply a linear combination of the starting model, $\mathbf{m}_0$ and the model update, $\delta\mathbf{m}$; an assumption that is made in the framework of the Born approximation.

One final point of discussion is that of starting model accuracy, and how it is related to both the gradient function as well as the gradient algorithm for inversion. When defining the wavefield data residuals within the frequency domain, one has the option of incorporating information regarding the wave amplitude, phase or both. For the inversions performed in Chapters 3 and 4, I have performed 'phase-only' inversions, the reasons for which are two-fold: Firstly, the accuracy of amplitudes modelled by the anisotropic wave equation (described in Section 2.3) that I have adapted has not been studied or quantified to date (Operto et al., 2009). Furthermore, it has been shown in the literature that, for rapidly varying heterogenous media such as the ones described in Chapter 4, the phase-only adaptation of the misfit function is more robust as the model gradient is extremely sensitive to strong amplitude variations (Kamei et al., 2014).

Our choice of phase-only inversions does not overcome the problem of non-linearity due to the oscillatory nature of the phase. The 'half-cycle criterion', described by Sirgue and Pratt (2004), suggests that the gradient algorithm will converge to an erroneous final model if the wavefield produced by the starting model ($\mathbf{u}_{est}$) is more than $\pi/2$ out-of-phase with the observed wavefield ($\mathbf{d}$). Luckily, the quality control check for this is rather simple, and can be determined empirically by visual inspection of the time- or frequency-domain waveforms produced in the starting model. Nevertheless, it is nearly impossible to generate a starting model that satisfies the half-cycle criterion at every point, so some mismatch manifests in the models produced

from inversion as artifacts, an occurance that has become commonly referred to in FWI studies as the 'cycle-skipping' phenomena (Sirgue and Pratt, 2004; Brenders and Pratt, 2007; Virieux and Operto, 2009).

It should be noted that FWI is a major field of research, therefore the techniques described in this thesis may differ from those explained elsewhere. I refer the reader to Virieux and Operto (2009) for a detailed, modern overview of FWI, and how it has developed over the past several decades. The inversion techniques described in this Section have been commonly referred to as 'conventional' FWI in recent years (Leeuwen et al., 2013), and several areas of research have been aimed at improving all aspects of FWI, ranging from alternative calculations of the misfit functional (Luo and Schuster, 1991; Leeuwen and Mulder, 2010) to expansion and reformulation of the model search space (Haber et al., 2000; Leeuwen et al., 2013). The primary focus of these works has been the non-linear nature of the seismic waveform problem, and how it relates to final model accuracy. For this thesis, I adopt the conventional FWI approach and I assume that the models generated using anisotropic traveltime tomography are accurate enough to serve as starting points for anisotropic FWI using the wave equation and finite-difference approach described in the following section.

2.3 The Acoustic Anisotropic Wave equation for Transversely Isotropic Media

In this subchapter, I show the evolution of the anisotropic wave equation for TI media from its conception by Alkhalifah (1998) to the system of equations proposed by Operto et al. (2009). I have integrated this system into the inversion strategies of Pratt (1990) and Pratt et al. (1998). Here I show the full derivation as the equation has been modified by several authors throughout the literature for several applications, as well as to be compatible with numerous modelling approaches. The objective of the derivation is to arrive at a set of coupled second-order equations that describe anisotropic wave propagation in an acoustic medium as a function of the heterogenous VTI parameters ϵ, δ, θ_0 and V_{P_0} (equations (2.100) and (2.101)). In Section 3.2.1, I describe the implementation that I have adapted to integrate this equation system into the FWI workflow which includes a description of both the forward operator as well as the boundary conditions.

Despite its comparative simplicity to its elastic counterpart, deriving an anisotropic equivalent to the acoustic wave equation is non-trivial. For one, it describes a non-physical process as anisotropic media are inherently elastic in nature (the pressure field is a scalar, therefore it should not vary with direction). Nevertheless, Alkhalifah (1998, 2000) proposed an acous-

tic wave equation for VTI media by setting the S-wave velocity along the vertical symmetry axis equal to zero. Starting with his equation relating vertical and horizontal slowness in VTI media:

$$V_{P_0}^2 p_z^2 = 1 - \frac{V_{nmo}^2 p_h^2}{1 - 2V_{nmo}^2 \eta p_h^2}, \tag{2.71}$$

he derived a fourth-order time-domain acoustic wave equation for VTI media:

$$\frac{\partial^4 F_t}{\partial t^4} - (1+2\eta)v^2\left(\frac{\partial^4 F_t}{\partial x^2 \partial t^2} + \frac{\partial^4 F_t}{\partial y^2 \partial t^2}\right) = v_v^2 \frac{\partial^4 F_t}{\partial z^2 \partial t^2} - 2\eta v^2 v_v^2 \left(\frac{\partial^4 F_t}{\partial x^2 \partial z^2} + \frac{\partial^4 F_t}{\partial y^2 \partial z^2}\right), \tag{2.72}$$

with

$$\eta = \frac{1}{2}\left(\frac{V_h^2}{V_{nmo}^2} - 1\right) = \frac{\epsilon - \delta}{1 + 2\delta}, \tag{2.73}$$

where F_t (x,y,z,t) is the pressure wavefield in the time-domain, V_{P_0} is the vertical p-wave velocity, V_{nmo} is the normal move-out (NMO) velocity, p_h is horizontal slowness, p_z is vertical slowness, v_v is vertical velocity and v is the interval NMO velocity in isotropic and TI media. For a finite-difference implementation, equation (2.72) is reduced from fourth order to second order in t by introducing $P = \partial^2 F_t / \partial t^2$ and re-arranging in terms of P:

$$\frac{\partial^2 P}{\partial t^2} = (1+2\eta)v^2\left(\frac{\partial^2 P}{\partial x^2} + \frac{\partial^2 P}{\partial y^2}\right) + v_v^2 \frac{\partial^2 P}{\partial z^2} - 2\eta v^2 v_v^2 \left(\frac{\partial^4 F_t}{\partial x^2 \partial z^2} + \frac{\partial^4 F_t}{\partial y^2 \partial z^2}\right), \tag{2.74}$$

for heterogenous ϵ and δ (and therefore heterogeneous η).

For subsurface models represented by VTI symmetry, Alkhalifah (2000) showed that equation (2.74) is kinematically accurate when compared to the elastic wave equation for VTI media. There are, however, some limitations to this method that must be discussed. Firstly, this equation is unable to accurately predict amplitudes, which has significance when inverting for the Quality factor, Q, of the medium. Furthermore, although the S-wave velocity is set to zero in the initial formulation of the phase velocity, S-wave energy is still excited within numerical simulations of the anisotropic acoustic equation (first observed in Alkhalifah, 2000, some other examples include Zhang et al., 2005 and Operto et al., 2009). These waves are believed to be generated at a source located in a VTI or TTI medium, or converted from P-waves at seismic boundaries. Herein, these S-wave artifacts will be referred to as 'spurious S-waves' and will be discussed in further detail later on.

Let us factor V_{nmo}^2 from each of the terms in equation (2.71):

$$V_{P_0}^2 p_z^2 = V_{nmo}^2 \left(\frac{1}{V_{nmo}^2} - \frac{p_h^2}{1 - 2V_{nmo}^2 \eta p_h^2} \right), \tag{2.75}$$

and divide both sides by $V_{P_0}^2$:

$$p_z^2 = \frac{V_{nmo}^2}{V_{P_0}^2} \left(\frac{1}{V_{nmo}^2} - \frac{p_h^2}{1 - 2V_{nmo}^2 \eta p_h^2} \right). \tag{2.76}$$

We can then set $p_h = \sqrt{p_x^2 + p_y^2}$ for $\mathbf{p} = (p_x, p_y, p_z)$ since the slowness in TI media will behave isotropically in the transverse plane, so that equation (2.76) now becomes:

$$p_z^2 = \frac{V_{nmo}^2}{V_{P_0}^2} \left(\frac{1}{V_{nmo}^2} - \frac{p_x^2 + p_y^2}{1 - 2V_{nmo}^2 \eta (p_x^2 + p_y^2)} \right). \tag{2.77}$$

Finally, letting $\mathbf{k} = \omega \mathbf{p}$ for $\mathbf{k} = (k_x, k_y, k_z)$ where k is the wavenumber and ω is the angular frequency, equation (2.77) becomes:

$$k_z^2 = \frac{V_{nmo}^2}{V_{P_0}^2} \left(\frac{\omega}{V_{nmo}^2} - \frac{\omega (k_x^2 + k_y^2)}{\omega - 2V_{nmo}^2 \eta (k_x^2 + k_y^2)} \right). \tag{2.78}$$

From equation (2.78), the higher-order partial differential equations described by equations (2.72) and (2.73) can be obtained, and have since been successfully implemented for several seismic migration and modelling applications (Alkhalifah, 2000; Grechka et al., 2004).

Zhou et al. (2006b) proposed an alternative anisotropic acoustic wave equation that is based on the same fundamental assumption as the one described by Alkhalifah (2000). Let us begin by re-arranging equation (2.78) and introducing the relation described by equation (2.73):

$$k_z^2 = \frac{\omega^2}{V_{P_0}^2} - (1 + 2\delta)(k_x^2 + k_y^2) - (1 + 2\delta)(k_x^2 + k_y^2) \left(\frac{2V_{P_0}^2 (\epsilon - \delta)(k_x^2 + k_y^2)}{\omega^2 - 2V_{P_0}^2 (\epsilon - \delta)(k_x^2 + k_y^2)} \right). \tag{2.79}$$

By multiplying both sides of equation (2.79) by the wavefield function in Fourier space, $\mathrm{p}(\omega, k_x, k_y, k_z)$, the fourth order partial differential equation described by equation (2.73) can be reduced to a coupled system of second-order equations by introducing an auxiliary wavefield function, $\mathrm{q}(\omega, k_x, k_y, k_z)$, such that:

$$k_z^2 p(\omega, k_x, k_y, k_z) = \left[\frac{\omega^2}{V_{P_0}^2} - (1+2\delta)(k_x^2 + k_y^2)\right] p(\omega, k_x, k_y, k_z) - (1+2\delta)(k_x^2 + k_y^2) q(\omega, k_x, k_y, k_z), \tag{2.80}$$

for

$$q(\omega, k_x, k_y, k_z) = \frac{2V_{P_0}^2(\epsilon - \delta)(k_x^2 + k_y^2)}{\omega^2 - 2V_{P_0}^2(\epsilon - \delta)(k_x^2 + k_y^2)} q(\omega, k_x, k_y, k_z). \tag{2.81}$$

Finally, if we take an inverse Fourier transform of both sides of equations (2.80) and (2.81), we obtain the final expressions for the anisotropic acoustic wave equation for VTI media:

$$\frac{1}{V_{P_0}^2}\frac{\partial^2 p}{\partial t^2} - (1+2\delta)\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right) p - \frac{\partial^2 p}{\partial z^2} = (1+2\delta)\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right) q, \tag{2.82}$$

and

$$\frac{1}{V_{P_0}^2}\frac{\partial^2 q}{\partial t^2} - 2(\epsilon - \delta)\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right) q = 2(\epsilon - \delta)\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right) p. \tag{2.83}$$

equations (2.82) and (2.83) are easier to implement (compared to equations (2.73) and (2.75)), and are kinematically accurate for modelling P-wave propagation in VTI media when compared to an elastic solution (Zhou et al., 2006b).

Until this point, we have derived all of our anisotropic wave equations under the assumption that the symmetry axis is parallel to the vertical (z-axis). The data cases in this thesis are limited to VTI media, however I still implement an equation than can handle a tilted symmetry axis for the purpose of future studies. Following on their earlier work, Zhou et al. (2006a) developed an anisotropic acoustic wave equation for 2-D TTI media that utilizes the same coupled strategy as their VTI method. Let us begin with the phase velocity relation for TI media developed by Tsvankin (1996):

$$\frac{V^2(\theta)}{V_{P_0}^2} = \epsilon \sin^2\bar{\theta} - \frac{f}{2} + -\frac{f}{2}\sqrt{\left(1 + \frac{2\epsilon \sin^2\bar{\theta}}{f}\right)^2 - \frac{8(\epsilon - \delta)\sin^2\bar{\theta}\cos^2\bar{\theta}}{f}}, \tag{2.84}$$

for

$$f = 1 - \frac{V_{S_0}^2}{V_{P_0}^2}, \tag{2.85}$$

where $V(\theta)$ is the phase velocity, $\bar{\theta}$ is the phase angle with respect to the symmetry axis, and $V_{S_0}^2$ is the S-wave velocity along the symmetry axis. Before going any further, note that the velocity relation for VTI media can be obtained from (2.84) by substituting the eikonal solution for phase velocity:

$$V^2(\theta) = \frac{1}{p_h^2 + p_h^2}, \tag{2.86}$$

such that

$$p_h = \frac{\sin\theta}{V(\bar{\theta})}, \tag{2.87}$$

$$p_v = \frac{\cos\theta}{V(\bar{\theta})}. \tag{2.88}$$

Returning to the derivation of the TTI wave equation, if we set $V_{S_0}^2 = 0$, equation (2.84) can be re-arranged as follows:

$$\frac{\cos^2\bar{\theta}}{V^2(\theta)} = \frac{1}{V_{P_0}^2} - (1+2\delta)\left(\frac{\frac{\sin^2\bar{\theta}}{V^2(\theta)}\frac{1}{V_{P_0}^2}}{\frac{1}{V_{P_0}^2} - 2(\epsilon-\delta)\frac{\sin^2\bar{\theta}}{V^2(\theta)}}\right). \tag{2.89}$$

Next, we introduce the following relations:

$$\sin\theta = \frac{V(\theta)k_x}{\omega}, \cos\theta = \frac{V(\theta)k_z}{\omega}, \tag{2.90}$$

where θ is the phase angle with respect to the z-axis. If we let θ_0 be equal to the angle between the symmetry axis and the vertical, we have $\bar{\theta} = \theta - \theta_0$. Substituting this expression and equation (2.90) into equation (2.89) gives us the dispersion relation for TTI function as a function of the angle of symmetry axis:

$$f_1 = \frac{\omega^2}{V_{P_0}^2} - (1+2\delta)f_2 - (1+2\delta)2(\epsilon-\delta)\frac{f_1 f_2}{\frac{\omega^2}{V_{P_0}^2} - 2(\epsilon-\delta)f_2}, \tag{2.91}$$

for

$$f_1(k_x, k_z, \theta) = k_z^2 \cos^2\theta - k_x k_z \sin 2\theta 0 + k_x^2 \sin^2\theta, \tag{2.92}$$

and

$$f_2(k_x, k_z, \theta) = k_x^2 \cos^2\theta + k_x k_z \sin 2\theta 0 + k_z^2 \sin^2\theta, \tag{2.93}$$

where the variables f_1 and f_2 can be introduced for clarity. As in Zhou et al. (2006b), we now multiply both sides of equation (2.91) by the wavefield function, p(ω,k_x,k_z), and introduce the auxiliary wavefield function, q(ω,k_x,k_z), such that

$$f_1 p = \left[\frac{\omega^2}{V_{P_0}^2} - (1+2\delta)f_2\right]p - (1+2\delta)f_2 q, \tag{2.94}$$

for

$$q(\omega, k_x, k_z) = \frac{2V_{P_0}^2(\epsilon-\delta)f_2}{\omega^2 - 2V_{P_0}^2(\epsilon-\delta)f_2}p(\omega, k_x, k_z). \tag{2.95}$$

Finally, we take the inverse Fourier transform of both sides of equations (2.94) and (2.95) to obtain the final coupled system of partial differential equations for 2-D TTI media:

$$\frac{1}{V_{P_0}^2}\frac{\partial^2 p}{\partial t^2} - (1+2\delta)\bar{H}p - \bar{H}_0 p = (1+2\delta)\bar{H}q, \tag{2.96}$$

$$\frac{1}{V_{P_0}^2}\frac{\partial^2 q}{\partial t^2} - 2(\epsilon-\delta)\bar{H}q = 2(\epsilon-\delta)\bar{H}p, \tag{2.97}$$

where two new differential operators, $\bar{H}$ and $\bar{H}_0$ can be defined as follows

$$\bar{H} = \cos^2\theta_0 \frac{\partial^2}{\partial x^2} + \sin^2\theta_0 \frac{\partial^2}{\partial z^2} - \sin 2\theta_0 \frac{\partial^2}{\partial x \partial z}, \tag{2.98}$$

$$\bar{H}_0 = \sin^2\theta_0 \frac{\partial^2}{\partial x^2} + \cos^2\theta_0 \frac{\partial^2}{\partial z^2} + \sin 2\theta_0 \frac{\partial^2}{\partial x \partial z}. \tag{2.99}$$

Note that, if we set θ_0=0°, equations (2.95),(2.96), and (2.97) are reduced to their 2-D VTI equivalents (equations (2.81),(2.82), and (2.83) respectively. Zhou et al. (2006a) implement equations (2.96) and (2.97) in the time-domain and show good kinematic agreement with the elastic equivalent for θ_0= 0°, 45° and 90°.

Up until this point, we have derived all of our equations for a constant density medium. This is also a poor assumption as density will vary based on several factors such as mineral composition and depth (Gardner et al., 1974). Operto et al. (2009) modify the coupled system of Zhou et al. (2006a) to allow for heterogenous density. By introducing heterogenous buoyancy (the inverse of density) into the $\bar{H}$ and $\bar{H}_0$ operators and remembering that the isotropic acoustic P-wave velocity, $V_{P_0} = \sqrt{\kappa_0 b}$, where κ_0 is the bulk modulus along the symmetry axis and b is buoyancy, the system of equations becomes

$$\frac{1}{\kappa_0}\frac{\partial^2 p}{\partial t^2} - (1+2\delta)\bar{H}p - \bar{H}_0 p = (1+2\delta)\bar{H}q, \tag{2.100}$$

$$\frac{1}{\kappa_0}\frac{\partial^2 q}{\partial t^2} - 2(\epsilon-\delta)\bar{H}q = 2(\epsilon-\delta)\bar{H}p, \tag{2.101}$$

with

$$\bar{H} = \cos^2\theta_0 \frac{\partial}{\partial x} b \frac{\partial}{\partial x} + \sin^2\theta_0 \frac{\partial}{\partial z} b \frac{\partial}{\partial z} - \frac{\sin 2\theta_0}{2}\left(\frac{\partial}{\partial x} b \frac{\partial}{\partial z} + \frac{\partial}{\partial z} b \frac{\partial}{\partial x}\right), \tag{2.102}$$

and

$$\bar{H}_0 = \sin^2\theta_0 \frac{\partial}{\partial x} b \frac{\partial}{\partial x} + \cos^2\theta_0 \frac{\partial}{\partial z} b \frac{\partial}{\partial z} + \frac{\sin 2\theta_0}{2}\left(\frac{\partial}{\partial x} b \frac{\partial}{\partial z} + \frac{\partial}{\partial z} b \frac{\partial}{\partial x}\right). \tag{2.103}$$

Before going further, a couple of important cases can be mentioned here. By setting both ϵ and δ equal to zero, equation (2.101) goes to zero and equation (2.100) is reduced to the second order isotropic wave equation. Furthermore, if ϵ=δ, such as in the case of elliptical transverse isotropy, equation (2.101) will again vanish. In this sense, equation (2.100) can be thought of as a measure of ellipticity, while equation (2.101) can be thought of as a compensation or correction term for the changing wavefront in both the lateral and depth dimensions (Zhou et al., 2006a).

Operto et al. (2009) then recast the system of second-order equations into a hyperbolic system of first-order equations by introducing the auxiliary wavefields, p_x, p_z, q_x and q_z, such that:

$$\frac{1}{\kappa_0}\frac{\partial p}{\partial t} = A_x\frac{\partial p_x}{\partial x} + B_x\frac{\partial p_z}{\partial x} + C_x\frac{\partial q_x}{\partial x} + D_x\frac{\partial q_z}{\partial x} + A_z\frac{\partial p_x}{\partial z} + B_z\frac{\partial p_z}{\partial z} + C_z\frac{\partial q_x}{\partial z} + D_z\frac{\partial q_z}{\partial z}, \quad (2.104)$$

$$\frac{1}{\kappa_0}\frac{\partial q}{\partial t} = E_x\frac{\partial p_x}{\partial x} + F_x\frac{\partial p_z}{\partial x} + G_x\frac{\partial q_x}{\partial x} + H_x\frac{\partial q_z}{\partial x} + E_z\frac{\partial p_x}{\partial z} + F_z\frac{\partial p_z}{\partial z} + G_z\frac{\partial q_x}{\partial z} + H_z\frac{\partial q_z}{\partial z}, \quad (2.105)$$

$$\frac{\partial p_x}{\partial t} = b\frac{\partial p}{\partial x}, \frac{\partial p_z}{\partial t} = b\frac{\partial p}{\partial z}, \quad (2.106)$$

$$\frac{\partial q_x}{\partial t} = b\frac{\partial q}{\partial x}, \frac{\partial q_z}{\partial t} = b\frac{\partial q}{\partial z}, \quad (2.107)$$

where the following coefficients are introduced for clarity and compactness:

$$\begin{aligned}
&A_x = 1 + 2\delta\cos^2(\theta_0), B_x = -\delta\sin(2\theta_0),\\
&C_x = (1+2\delta)\cos^2(\theta_0), D_x = -(1+2\delta)\frac{\sin(2\theta_0)}{2},\\
&A_z = B_x, B_z = 1 + 2\delta\sin^2(\theta_0),\\
&C_z = D_x, D_z = (1+2\delta)\sin^2(\theta_0),\\
&E_x = 2(\epsilon-\delta)\cos^2(\theta_0), F_x = -(\epsilon-\delta)sin(2\theta_0),\\
&G_x = E_x, H_x = F_x,\\
&E_z = F_x, F_z = 2(\epsilon-\delta)\sin^2(\theta_0),\\
&G_z = F_x, H_z = F_z. \quad (2.108)
\end{aligned}$$

Next, we introduce 1D damping functions, ξ_x and ξ_z for convolutional Perfectly Matched Layers (c-PML) absorbing boundary conditions, which are discussed in detail in Drossaert and Giannopoulos (2007) and Komatitsch and Martin (2007). These functions define a zone around the boundaries of our model in which the waves are gradually damped as they travel further into the PML layer, thus mitigating erroneous reflections off boundary edges. In all other areas of the grid, $\xi_x = \xi_z = 1$. Now, if we take an inverse Fourier transform of the hyperbolic system of first order equations, we arrive at the frequency-domain 2-D acoustic wave equation for TI media that we will implement into our anisotropic FWI approach:

$$\frac{-i\omega}{\kappa_0}p = \frac{1}{\xi_x}(A_x\frac{\partial p_x}{\partial x} + B_x\frac{\partial p_z}{\partial x} + c_x\frac{\partial q_x}{\partial x} + D_x\frac{\partial q_z}{\partial x}) + \frac{1}{\xi_z}(A_z\frac{\partial p_x}{\partial z} + B_z\frac{\partial p_z}{\partial z} + c_z\frac{\partial q_x}{\partial z} + D_z\frac{\partial q_z}{\partial z}), \tag{2.109}$$

$$\frac{-i\omega}{\kappa_0}q = \frac{1}{\xi_x}(E_x\frac{\partial p_x}{\partial x} + F_x\frac{\partial p_z}{\partial x} + G_x\frac{\partial q_x}{\partial x} + H_x\frac{\partial q_z}{\partial x}) + \frac{1}{\xi_z}(E_z\frac{\partial p_x}{\partial z} + F_z\frac{\partial p_z}{\partial z} + G_z\frac{\partial q_x}{\partial z} + H_z\frac{\partial q_z}{\partial z}), \tag{2.110}$$

$$\begin{aligned} -i\omega p_x &= \frac{b}{\xi_x}\frac{\partial p}{\partial x}, \quad -i\omega p_z = \frac{b}{\xi_z}\frac{\partial p}{\partial z}, \\ -i\omega q_x &= \frac{b}{\xi_x}\frac{\partial q}{\partial x}, \quad -i\omega q_z = \frac{b}{\xi_z}\frac{\partial q}{\partial z}, \end{aligned} \tag{2.111}$$

for

$$\xi_x = \gamma_x + \frac{d_x}{\chi_x + i\omega}, \tag{2.112}$$

$$\xi_z = \gamma_z + \frac{d_z}{\chi_z + i\omega}, \tag{2.113}$$

where d represents the distance into the PML layer and χ and γ are damping functions (Drossaert and Giannopoulos, 2007; Komatitsch and Martin, 2007).

Although we are able to utilize the inversion strategies of Pratt et al. (1998) for the works presented in this thesis, the anisotropic wave equation is not compatible with the differencing operator and matrix factorization scheme that is typically coupled with this technique. Most studies by Dr. Pratt and his collaborates to date have used the mixed-grid approach of Jo et al. (1996), a nine-point isotropic FD operator, to perform isotropic or elliptically isotropic FWI studies (Pratt et al., 2005; Brenders and Pratt, 2007; Afanasiev et al., 2014). A simple comparison between the system of anisotropic equations above, to the isotropic wave equation described in equation (2.19) shows that there is an additional wavefield (**q**) present in the anisotropic case. If we recall the matrix formulation in equation (2.57), this additional wavefield must be accounted for within the **u** term, which in turn must be accounted for within the structure of the **S** and **f** terms as well. In the next subsection, I describe the finite-difference frequency-domain implementation that I have adopted to allow the inversion strategies of Pratt et al. (1998) to be compatible with the anisotropic wave equation for TI media (equations (2.109) and (2.110)).

2.3.1 The Frequency-Domain Finite-Difference Implementation for Anisotropic Media

For accurate simulations, we require a differencing operator that can account for changes to the anisotropic parameters in both directions. Furthermore, it must be compatible with the first-order velocity-stress hyperbolic system of equations (2.109) and (2.110). As such, I adopt the 'parsimonious mixed-grid approach' proposed by Operto et al. (2009). This discretization is based on the original mixed-grid method proposed by Jo et al. (1996), but recast into the framework of the parsimonious staggered-grid method proposed by Hustedt et al. (2004). In the original formulation proposed by Jo et al. (1996), spatial derivatives of a second-order wave equation (such as equation (2.19)) are computed on two separate coordinate systems, the results of which are then combined linearly to form one compact stencil. In the parsimonious approach, the wave equation is instead written as a first-order velocity-stress hyperbolic system and discretized on staggered-grid stencils for the same two coordinate systems as the mixed-grid approach (Virieux, 1986). In the parsimonious case, the particle wavefields are eliminated during discretization, such that the resulting solution on each coordinate system is the second-order pressure wavefield (in the acoustic case). Finally, the result on each coordinate system can be combined linearly to form one discrete wave equation, provided that the wavefields kept after elimination are discretized onto the same coordinate system (Operto et al., 2009).

Following Operto et al. (2009), I discretize equations (2.109) and (2.110) on two coordinate systems. The first is the well-known Cartesian System (CS), and the second is a coordinate system rotated 45° from the cartesian system (RS). All further mentions of the cartesian and rotated coordinate systems with be denoted by the subscripts 'c' and 'r' respectively. After discretization and linear combination of the two stencils, the particle wavefields p_x, p_z, q_x, q_z are eliminated, and the system of equations can be written in matrix form for the source-free case as follows:

$$\begin{pmatrix} \mathbf{M}_p + w_1\mathbf{A}_r + (1-w_1)\,\mathbf{A}_c & w_1\mathbf{B}_r + (1-w_1)\,\mathbf{B}_c \\ w_1\mathbf{C}_r + (1-w_1)\,\mathbf{C}_c & \mathbf{M}_q + w_1\mathbf{D}_r + (1-w_1)\,\mathbf{D}_c \end{pmatrix} x \begin{pmatrix} \mathbf{p} \\ \mathbf{q} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \tag{2.114}$$

where $\mathbf{M}_p$ and $\mathbf{M}_q$ denote the diagonals of the mass matrix for the p- and q-wavefields, blocks $\mathbf{A}_r$, $\mathbf{B}_r$, $\mathbf{C}_r$, $\mathbf{D}_r$ and $\mathbf{A}_c$, $\mathbf{B}_c$, $\mathbf{C}_c$, and $\mathbf{D}_c$ are the stiffness matrices for the RS and CS stencils, and w_1 is the experimentally determined coefficient that weights the relative components of each of the two stencils (Operto et al., 2009). It is important to note here that the notation of the stiffness matrices should not be confused with the alphabetical notation used to describe the coefficients in the system of anisotropic wave equations (equation 2.108). Rather, each of

the stiffness matrices is comprised of nine-point stencils that utilize a unique subset of these alphabetical coefficients. I refer the reader to Appendices A and B of Operto et al. (2009) for a full description of these stencil coefficients, as they are quite lengthy, and therefore impractical to show here. To improve stencil accuracy, we adopt the 'anti-lumped mass' technique proposed by Stekl and Pratt (1998), whereby the mass term is averaged across the nine points of the stencil. For the *ij*th term in the mass diagonals $\mathbf{M}_p$ and $\mathbf{M}_q$, the diagonal is replaced by its weighted average, such that:

$$\begin{aligned}\frac{\omega}{\kappa_{ij}} \rightarrow w_{m1}\frac{\omega}{\kappa_{ij}} + \frac{w_{m2}}{4}\left(\frac{\omega}{\kappa_{i+1,j}} + \frac{\omega}{\kappa_{i\text{-}1,j}} + \frac{\omega}{\kappa_{i,j+1}} + \frac{\omega}{\kappa_{i,j\text{-}1}}\right) \\ + \frac{(1 - w_{m1} - w_{m2})}{4}\left(\frac{\omega}{\kappa_{i+1,j+1}} + \frac{\omega}{\kappa_{i\text{-}1,j\text{-}1}} + \frac{\omega}{\kappa_{i\text{-}1,j+1}} + \frac{\omega}{\kappa_{i\text{-}1,j\text{-}1}}\right),\end{aligned} \tag{2.115}$$

where the arrow denotes a replacement and w_{m1} and w_{m2} are weighting coefficients determined jointly with w_1 during dispersion analysis. As we did not perform these analyses, we use the weighting coefficients recommended by Operto et al. (2009), the values of which are summerized in Table 2.1. The reader should note that these coefficients are subtly different to those proposed by Jo et al. (1996), which is probably related to the incorporation of anisotropy into the modelling process.

Let us pause briefly to carefully examine the structure of equation (2.114). First, we can clearly see it is analogous to the forward problem presented in equation (2.57). That is, the first term forms the operator (**S**) which contains the 'wave physics', and the second term contains the p- and q-wavefields (**u**) for a given source (**f**). It is also important to note that the mass diagonals are only applied to the first and fourth quadrants. This is because each row of every quadrant represents one point in the model, so applying the mass term to the second and third quadrants as well would double count the mass effect for the p- and q-wavefields respectively. Setting $\epsilon = \delta = \theta = 0$, or $\epsilon = \delta \neq 0$, we see that equation (2.110) vanishes (the q-wavefield is nil), and equation (2.114) is reduced to the upper left block only, or:

$$\left(\mathbf{M}_p + w_1\mathbf{A}_r + (1 - w_1)\,\mathbf{A}_c \right)\mathbf{p} = 0. \tag{2.116}$$

Let us now examine the size of the matrices in equations (2.114) and (2.116) in more detail. For simplicity's sake, we will only consider a single-frequency, single-source case so that the arrays themselves are 2-D. For multi-source multi-frequency problems (Chapters 3 and 4), I store the source information in the third array dimension, and parallelize the computation over frequencies. Let us first consider the isotropic case. Letting *nx* equal the number of gridpoints in the x-dimension and *nz* equal the number of gridpoints in the z-dimension, each of the terms

in the forward operator in equation (2.116) will have *nx*nz* x *nx*nz* complex-valued elements, the p-wavefield will have *nx*nz* x 1 elements, and the source term also having *nx*nz* x 1 elements. Let us note here that this would also be the case if we used the isotropic wave equation with the mixed-grid method of Jo et al. (1996). If we now assume elliptical anisotropy, the number of matrix elements remains the same (thus, the runtimes will be unaffected), however the use of the isotropic wave equation with the mixed-grid method would no longer be justified, as this does not account for changes to the anisotropic parameters in both directions.

Finally, if we consider the anelliptical case ($\epsilon \neq \delta$), the additional matrices are non-zero and we must utilize the entire system described by equation (2.114). First, we have an additional wavefield parameter, **q**, which must be accompanied by an additional source term ($\mathbf{f} = (f_p, f_q)^T$). For all of the works presented in this thesis, the source term for the q-wavefield (f_q) is set to zero. Nevertheless, we account for this additional source term by doubling the length of the wavefield and source vectors, such that the **u** and **f** terms now comprise 2**nx*nz* x 1 elements. Since each of the four quadrants in the differencing operator are now non-zero, its size is significantly increased when compared to the isotropic/elliptical case, increasing to a total size of 2**nx*nz* x 2**nx*nz* elements. The overall increase in matrix components has a significant impact on computation runtimes, as well as computational storage for the anisotropic case. However, these increases are necessary as we are introducing a more thorough modelling approach that better represents anisotropic wave behaviour. For the anisotropic FWI results presented in Chapters 3 and 4, I implement the FDFD system described by equation (2.114) into a seismic waveform modelling framework developed in Python (Smithyman et al., 2015), which is part of an open-source framework for waveform inversion called *Zephyr* (https://zephyr.space).

w_1	w_{m1}	w_{m2}
0.4382634	0.6287326	0.3712667

Table 2.1: Optimal weighting coefficients for equations 2.114 and 2.115 determined jointly during dispersion analysis by Operto et al. (2009).

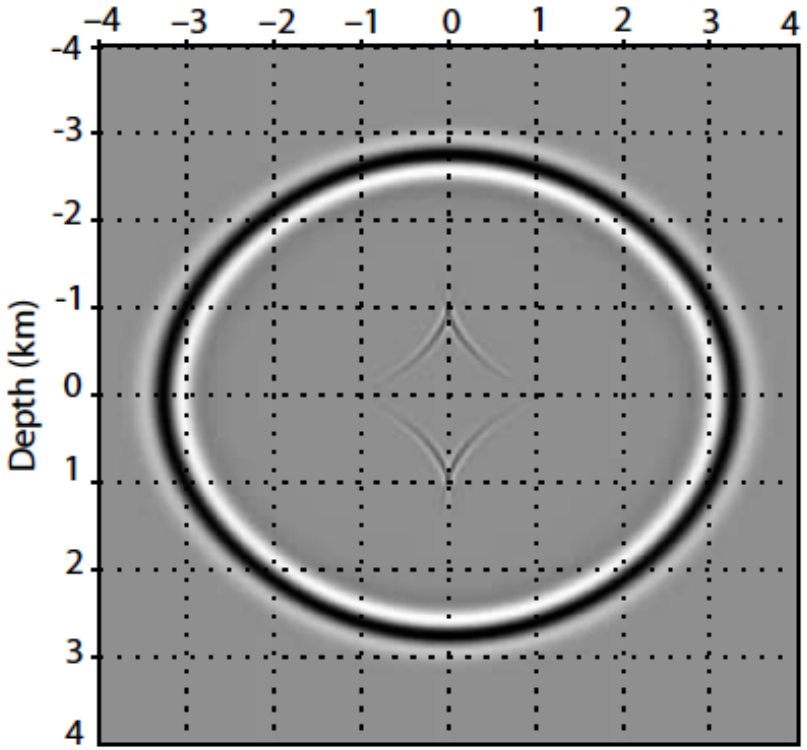
To ensure that I implemented Operto et al. (2009)'s technique correctly, I replicate some of the time-domain forward modelling results published in their paper (their Figure 3 on page T82). Table 2.2 summerizes the forward modelling parameters used for this simulation, and Figure 2.2 shows the time-domain snapshot for a source located at the very centre of the grid. My results in Figure 2.2 are most comparable with panel (c) of their results, as no anisotropic weighting of the source region was used, so I show this panel for comparison. Operto et al. (2009) shows that for their modelling technique, a secondary wave that travels at approximately one-third the speed of the pressure wave is erroneously injected when the source is located in a medium where $\epsilon > \delta$. These slow-moving waves have become known as 'spurious S-wave

artifacts ' in the literature (Alkhalifah, 2000; Zhou et al., 2006b; Operto et al., 2009). In Chapter 3 I discuss some of the strategies that have been developed to address these artifacts, as well as the ones that I have adopted for these analyses. For this simulation, the source is a Ricker wavelet with a dominant frequency of 4 Hz, and $\epsilon > \delta$ such that the spurious S-wave energy can observed. By comparing the two images, it is obvious that the wave modelling techniques are the same.

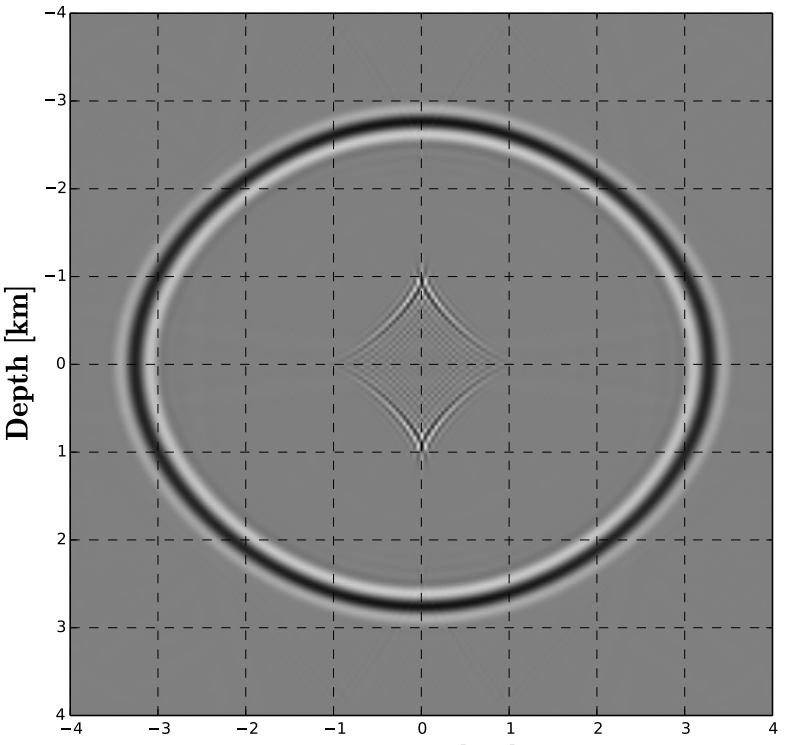
Using the same modelling parameters, I re-perform this analysis for several unique combinations of ϵ and δ to show how the wavefront differs from its isotropic counterpart in each case (Figure 2.3). Note how the spurious S-wave energy is only injected where $\epsilon > \delta$, as expected. The first four panels of Figure 2.3 were computed for VTI symmetry ($\theta_0 = 0°$), while the fifth and sixth panels were computed for TTI symmetry ($\theta_0 = 45°$). The primary reason that I have restricted my analyses in Chapters 3 and 4 of this thesis to the VTI case is related to the spurious S-wave artifacts. Operto et al. (2009) showed that when $\epsilon > \delta$ and $\theta_0 \neq 0$, significant instabilities are observed within the source and PML regions of the model (their Figure 9 on page T87) that become more prominent as the number of grid points per wavelength increases. These instabilities are clearly present in panel (e) of my Figure 2.3, but not panel (f) as $\delta > \epsilon$ in this case. It is also apparent that the magnitude of anellipticity (the difference between ϵ and δ) is positively correlated with the amplitude of the S-wave artifacts. This is not a published relationship, but rather an observation that I have made over the course of this thesis. This has implications for some of my results shown in Chapter 4, but more on this later.

Parameter	Value
Dominant Frequency [Hz]	4
$V_{P0}[km/s]$	2.0
ϵ	0.2
δ	0.1
Grid Interval [m]	10

Table 2.2: Forward modelling parameters used in Figures 2.2 and 2.3.

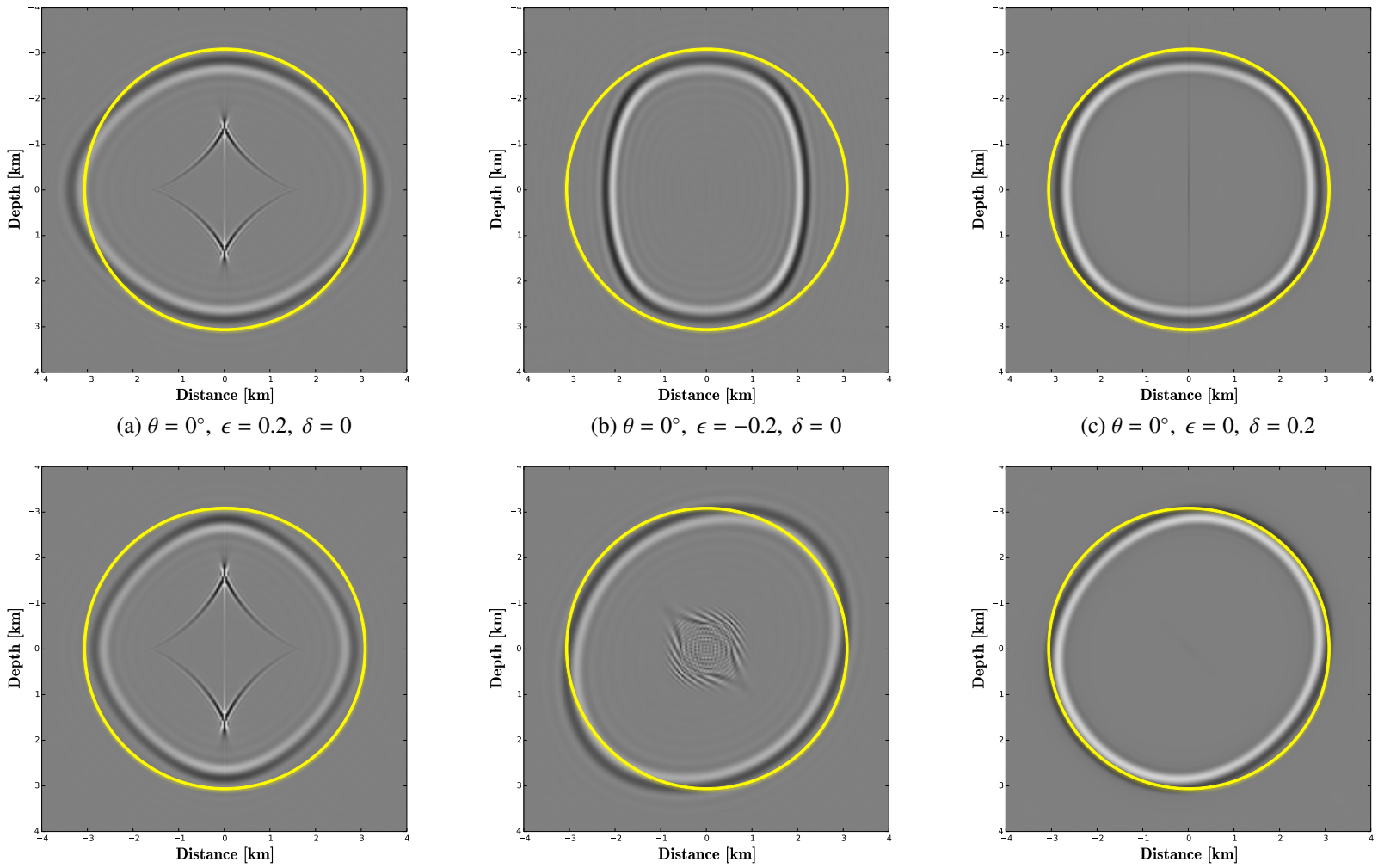


(a) Operto et al. (2009)'s Result



(b) My Result

Figure 2.2: Forward modelled time-domain snapshot for comparison to Operto et al. (2009)'s results. Modelling parameters for this simulation are summarized in Table 2.2. The source is located at the very centre of the grid.



(a) $\theta = 0°,\ \epsilon = 0.2,\ \delta = 0$ (b) $\theta = 0°,\ \epsilon = -0.2,\ \delta = 0$ (c) $\theta = 0°,\ \epsilon = 0,\ \delta = 0.2$

(d) $\theta = 0°,\ \epsilon = 0,\ \delta = -0.2$ (e) $\theta = 45°,\ \epsilon = 0.2,\ \delta = 0.1$ (f) $\theta = 45°,\ \epsilon = 0.1,\ \delta = 0.2$

Figure 2.3: Forward modelled time-domain snapshots for several combinations of ϵ and δ. An overlay in yellow representing a perfect circle corresponding to the isotropic case ($\delta = \epsilon = 0$) is provided for reference. In panels (a) and (b) we see that the effect of ϵ is analgous to stretching or compressing perpendicular to the symmetry axis. In panels (c) and (d) we see that δ affects the curvature of the wavefront as a function of the distance away from the symmetry axis. If we observe the wavefront as it moves from the vertical to the horizontal, we see that for $\delta > 0$ it appears to overtake the isotropic wavefront, whereas for $\delta < 0$ it appears to fall behind. For panels (e) and (f) we see that the axis of symmetry is tilted from the vertical, and that the S-wave energy in panel (e) is less stable when compared to the energy observed in panels (a) and (d).

Chapter 3

Synthetic Crosshole Study

In this Chapter, I apply the joint methods of anisotropic traveltime tomography and VTI FWI to a synthetically generated dataset. The primary objective of this study is to determine if the traveltime methods of Chapman and Pratt (1992) and Pratt and Chapman (1992) are suitable for generating adequate starting models for FWI using the inversion strategies of Pratt et al. (1998) and the FDFD approach of Operto et al. (2009). More specifically, I tackle the question of whether or not the traveltime inversion is capable of producing multi-parameter (V_V, ϵ, δ) results that satisfy the half cycle criterion for anisotropic FWI. First, I provide a brief introduction to the survey, as well as a review of parameterization and resolution studies from recent years. Next, I show the traveltime tomography results for several cases, as well as the corresponding FWI results produced from these starting models. Finally, I conclude with a discussion regarding the compatibility of the two techniques as well as the implications of applying this joint method to field data.

3.1 Introduction

The feasibility of multiparameter inversions and parameter resolution tradeoff have been topics of significant interest in modern FWI studies (Plessix and Cao, 2011; Kamei et al., 2013; Gholami et al., 2013b). Due to the extremely ill-posed nature of FWI, non-uniqueness (in which multiple models fit the data equally well) is a significant issue that is further amplified by the introduction of additional inversion parameters. The inversion results are contaminated by parameter 'crosstalk', whereby the poor resolution of one parameter can impact the resolution of another and vice-versa (resulting in an *a posteriori* dependance between model parameters). This leads to ambiguities when claiming that one has simultaneously inverted for two or more parameters. There are no studies to date that claim to have fully overcome the problems related to parameter crosstalk. Rather, modern studies have adapted several approximations or

imposed constraints on the inversion with the goal of mitigating crosstalk in multiparameter FWI. One such example of this is the hierarchical approach, in which preliminary inversions are focused around resolving the most impactful parameter first before incorporating other parameters into the inversion process (Shipp and Singh, 2002; Plessix and Cao, 2011; Kamei et al., 2013).

In the simplest of cases (acoustic isotropic media), FWI is restricted to inversions for P-wave velocity and/or P-wave attenuation. Kamei et al. (2013) investigated the resolution trade-off between velocity and attenuation by conducting several synthetic crosshole experiments using two approaches. The first was a simultaneous approach with a scaling term applied to the gradient of the attenuation parameter, and the second was termed a 'sequential' approach, in which initial inversions were conducted for velocity only, followed by the incorporation of attenuation inversions once an acceptable velocity model was obtained (a version of the hierarchical approach described above). In both cases, Kamei et al. (2013) showed the presence of parameter crosstalk, but further showed this can be mitigated by applying the regularization term to the gradient of the attenuation parameter, or by adapting the sequential inversion technique.

Moving from an isotropic to an anisotropic medium introduces additional parameters per spatial location depending on the selected parameterization. Plessix and Cao (2011) and Gholami et al. (2013b) performed sensitivity and trade-off analyses for acoustic FWI in a VTI medium for surface reflection data. In both cases, the diving and reflected waves were found to be more sensitive to the velocity parameter classes than the anisotropy parameter classes, with δ having the least influence on the kinematics of the propagating waves. Furthermore, Gholami et al. (2013b) explored several possible combinations of parameters in order to determine the optimal parameterization for FWI. They concluded that, for the crosshole case, a parameterization consisting of the velocity along the symmetry axis,V_{P_0}, and Thomsen's parameters ϵ and δ is the most suitable provided that the long wavelength features of the background ϵ are a sufficient representation of the subsurface.

It is important to acknowledge, however, that these studies were initially conducted for surface reflection data and the conclusions were extended to the crosshole case. There has been no conclusive study on the optimal parameterization for acoustic FWI in VTI media for crosshole data. Therefore, I adopt the suggested parameterization of Gholami et al. (2013b) and make the assumption that the ϵ model provided by anisotropic traveltime tomography satisfies the success criteria outlined in their paper. In regards to mono- versus multiparameter FWI, Plessix and Cao (2011) suggested that a monoparameter approach for updating velocity could be suitable if traveltime tomography provides accurate models for the Thomsen parameters. Furthermore, Gholami et al. (2013b) concluded that the aforemetioned parameterization (V_V, ϵ

and δ) is best coupled with a monoparameter FWI approach, as significant tradeoff is observed between velocity and Thomsen's anisotropy parameters.

I test my joint approach using a modified version of the synthetic crosshole experiment first presented by Pratt and Chapman (1992). The 'true' models consist of a smooth background velocity ($V_{background} = 3000\,m/s$) with a square anomaly ($V_{anomaly} = 3200\,m/s$) in the centre, with homogenous models for the anisotropy parameters ($\epsilon = 0.2$, δ=0.1, $\theta_0 = 0°$). The original dimensions are 100 m by 200 m, however I have padded the model by 15 m in all directions to minimize internal reflections from the PML boundaries. Source and receiver locations were chosen at 0.5 m intervals spanning the entirety of the depth dimension on either side of the model before padding was applied, for a total of 201 sources, 201 receivers and thus 40401 traces.

Figure 3.1 shows the true vertical velocity model with the source and receiver geometries overlaid. This is a relatively simple test case, and obtaining satisfactory results will be greatly aided by the level of *a priori* information that one can infer from the true models (the expectation of VTI symmetry). Furthermore, the modelling approach used to generate the true models is the same one that was used for forward modelling during inversion, which is commonly referred to as committing the 'inverse crime' (Kaipio and Somersalo, 2007). Here, the intermediate step involves raytracing which is based on a different algorithm, so perhaps this is only partially an inverse crime. Nevertheless, the primary objective of this study is to observe and quantify the level of tradeoff between parameters, and it is critical to utilize controlled experiments like this one for testing specific additions and alterations to the FWI process (Kamei et al., 2013; Gholami et al., 2013a; Afanasiev et al., 2014).

Before one can develop an efficient strategy for traveltime tomography or FWI, selection of appropriate modelling parameters is an important consideration. As this is a synthetic test, we must first forward model with the true models and perform first-break picking on these data. The picked traveltimes are then input into traveltime tomography as the observed data, $\mathbf{d}$. For this I take advantage of the fact that this is a synthetic test, therefore the data are 'noise-free'. As such, I utilize an auto-picker that is based off of the short-term and long-terms averages (STA/LTA) of the traces (Allen, 1978) to perform first break picking reliably and efficiently. Table 3.1 summarizes the selected model parameters for generating the synthetic crosshole data. The spacing between model grid points is an important consideration, as finer grid sizes allow for higher frequencies and improved resolution of the target structures, but consequently increase inversion runtimes. The maximum frequency was chosen in accordance with the minimum wavelength required to resolve the central velocity anomaly based on the chosen grid size. Synthetic shot gathers were generated in the time-domain using a Keuper wavelet with a dominant frequency of 533 Hz and a maximum frequency of 1600 Hz. Here, I model in

reduced time ($t_{reduced} = t - \text{offset}/v$) to mitigate computational costs. For FDFD modelling, I use a Kaiser windowed sinc function to spatially distribute the source term (Hicks, 2002).

At this point the choice was made to keep $\epsilon > \delta$. By choosing these values the S-wave artifacts are purposely injected into the modelled data. The condition that $\epsilon > \delta$ is almost guaranteed to occur for heterogenous media as δ is known to be rapidly varying in realistic geological formations. Therefore, I purposely inject them in the synthetic test so that I can observe their behaviour and develop strategies to address them for the field data case. Figure 3.2 shows a forward modelled shot gather with the true models. Several studies have been focused on eliminating or attenuating the S-wave artifacts discussed in Chapter 2. In the simplest of cases, these S-waves can be removed by placing the source in a thin, isotropic or elliptically isotropic layer, however this introduces modelling inaccuracies. Fletcher et al. (2009) introduced a small amount of shear wave velocity along the symmetry axis to stabilize the simulation of S-waves for reverse time migration in TTI media. Métivier et al. (2014) implemented a small zone around the source in which strong S-wave damping is imposed. In our implementation, we attempt to attenuate the excited S-waves using Laplace-domain damping (as described in Section 2.2), and by applying an anisotropic weighting term to the source region, as suggested by Operto et al. (2009).

Model Frequencies	80-1600 Hz
Frequency interval, Δf	80 Hz
Total Modelled Time Interval	12.5 ms (reduced)
Time Sampling, Δt	0.05 ms
Grid Spacing in x-dimension, Δ_x	0.5 m
Grid Spacing in z-dimension, Δ_z	0.5 m
Number of Gridpoints in x-dimension, N_x	131
Number of Gridpoints in z-dimension, N_z	231
Minimum Wavelength, λ_{min}	1.875 m
Reduction Velocity	3000 m/s
Number of Sources	201
Number of Receivers	201
Source-receiver spacing	0.5 m

Table 3.1: Waveform modelling parameters for the synthetic crosshole study.

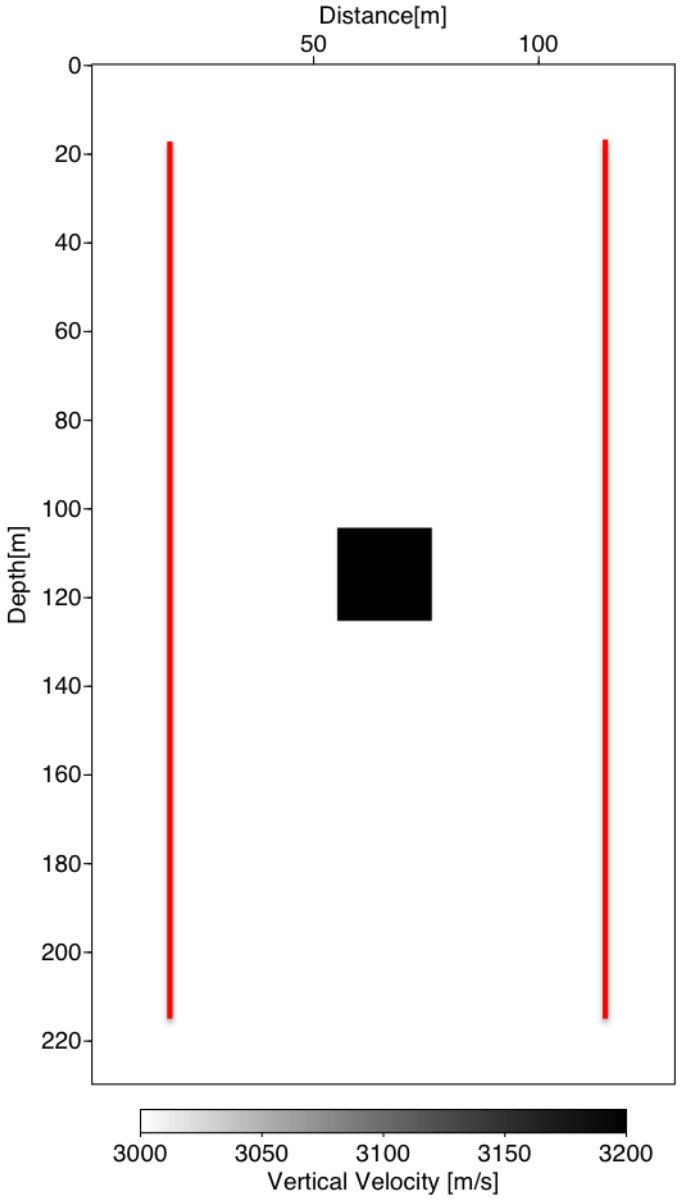
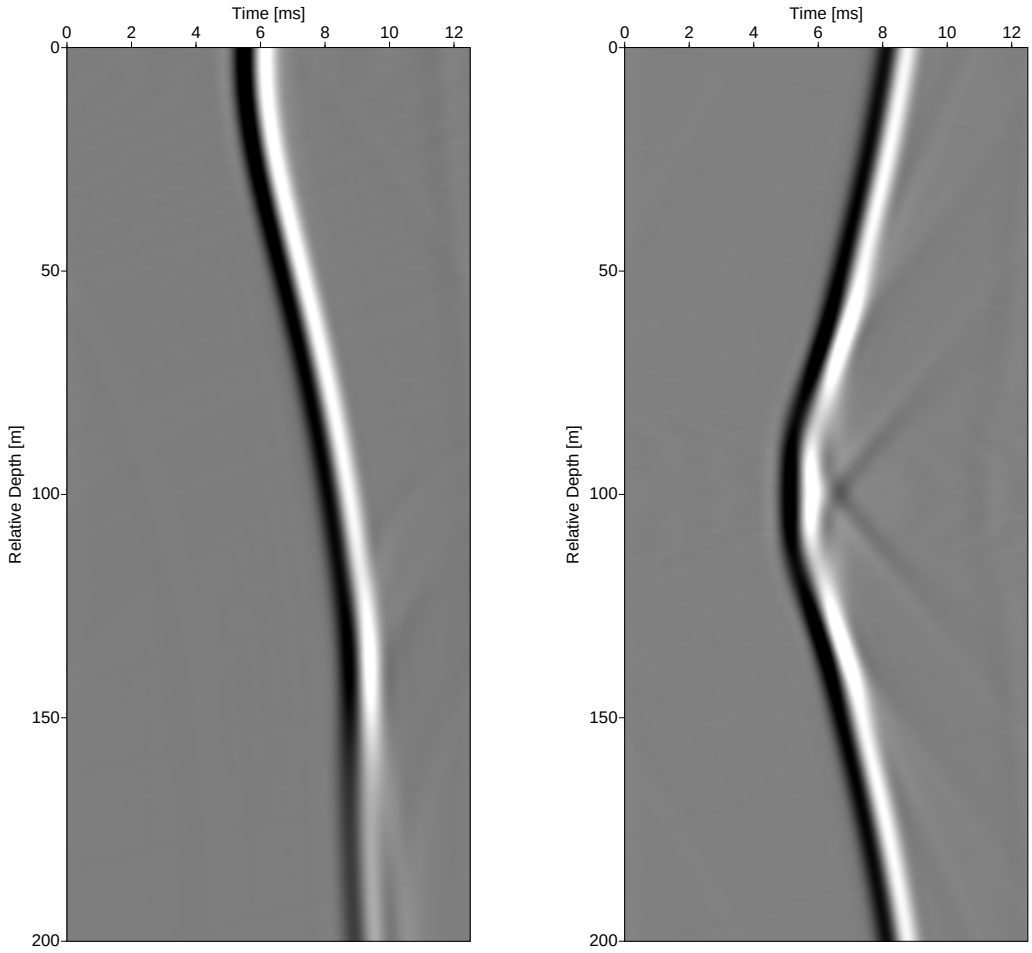


Figure 3.1: True Vertical Velocity Model for the Synthetic Crosshole Study. The source and receiver geometry is shown in red.



(a) Shot Depth = 20m

(b) Shot Depth = 120m

Figure 3.2: Forward modelled shot gathers with the true models. Note the presence of diffractions off the central anomaly in panel (b), as well as the absence of S-wave artifacts within the modelled time interval. The data are modelled in reduced time with a reduction velocity of 3000 m/s.

Following traveltime picking, I apply the anisotropic traveltime tomography method described in Chapter 2.2.1 to recover best-fitting models of V_V, ϵ and δ. For traveltime tomography I use the unpadded model (100 m by 200 m) as raytracing is robust with regards to boundary conditions. The final models are then interpolated onto a grid that is consistent with the padded dimensions of the starting models for FWI. Finally, I perform monoparameter FWI to update the starting V_V model only, keeping the ϵ and δ models fixed. FWI results will be shown for three separate cases, each using a different strategy for the anisotropy models. These three cases are as follows:

1. FWI using the anisotropy models as recovered from traveltime tomography (unaltered)
2. FWI using the anisotropy models recovered from traveltime tomography with a spatial Guassian smoothing filter applied
3. FWI using the 'true' anisotropy models

The first case can be viewed as the 'optimistic' approach, as I am using the anisotropy models as recovered from traveltime tomography. Several studies related to anisotropic FWI suggest smoothing of the anisotropy models, or imposing smoothing during the preliminary inversions of the anisotropy parameters as a candidate for improving velocity model convergence (Operto et al., 2009; Gholami et al., 2013a; Alkhalifah and Plessix, 2014). This is due to the fact that, despite the efforts to improve modelling and inversion algorithms, the simultaneous recovery of at least three parameters is still extremely challenging for anisotropic FWI. Thus, the second case can be viewed as the 'realistic' approach. For field data, the true anisotropy models are not known, therefore I perform inversions using the true anisotropy models solely to obtain a qualitative estimate of the success of the first two inversion cases. In other words, the final case should only be viewed as a benchmark for the results obtained from the previous two.

3.2 Traveltime Tomography

Table 3.2 summarizes the survey specifications for traveltime tomography. Following the suggestions of Chapman and Pratt (1992) and Pratt and Chapman (1992), the inversion strategy was as follows: I performed several initial tests using a wide variance of the three regularization parameters described in Section 2.2.1 (β, ν and ε), as well as for different starting velocity models to observe the tradeoffs between vertical velocity and ϵ. Based on empirical observations of the resulting tomograms, the range of the parameters was then reduced to encompass the 'optimal' values for each. All traveltime inversions were conducted using 400 iterations of the LSQR solver, which were computed quickly due to the relatively small size of the synthetic

survey. It is common during the traveltime tomography process to re-trace the raypaths after initial updates to the model parameters. During the preliminary testing stages, this step was performed however it was eventually removed from the inversion workflow as it did not appear to be providing substantial improvements to the raypaths (which was most likely due to the inherent simplicity of the true models).

Number of grid points in x-dimension	66
Number of grid points in z-dimension	116
Grid size in x-dimension [m]	2
Grid size in z-dimension [m]	2
Number of sources	130
Number of receivers	130

Table 3.2: Traveltime tomography model parameters for the synthetic crosshole study.

Table 3.3 shows the combinations of parameters tested in the initial inversion passes, and Figure 3.3 shows the resulting tomographic models for horizontal velocity, V_H. Note here that the V_V model is obtained by dividing the V_H model obtained from traveltime tomography by the corresponding ϵ values. Let us discuss the choices made regarding the ranges of regularization parameters that were tested. Recalling that ν penalizes large deviations from the starting model, I only test negligible ($\nu = 0.001$) values to allow the inversion to deviate significantly from the starting model if required. The reason for this is that traveltime tomography is often the first step in the model building process, therefore the starting model chosen is simply a 'best guess' as to what the true values might be. Next, I test two very different values of β; the first ($\beta = 0.001$) places little penalty on anisotropic solutions, whereas in the second case ($\beta = 1.0$), I am effectively forcing the tomographic algorithm to obtain an isotropic solution (by placing a strong penalty on anisotropic solutions). As shown in both the traveltime residuals as well as by visual inspection of the tomographic models, the anisotropic solutions are superior which is expected for these data (I show the results of $\beta = 1.0$ for illustrative purposes only). Finally, testing of the 'roughness' or ε parameter requires consideration of the predetermined knowledge that anisotropic FWI favors smooth anisotropy models to improve model convergence, but also the *a priori* knowledge that a velocity anomaly exists near the centre of the model. Therefore, I investigate ε values that provide relatively smooth anisotropy models (but also acceptable traveltime residuals) while still resolving the anomaly present within the velocity solution.

From this analysis, the optimal suite of parameters was determined to lie somewhere between the values for panels (b) and (c) in Figure 3.3. A final inversion pass was then conducted to obtain the best-fitting models of V_V, ϵ and δ for anisotropic FWI. Table 3.4 shows the optimal traveltime parameters, and Figures 3.4 and 3.5 show the models obtained from traveltime

tomography for velocity and Thomsen's anisotropy parameters respectively, together with the true velocity model alongside the recovered model for comparison.

Panel	ν	ε	β	$E(\mathbf{m})$ [ms]
a	0.001	0.01	0.001	0.0146
b	0.001	0.3	0.001	0.0174
c	0.001	2.0	0.001	0.0256
d	0.001	0.01	1.0	0.1885
e	0.001	0.3	1.0	0.3346
f	0.001	2.0	1.0	0.4837

Table 3.3: Traveltime inversion parameters investigated during preliminary inversions for the synthetic crosshole study.

ν	ε	β	$E(\mathbf{m})$ [ms]
0.001	0.400	0.001	0.0196

Table 3.4: Optimal traveltime inversion parameters determined for the synthetic crosshole study.

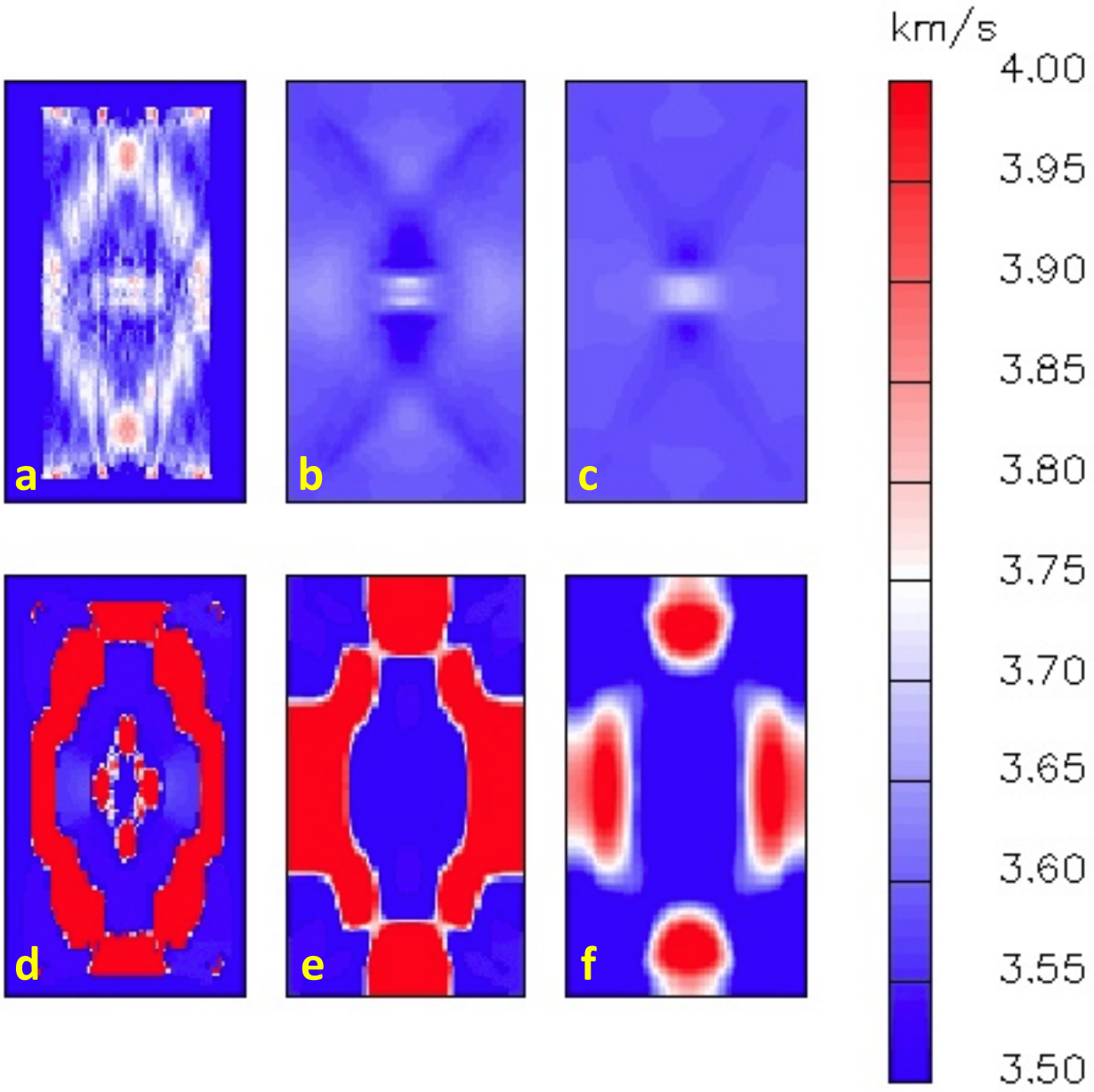
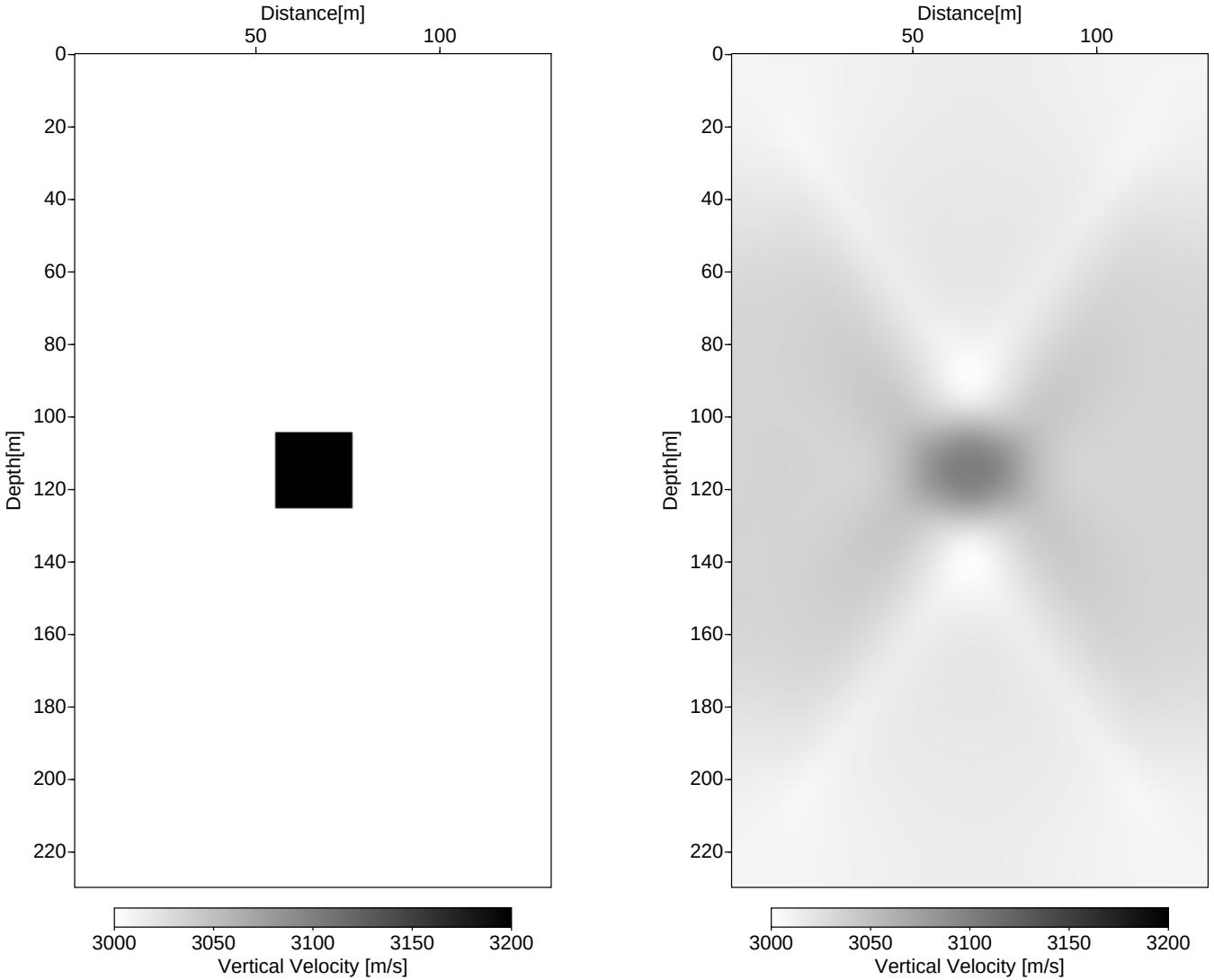


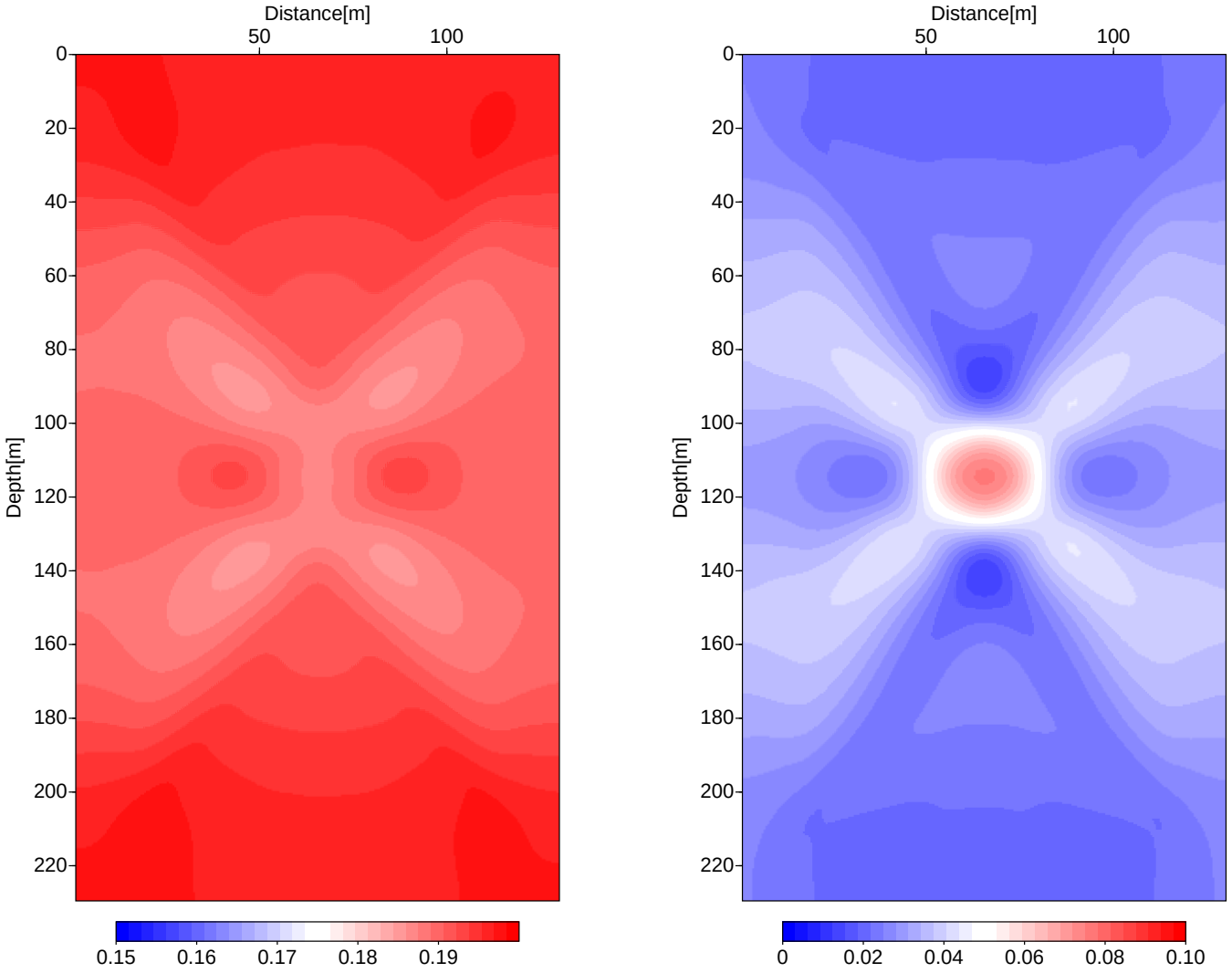
Figure 3.3: Preliminary traveltime tomography results for the parameters summarized in Table 3.3. From left to right: $\varepsilon = 0.01$, 0.3 and 2.0. From top to bottom: $\beta = 0.001$ and 1.0. For all cases, $\nu = 0.001$. Note that these are horizontal velocities (V_H) and the units are in km/s, as these are the defaults for the traveltime tomography software.



(a) True Model

(b) Recovered Model

Figure 3.4: Best-fitting vertical velocity model, V_V, obtained from traveltime tomography for the synthetic crosshole study. The true model, which consists of a background velocity of 3000 m/s with a positive central anomaly of 200 m/s is also shown for comparison.



(a) ϵ (b) δ

Figure 3.5: Best-fitting ϵ and δ models obtained from traveltime tomography for the synthetic crosshole study. It should be noted that the true models are homogenous with ϵ=0.2 and δ=0.1.

The results show that the velocities obtained from traveltime tomography are closer to the true background velocity when compared to the starting model, and the positive anomaly in the centre of the model is now apparent. However, the overall resolution of the anomaly is poor, and there are low-velocity 'lobes' directly above and below the central anomaly. Furthermore, there is a cross-shaped artifact thats extends from the four corners of the model that slightly resembles the shape of the first Fresnel zone. These artifacts can be attributed to the resolution limits of traveltime tomography coupled with the lack of ray coverage near the extremities of the model region as well as directly adjacent to higher velocity regions. If the source and receiver lines were instead oriented in the x-direction, the low-velocity lobes would mirror this transition.

Recalling the values of the true anisotropy models (ϵ = 0.2, δ=0.1), we see that the ϵ model is well-recovered (with values ranging from 0.18 to 0.195) but has been contaminated by the presence of the central anomaly. Comparitavely, the δ model is less resolved and shows more significant contamination from the central anomaly. Gholami et al. (2013b) showed that inaccuracies in the δ model have the least impact on wave propagation accuracy. I chose to accept these results as my starting models for anisotropic FWI on the basis of the accuracy in the final velocity and ϵ models, and I will evaluate whether the inaccuracies in the δ model obstruct the success of FWI.

3.3 Full Waveform Inversion

In this Section I describe the application of FWI for each of the aforementioned subcases of anisotropy models. For this, I use the same inversion strategy in all three cases so that the results are directly comparable. As FWI is an ill-posed and non-linear process, it is possible to take advantage of several regularization strategies to help guide the inversion process. Higher inversion frequencies improve the resolution of the sharper features, but are at higher risk of introducing artifacts due to cycle-skipping. The risk of introducing these artifacts can be mitigated by the initial use of lower frequencies. It has become common practice to adopt a multi-scale approach, whereby preliminary inversions are conducted on a coarse grid or with stronger smoothing while inverting for the lower frequencies. Once a satisfactory model update has been obtained, the model is then interpolated onto a finer grid, and inversions are conducted for the higher frequencies within the data (Bunks et al., 1995; Sirgue and Pratt, 2004). I use a modified multi-scale approach, as I do not refine the model grid size when transitioning to the higher frequencies.

Spatial tapering of the gradient in either direction allows greater or lesser emphasis to be placed on certain regions of the model. One such approach to gradient tapering is the layer-

stripping approach, whereby the model is resolved in sequential layers, typically beginning with the shallower parts and extending to the deeper model layers thereafter (Pratt et al., 1996; Shipp and Singh, 2002; Virieux and Operto, 2009); this is only applicable for surface seismic data, as overburden effects can be isolated using only near offset data. Due to the relative simplicity in the synthetic survey geometry, I do not apply tapering of the gradient to these analyses, however we will see in Chapter 4 that tapering of the gradient is an important regularization tool for real data, especially in the anisotropic case.

The use of offset weighting allows control over which data are input into the inversion process based on their relative travel distance from the source to receiver location. In surface seismic data, the near offsets typically contain elastic noise such as ground roll, while the far offsets carry a greater risk of introducing cycle-skipping (Brenders and Pratt, 2007). In the crosshole case, the offset is defined as the absolute distance between the source and receiver position. As these data comprise transmitted waves, they do not suffer from the presence of surface waves in the near offsets. Therefore, it is important to utilize the full angular coverage of the data as the far offset (wide-aperture) data will yield important information about lateral variations in velocity. For this synthetic test, I do not apply offset weighting to the data as the primary focus is to observe the resolution of the anomaly. We will see in Chapter 4, however, that sequentially increasing the range of data offsets considered in the inversion process is critical to avoid converging to local minima.

There are two aspects of the wavefield that can be taken into account when defining the objective function, E($\mathbf{m}$): wave amplitude and phase. Amplitude modelling in the acoustic framework for FWI studies is somewhat problematic due to the shortcomings of the acoustic model in representing the true earth (Alkhalifah and Plessix, 2014). Furthermore, in acoustic FWI we are only interested in P-waves, but field data contain several other wave types ranging from surface waves and converted waves in the surface reflection case to tube waves for the crosshole case. As such, an important step in waveform preprocessing is the removal of these sources of noise by applying high- or low-pass filters and/or mutes to the traces, such that a small window containing only the P-wave energy remains. This is significant for amplitude modelling, as wave amplitudes are almost certainly affected by these pre-processing steps. Nevertheless, Brenders and Pratt (2007) formulate their objective function for acoustic isotropic FWI with both wave amplitude and phase by applying amplitude scaling to the preprocessed waveforms. The incorporation of anisotropy adds another layer of complexity to this problem. Operto et al. (2009) concluded that it is difficult to quantify the level of amplitude mismatch with their modelling method. Plessix and Cao (2011) state that although the acoustic wave equation for anisotropic media is kinematically accurate for P-wave modelling, it is impossible to interpret the dynamic behaviour of the waveforms as it describes a non-physical

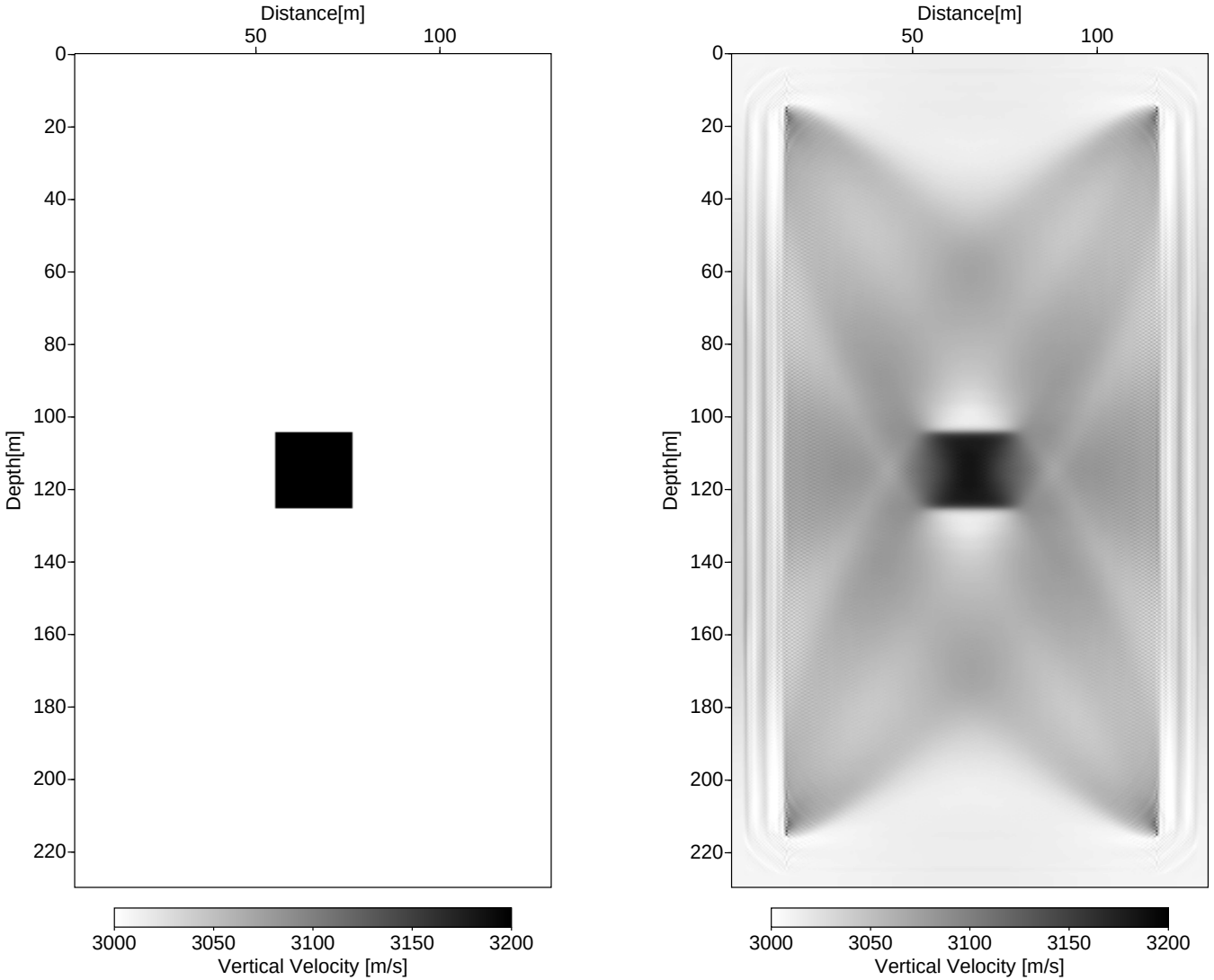
process. Therefore, I follow Kamei et al. (2014) and formulate the objective function using the phase-only approach.

The final preconditioning feature for discussion is that of wavenumber filtering. Wu and Toksöz (1987) showed that the maximum wavenumber spectral coverage, k_{max}, is theoretically bounded by the minimum observed velocity, v_{min}, as well as the maximum model frequency, f_{max} ($k_{max} = 2f_{max}/v_{min}$). It is therefore possible to filter out unstable high wavenumber components by applying a spatial low-pass filter in the wavenumber domain based on the maximum expected wavenumber per block of inversion frequencies; this is not a major concern for the synthetic data case, but we will see the importance of this preconditioning operator in Chapter 4.

Table 3.5 outlines each of the frequency blocks selected for anisotropic FWI. Inversion frequencies were chosen to cover the frequency range in the data. For each block of frequencies, I use a total of 5 FWI iterations. It is common to utilize a line search method for the first iteration of FWI to reduce the risk of converging to a local minimum. Line searches require multiple forward modelling runs, and are therefore expensive when compared to linear estimates. As the overall scale of the synthetic test is relatively small, I accept the increase in computational cost by applying a line search method for the first iteration of each frequency block only. Figures 3.6, 3.7, and 3.8 show the final FWI model of vertical velocity for the three subcases respectively.

Frequency Block	Inversion Frequencies [Hz]
1	80, 240, 400
2	560, 720, 880
3	1040,1200, 1360

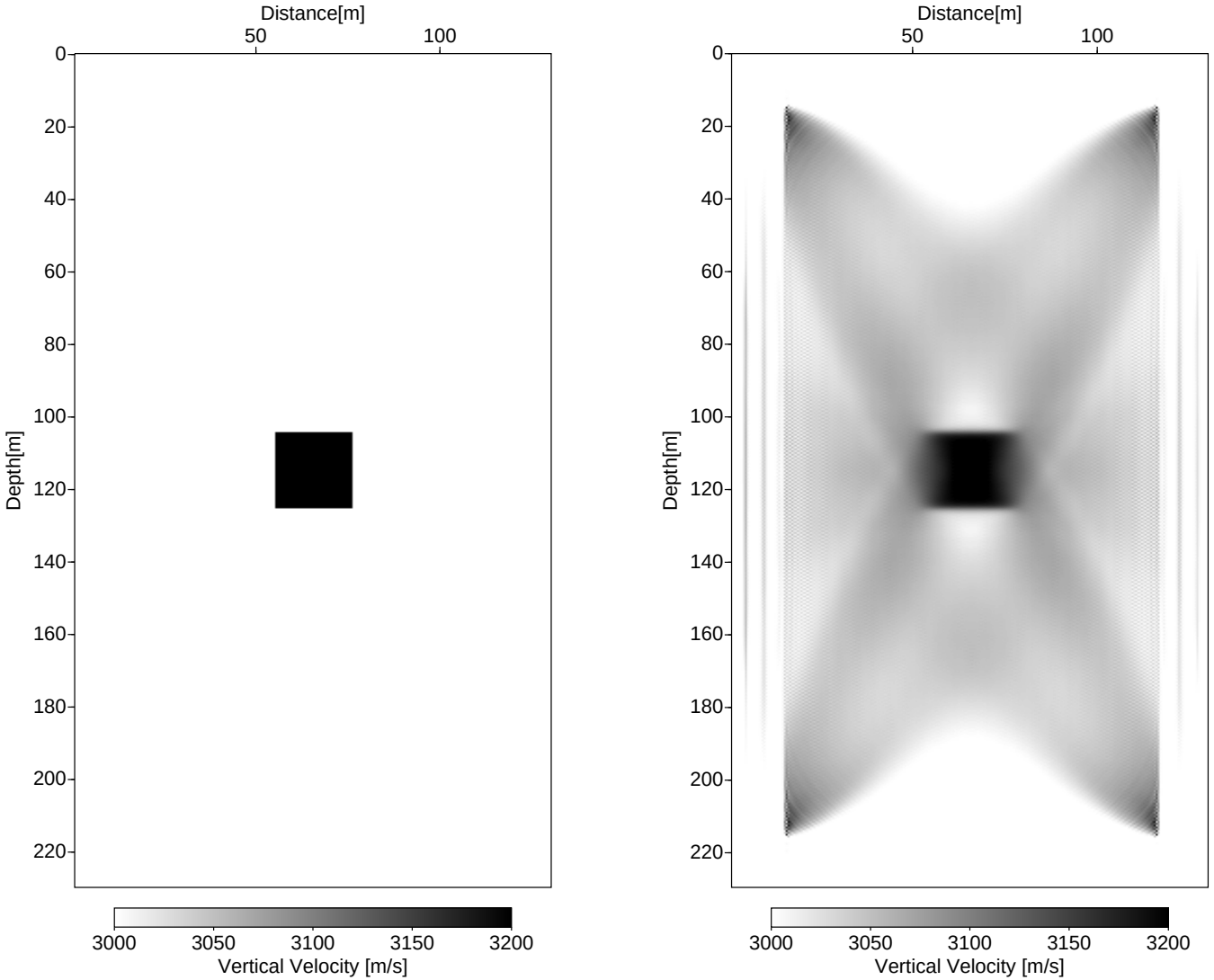
Table 3.5: Selected inversion frequencies for anisotropic FWI of the synthetic dataset.



(a) True Model

(b) FWI Model

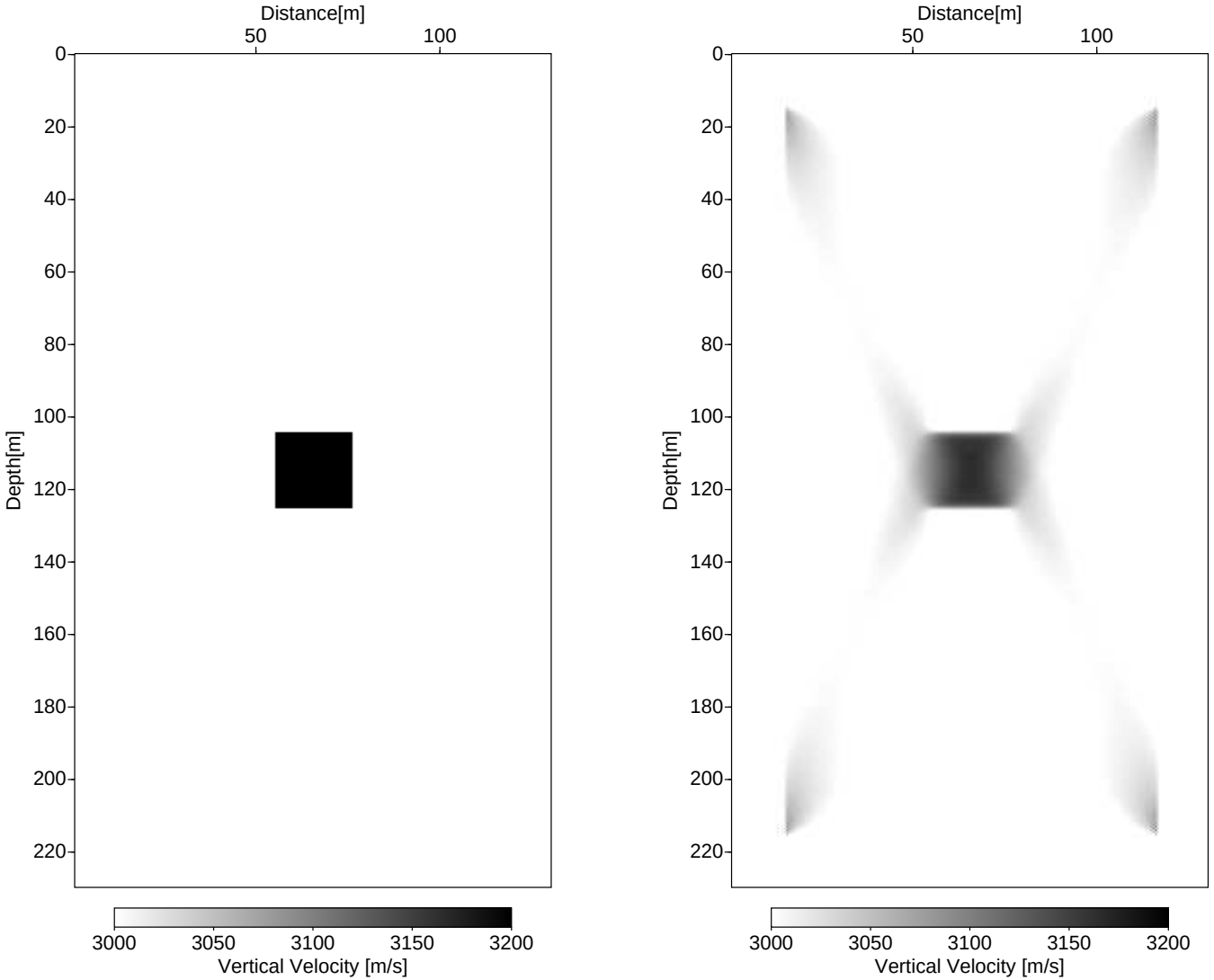
Figure 3.6: Final V_V model obtained from FWI using the δ and ϵ models as recovered from traveltime tomography. The true velocity model is also shown for comparison.



(a) True Model

(b) FWI Model

Figure 3.7: Final V_V model obtained from FWI using the δ and ϵ models with a Gaussian smoothing function applied. The true velocity model is also shown for comparison.



(a) True Model

(b) FWI Model

Figure 3.8: Final V_V model obtained from FWI using the true δ and ϵ models. The true velocity model is also shown for comparison.

The FWI results show that the algorithm was successful in delineating the outer edges of the velocity anomaly, as they are much closer to the true model than the traveltime results for all three cases. It is important to note here that the top and bottom edges of the anomaly are more clearly defined than the lateral edges, which is a direct consequence of the survey geometry. There are some minor artifacts present, such as those proximal to the source and receiver boreholes as well as the cross-shaped pattern introduced by traveltime tomography. The primary objective of this study was to see whether or not the anomaly could be clearly defined, and I conclude the presence of these artifacts to be less critical. I suggest that further inversions with some of the regularization features discussed above could aid in the removal of these features.

It is immediately obvious that the resolution of the central anomaly is improved when the starting anisotropy models are smoothed, as well as when the true anisotropy models are utilized. The latter is an obvious result, nevertheless the velocity model obtained using the true models highlights the expected tradeoffs in resolution between the velocity and anisotropy parameters, as the velocities recovered from this approach are closest in value to the true velocities used in forward modelling. The decision to smooth the anisotropy models obtained from traveltime tomography paves the path for an interesting discussion. In this case, we had *a priori* knowledge that the true anisotropy models were smooth, however for field data this information may not be available. For the FWI results shown in the next chapter, significant smoothing constraints were imposed during the traveltime inversions. This is both a fundamental limitation of anisotropic traveltime tomography, which requires strong smoothing regularization to achieve the desired parameter resolution (Pratt and Chapman, 1992), as well as a limitation of the FDFD modelling technique that I have adopted.

I have shown in this Chapter that the anisotropic traveltime tomography method developed by Chapman and Pratt (1992) and Pratt and Chapman (1992) is a suitable technique for generating starting models of velocity, ϵ and δ for anisotropic FWI. These FWI results serve as a cross-verification that the FDFD implementation is working as intended, as it is consistent with the well established traveltime tomography method. The implementation is compatible with the inversion structure and methods of Pratt et al. (1998). The comparison of the velocity models recovered for each of the three subcases highlights the expected tradeoffs between anisotropy and velocity, but also showcases the success of the inversion algorithm in recovering the structure of the true anomaly. I propose that the results of this work indicate that the joint methodology is a valid approach for TI media, therefore application to a field dataset is a reasonable next step.

Chapter 4

Field Data Case Study - Western Canada Crosshole Survey

In this Chapter I apply the joint techniques of anisotropic traveltime tomography and anisotropic FWI to field data from a crosshole survey located in Western Canada. These data were provided on the condition that the true depths and formation names be withheld. While the focus of Chapter 3 was to test the compatibility of the two methods and to observe the tradeoffs in resolution, the primary objectives of this study were to develop an efficient FWI strategy for anisotropic media, as well as to directly quantify the benefits of accounting for anisotropy in the modelling process. These data were previously processed by Pratt et al. (2008), however they used the isotropic modelling operator proposed by Jo et al. (1996) and then performed a coordinate stretch on the resulting wavefield to simulate 1-D elliptical VTI, after Dellinger (1991). Here, I use a more rigorous approach for handling anisotropy with the hope of improving the FWI results.

I will first describe the details of the survey including the objectives, the geological background as well as the survey parameters. Next, I will summarize the data preprocessing performed by Pratt et al. (2008), their strategies for traveltime tomography, and present the best-fitting starting models chosen for FWI. I will then discuss the FWI strategy that I developed which makes use of several of the regularization strategies discussed in Chapter 3. I apply this inversion strategy to the field data twice: once using the anisotropic approach described in Section 2.3 and then again with the elliptical isotropy approximation of Pratt et al. (2008). From here on in, I will refer to these two approaches as 'Anelliptical FWI' and 'Elliptical FWI' respectively. The final FWI models will then be directly compared to determine the benefits of properly accounting for anisotropy in this dataset, as well as to assess the overall success of applying FWI to these data.

4.1 Introduction

The crosswell seismic survey was conducted in Western Canada across finely layered sediments to identify the structures present within local sandstone reservoirs. The primary objective was to fully characterize the reservoir architecture by differentiating sandstone channels from impermeable shale barriers through expected contrasts in seismic velocity, and to use the velocity model obtained as an input for reservoir simulation and drainage area calculations. Data were collected using a piezoelectric source for a sweep of frequencies ranging from 100 to 2000 Hz across source and receiver boreholes spaced 160 m apart. The survey was conducted at a depth of over 2 km, with a source and receiver spacing of 1.5 m spanning the total survey depth of 420 m. In each borehole, a series of petrophysical logs were also provided (Pratt et al., 2008).

These data were selected for the application of anisotropic FWI for several reasons. The expected lithology consists of finely-layered sedimentary rocks (sandstones, shales and siltstones) and correlation of the formation tops interpreted from the geophysical well logs in each of the boreholes suggest that they are laterally continuous and horizontal to sub-horizontal in nature; these together suggest that the region can be well represented by VTI symmetry. Furthermore, these data are challenging as there is a low signal-to-noise ratio (SNR) and the gathers are plagued by the presence of artifacts such as 'tube waves', which travel within the boreholes before propagating through the rockmass or vice-versa. These data will therefore provide a good benchmark for the anisotropic waveform tomography (AWT) implementation. The final reason is that these data were previously analyzed by Pratt et al. (2008) using an elliptical isotropy model, however the presence of artifacts in their final models, as well as a significant mismatch observed between their predicted gathers and the field gathers at the furthest offsets indicated that a more complete anisotropy model was required. This provides further motivation for the use of the more generalized anisotropy approach to determine if an interpretable result is achievable for this case study.

There are 4 sedimentary formations of interest throughout the surveyed region, referred to as 'Formation A','Formation B','Formation C' and 'Formation D'. Figure 4.1 shows the sonic and gamma ray logs from the source and receiver boreholes with the interpreted formation tops and thicknesses. Based on the variations in the sonic logs, Formation A is expected to have more homogenous velocities, while the other three units appear to be more heterogenous. Furthermore, the gamma logs suggests that Formation A is most consistent with a sandstone lithology while the other three formations appear to be more shale-rich. For these reasons, Formation A is believed to be the targeted reservoir of this crosshole survey. It is important to note that there is an unconformity that lies at the boundary between Formations A and B,

which correlates to a change in dip structure from near horizontal to gently dipping (Pratt, 2007). Figure 4.2 shows every 50th ray obtained from raytracing in the starting model using the method described in Section 2.2, plotted with the sonic logs for comparison. Note here the dense ray coverage within the target reservoir, but also the sparsity in coverage in the regions directly above and below the reservoir.

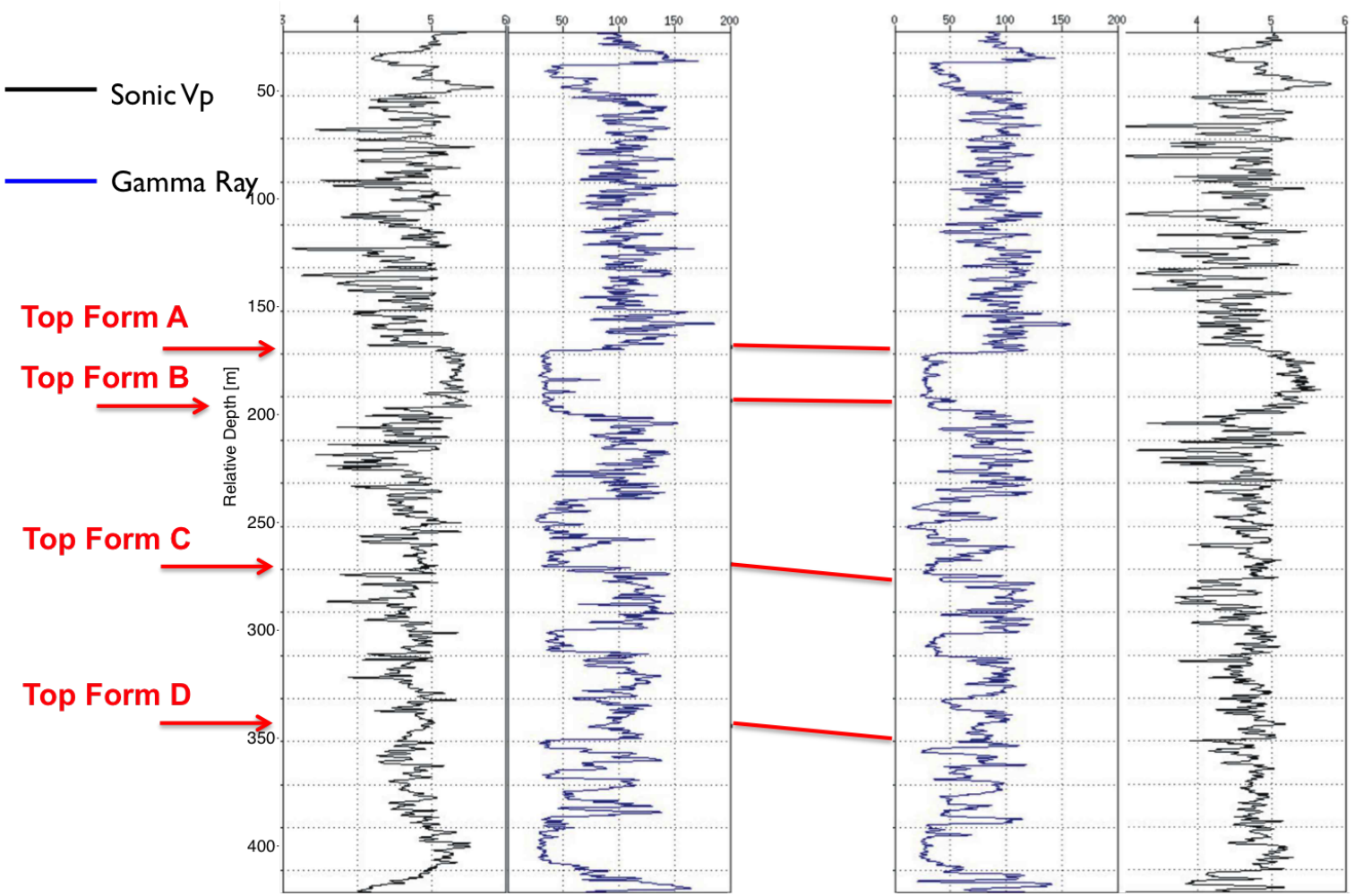


Figure 4.1: Borehole sonic (black) and gamma ray (blue) logs for both the source (left) and receiver (right) boreholes. The interpreted formation tops and thicknesses are overlaid in red. The distance between the two wells is approximately 160 m. Original image modified from Pratt et al. (2008).

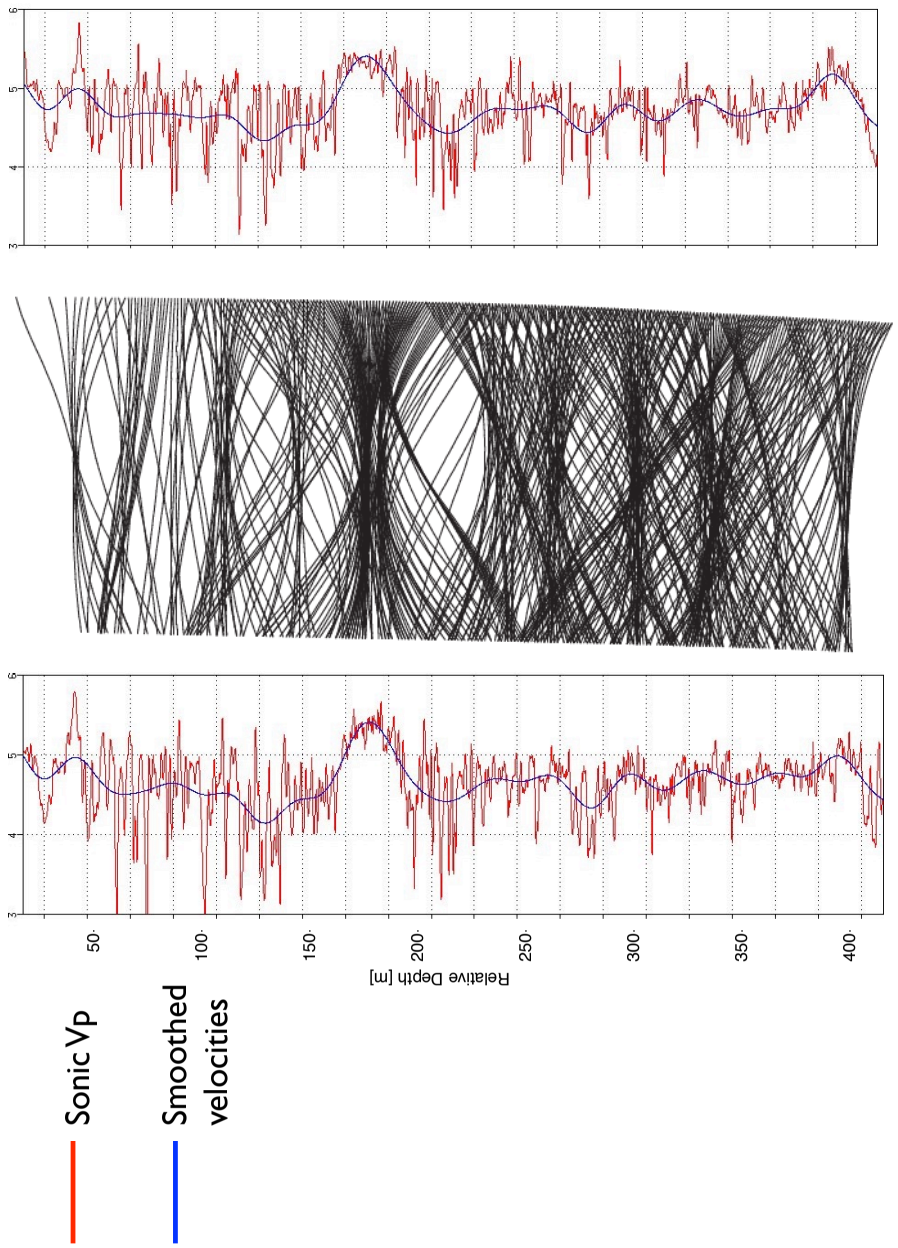


Figure 4.2: Every 50th ray traced in the starting model, plotted with the sonic velocities for comparison. The distance between the two wells is approximately 160 m. Original image modified from Pratt et al. (2008).

4.2 Traveltime Tomography

In this section I will briefly describe the anisotropic traveltime analyses conducted by Pratt (2007) and Pratt et al. (2008) for these data. They noted a marked increase in the SNR with depth, which corresponded to an increase in the number of pickable traces at larger offsets for the deeper shot and receiver positions. From a possible 45,000 traces, approximately 25,000 traveltime picks were made, ranging from ±60 m vertical offset for the shallower parts to as much as ±120 m offset in the deeper regions of the model. Although the traveltime inversion scheme is capable of handling general 2-D anisotropy systems, they imposed VTI symmetry based on the aforementioned *a priori* information (Pratt, 2007).

The starting model for vertical velocity was obtained by smoothing the sonic logs in each borehole, followed by lateral extrapolation and interpolation of the two 1-D profiles to generate a 2-D model. As in Chapter 3, the first stage of inversion was concerned with testing the various regularization parameters and observing the resulting tomograms and traveltime residuals. From visual inspection of the handpicked traveltimes, Pratt (2007) estimated the traveltime errors to be within 0.025 ms. Once an optimal set of regularization parameters were chosen, Pratt (2007) performed two further passes of traveltime inversion by first re-calculating the raypaths using the bending method of Um and Thurber (1987) and Davison (1991). Table 4.1 summarizes the optimal parameters selected for traveltime tomography (see Chapter 2 for a review of these parameter definitions), and Figure 4.3 shows the best-fitting models of vertical velocity, δ and ϵ obtained from traveltime tomography.

ν	ε	β	$E(\mathbf{m})$ [ms]
0.001	0.100	0.001	0.074

Table 4.1: Optimal traveltime inversion parameters determined for the Western Canada cross-hole survey.

Some comparisons can be drawn between the optimal parameters selected by Pratt (2007), and those presented in Chapter 3. We see that little penalty was placed on models that are different from the starting model (ν is small), as the interpolated starting model is only a representation of the rockmass directly adjacent to the boreholes and does not consider any lateral variations throughout the surveyed region. Becuase the formations were expected to be anisotropic in nature, the anisotropy penalty (β) was also relaxed to allow the inversion algorithm to explore anisotropic solutions. Finally, the roughness penalty (ε) was relaxed until the residuals were within the estimated traveltime errors.

The observed contrasts in seismic velocity correlate reasonably well with the major formations. Formation B appears to consist of geological units with both low ($\approx$ 4000 m/s) and

high ($\approx$ 5500 m/s) velocity regions, which is consistent with an interbedded shale and sandstone lithology, whereas Formation A mainly exhibits higher velocities, which is consistent with a massive sandstone lithology (Han et al., 1986). When comparing the anisotropy models to the well logs, correlations between the Thomsen's parameters and either velocity or shale content is more difficult. The overall trend of the ϵ model is reasonably consistent with the lithology, with the largest observed values (10 to 15 percent) proximal to the reservoir region. On the other hand, the δ model is more variable, which is consistent with the results obtained in Chapter 3. It is also important to note that the anellipticity (the difference between ϵ and δ) is greatest proximal to the reservoir region, which will be significant when we analyze the elliptical and anelliptical FWI results in the next section.

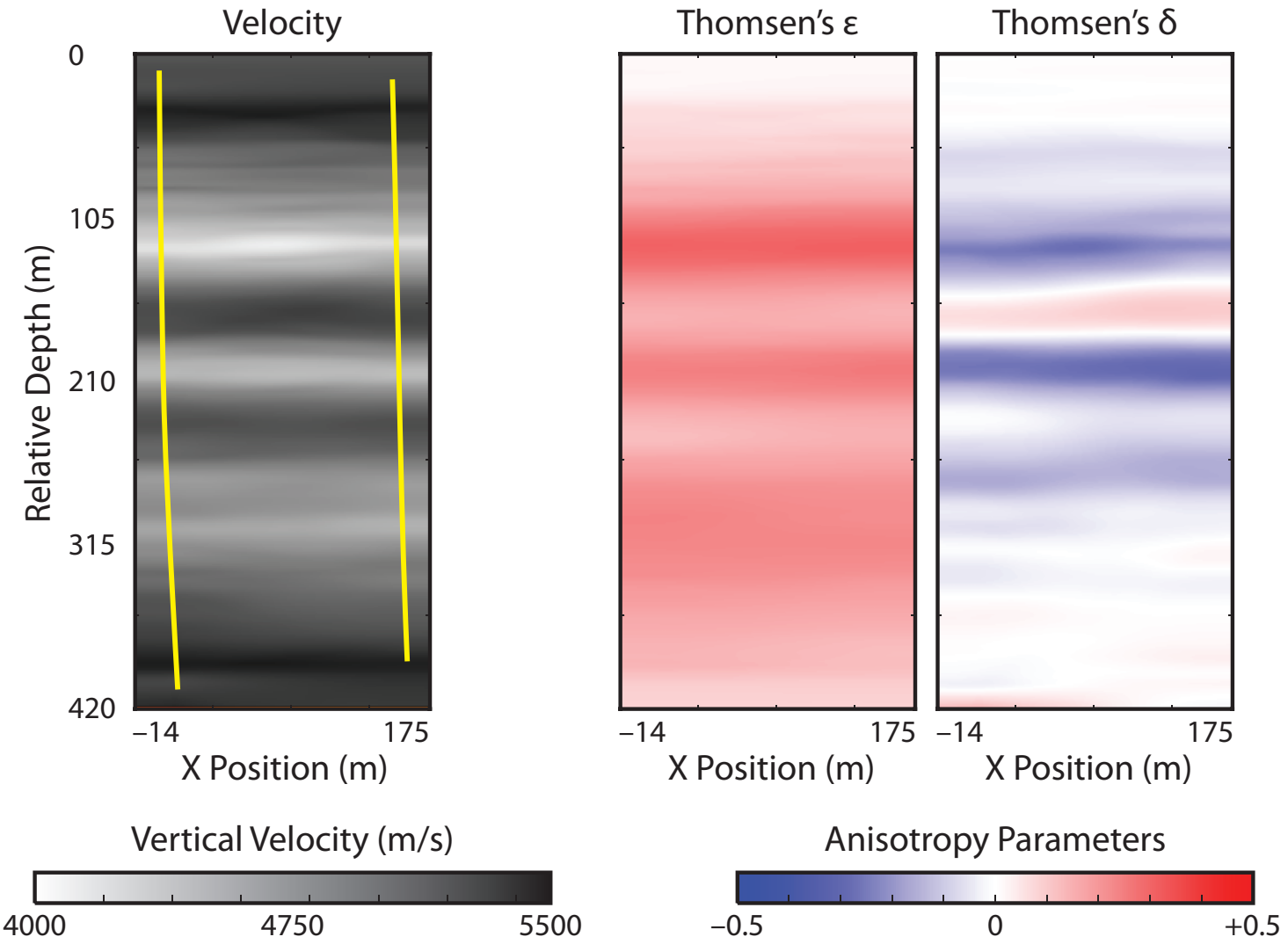


Figure 4.3: Best-fitting V_V, ϵ and δ models obtained from traveltime tomography by Pratt et al. (2008) for the Western Canada crosshole dataset. Source and receiver borehole locations are overlaid in yellow.

4.3 Full Waveform Inversion

In order for acoustic FWI to be successful at inverting field data, stringent preprocessing techniques must be applied to the data to remove any noise or artifacts that contaminate the pressure wave arrivals. For the crosshole case, we are only interested in modelling the transmitted P-waves which depart from the source and travel directly through the medium to the receiver. I will briefly summarize the preprocessing techniques performed by Pratt et al. (2008), as I have adopted the preprocessed dataset for my inversions. First, an f-k filter was applied to remove the tube waves which are easily characterized by their higher amplitudes, slower speeds and near-linear moveout. Next, top and bottom mutes (removal of all energy before or after a certain point in time) were applied such that the P-wave arrivals were contained within a 10 ms window. To achieve this, all energy before the picked traveltimes was removed, followed by the removal of any energy arriving later than 10 ms after the first arrival. Traces that did not contain a traveltime pick were 'killed' or disregarded for the inversion process (Pratt et al., 2008). Figure 4.4 shows the application of these preprocessing techniques to field gathers for sources located at depths of 90 m and 275 m, with an overlay in red of the handpicked traveltimes for reference. Note the marked increase in the SNR of the deeper shot gather (panels (c) and (d)).

The selection of an appropriate FWI approach for field data must be coupled with the selection of appropriate modelling parameters, as there are no 'true' models to help guide these decisions. Incorporating higher frequencies earlier on allows for greater resolution in the preliminary inversions but requires lowering the grid interval. For large field datasets, such as the one described here, computation costs can be mitigated by first inverting a smaller subset of the data. For this reason, I adopt a variant of the multi-scale approach by performing a preliminary stage of FWI, referred to from here on out as 'Stage 1 Inversion', using a filtered subset of the field gathers that contains only frequencies up to 1000 Hz. This data subset was generated by the application of a high-cut Ormsby minimum phase filter to obtain data that contained little energy above 800 Hz. Once an acceptable result was obtained, I performed a second pass of inversion that utilized the full frequency spectrum (up to 2000 Hz), referred to as 'Stage 2 Inversion'.

Tables 4.2 and Table 4.3 shows the selected survey parameters for Stage 1 and Stage 2 inversions respectively. The grid intervals were selected to allow the smallest useable wavelength to be represented across a minimum of four grid nodes, and the minimum expected velocity was estimated from the starting model (Figure 4.3) as well as from the sonic logs. I model in reduced time in order to realize the significant reductions to forward modelling costs; a reduction velocity of 4000 m/s was used. Finally, the time damping factor was estimated

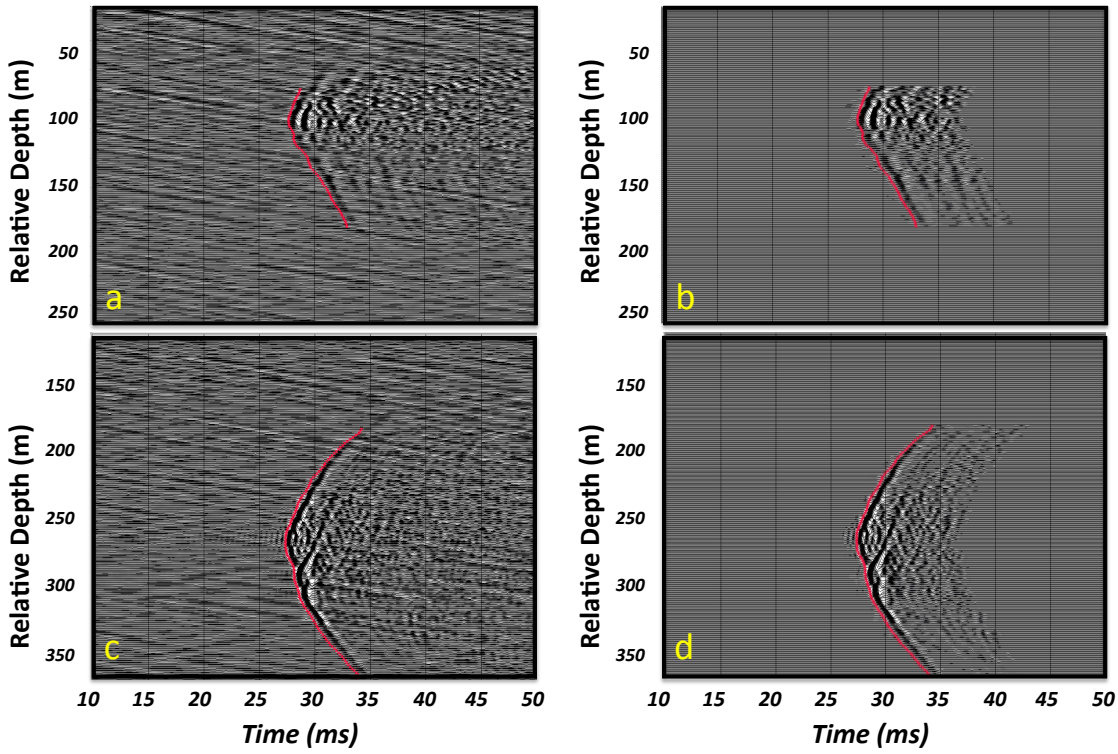


Figure 4.4: Left: Shot gathers for sources located at depths of 90 m (a) and 275 m (c) before preprocessing. Right: Shot gathers for sources located at depths of 90 m (b) and 275 m (d) following the suppression of tube wave energy and time windowing of the data guided by the handpicked traveltimes, which are shown in red. Note that much of the tube wave energy has been suppressed, with the exception of the highest frequencies. For these images, reduced time modelling was not used. Original image modified from Pratt et al. (2008).

from the maximum modelled time ($\tau = 0.4\ t_{max}$) which, after bottom muting was applied, was approximately 0.02 s in both cases.

Figure 4.5 shows a time-domain 'snapshot' as well as a shot gather for a source located at a depth of 170 m in the FWI starting model for each modelling technique. Note the differences in the propagating wavefront for each case, as well as the presence of the spurious S-waves in the anelliptical case. The manually picked traveltimes are overlaid for each shot gather, and clearly show that the anelliptical method better matches the picks in the starting model, which is a further indication that anisotropic FWI will lead to an improved result.

Parameter	Value for Preliminary Inversions
Model Size in the X direction, L_x [m]	188
Model Size in the Z direction, L_z [m]	420
Maximum Model Frequency, f_{max} [Hz]	1000
Minimum Velocity, c_{min} [m/s]	4000
Minimum Wavelength, λ_{min} [m]	4
Implemented Grid interval [m]	1
Number of Grid Points in the X Direction	189
Number of Grid Points in the Z Direction	421
Maximum Modelled Time, t_{max} [s]	0.02
Frequency Interval, δf [Hz]	50
Number of Available Frequencies N_f	20
Anti time-aliasing time domain damping, τ [s]	0.008

Table 4.2: FWI modelling parameters for Stage 1 inversion.

Parameter	Value for Sucessive Inversions
Model Size in the X direction, L_x [m]	188
Model Size in the Z direction, L_z [m]	420
Maximum Model Frequency, f_{max} [Hz]	2000
Minimum Velocity, c_{min} [m/s]	4000
Minimum Wavelength, λ_{min} [m]	2
Implemented Grid interval [m]	0.5
Number of Grid Points in the X Direction	379
Number of Grid Points in the Z Direction	441
Maximum Modelled Time, t_{max} [s]	0.02
Frequency Interval, δf [Hz]	50
Number of Available Frequencies N_f	40
Anti time-aliasing time domain damping, τ [s]	0.008

Table 4.3: FWI modelling parameters for Stage 2 inversion.

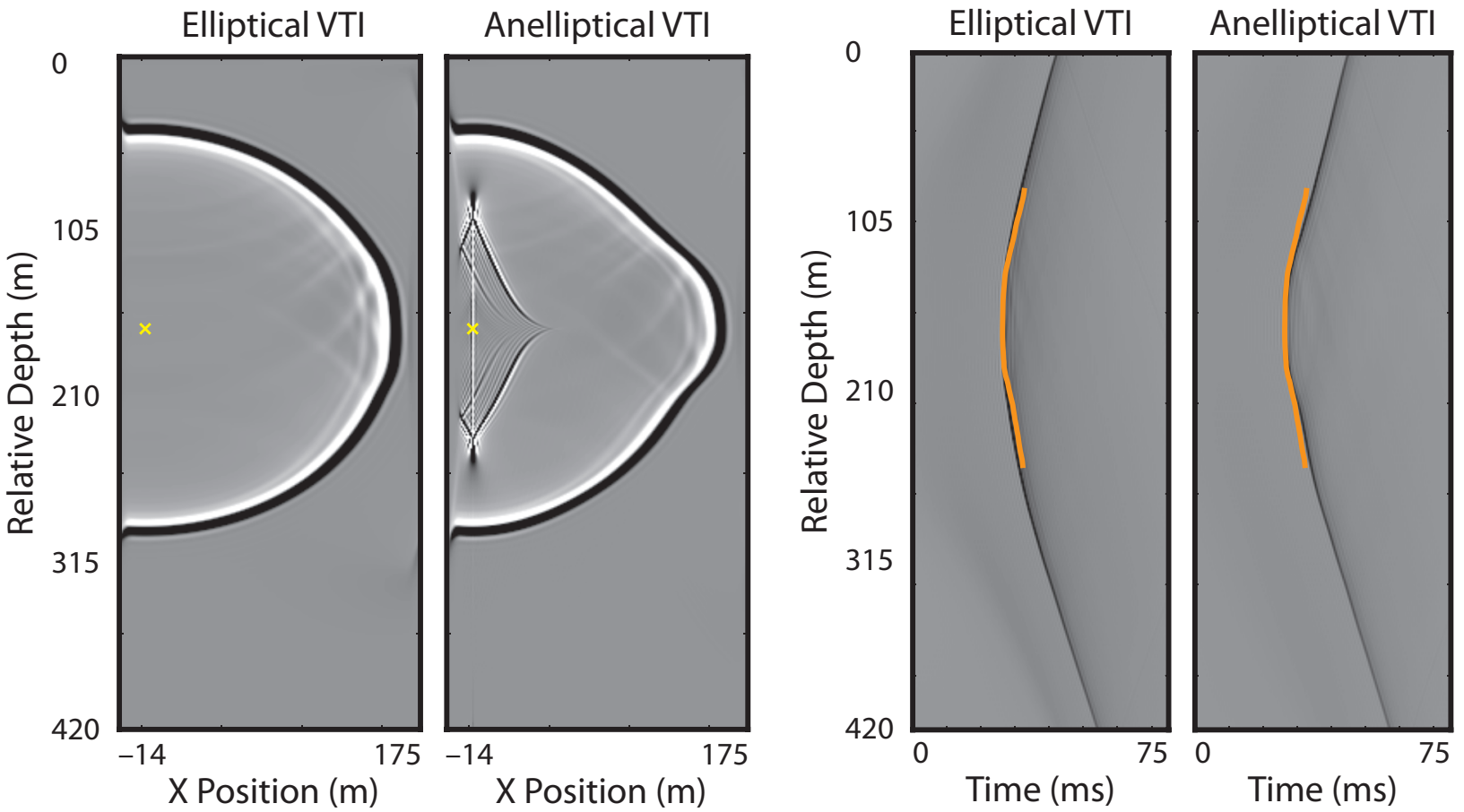


Figure 4.5: Left: Time-domain snapshots (t=0.025 s) of wave propagation for the elliptical (left) and anelliptical (right) starting model. The source is located at a depth of 170 m and is shown as a yellow cross on the panels. Note the presence of spurious S-wave energy in the anelliptical wavefield. Right: Shot gathers for the elliptical (left) and anelliptical (right) starting model. The traveltime picks are overlaid in orange. The source location is the same as for the snapshots. For these images, reduced time modelling was not used.

I performed many inversions testing different combinations of regularization parameters and techniques. The primary objective of this was to develop an efficient strategy for anisotropic acoustic FWI of crosshole data. I believe that the strategies presented here can be extended to future case studies provided that the survey region can be well represented by VTI symmetry. I will first discuss each of the parameters that I investigated, as well as the final inversion strategy. The results shown thereafter are obtained using the final approach, which was applied to both the elliptical and anelliptical cases. This strategy is based solely on the results from anisotropic inversions, therefore I am assuming that it is effective for the elliptical case as well.

Let us first discuss spatial weighting of the gradient function. Several preliminary inversions were conducted without tapering of the gradient. These results were reasonably consistent in terms of structure, however they were plagued with high velocity 'lobes' proximal to the boreholes, as well as near the very top and bottom of the model. These high velocity lobes were especially prevalent in the anelliptical case, with some predicted velocities as high as 7000 m/s, which is most likely related to the sensitivity of the velocities to the ϵ values. High velocities zones proximal to the boreholes is a common observation in crosshole FWI and is thought to be related to the amplitude mismatches between the predicted and real data. The high velocity regions near the top and bottom of the model are visible in the starting model, and are most likely a consequence of the lack of ray coverage for these regions. Since the primary targets of this survey are outside of these regions, I apply a tapering function to the gradient near the top and bottom of the model, as well as to the regions proximal to the borehole in order to obtain accurate velocity estimates for the target formations.

For wavenumber filtering, I use the filter values suggested by Brenders and Pratt (2007), who set the maximum vertical wavenumber equal to the maximum expected wavenumber ($k_{zmax} = k_{max}$), and the maximum horizontal wavenumber equal to one fifth the value of the maximum expected wavenumber ($k_{xmax} = 0.2\ k_{max}$), with a taper width equal to half the wavenumber maximum. They suggest more aggressive filtering of the maximum horizontal wavenumber at later stages, but I chose not to implement this as one of my primary goals was to resolve lateral variations in the lithology. Further inversion parameters tested included the number of iterations per block, as well as the number of inversion frequencies. Sirgue and Pratt (2004) proposed a strategy for selecting waveform inversion frequencies for a homogenous layer that is based on the theoretical limit of the wavenumber spectral coverage. For realistic data cases, it is more common to select frequencies that cover the spectrum of the data. I adopt this approach with some overlap in inversion frequencies, which is analogous to signal stacking is reflection seismics. Finally, I determined that five iterations per block of frequencies was the optimal choice when considering the tradeoff between model improvement and inversion runtimes, and I adopted a line search method for the first iteration of each block as in Chapter 3.

The last regularization parameter that was investigated was offset weighting of the data. The gradual increase of offset ranges utilized by successive inversions is a common technique for mitigating cycle-skipping during FWI, as modelling inaccuracies are most pronounced at the furthest offsets (Sirgue and Pratt, 2004; Brenders and Pratt, 2007). The rate at which this offset constraint is relaxed is often empirical, and can significantly increase computation costs if a conservative approach is taken when allowing more offsets into the inversion. Furthermore, restriction of the offsets can affect model resolution as wide-angle crosshole gathers contribute significantly to resolving the lateral structures within the model. I adopt the following strategy for offset weighting: for each frequency block, the first 5 iterations are performed with an offset taper from 155 m to 165 m, meaning that data with greater than 165 m offset were disregarded completely. The same frequencies were then inverted again using 5 iterations, except the offset restriction was relaxed to instead taper from 160 m to 185 m which was the largest observed offset with an acceptable SNR.

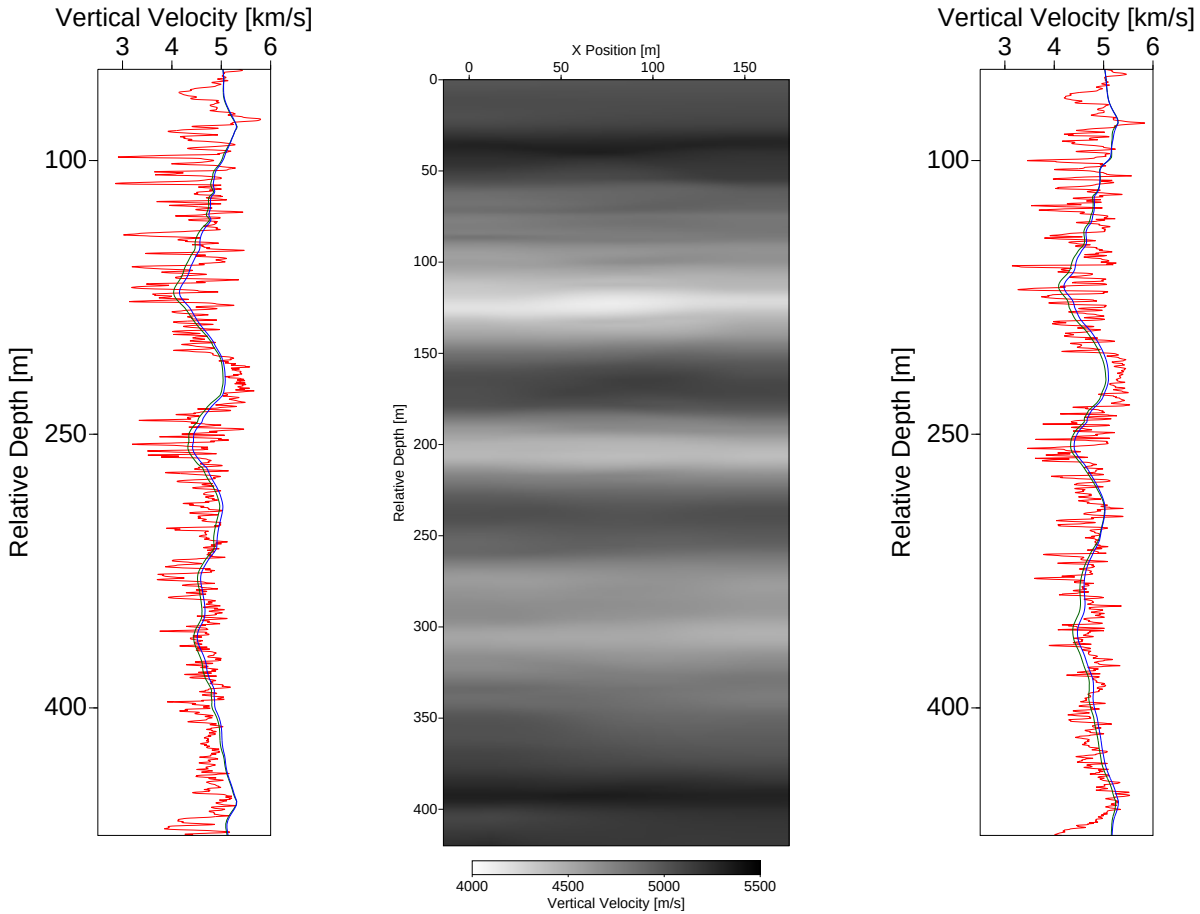
Table 4.4 summarizes the inversion frequencies for each stage of FWI. Recalling that each frequency block was inverted twice for a total of 10 iterations, the entire FWI strategy comprises 60 iterations. Preliminary inversions below 400Hz offered no significant contributions to the model resolution and were therefore disregarded to minimize computation costs. Each frequency block was limited to three frequencies based on the memory limits of the system. A source inversion was performed before and after each Stage in an attempt to recover the true source signatures, which could also improve the results themselves. The models obtained from Stage 1 Inversion were then interpolated onto the finer grid and used as the starting models for Stage 2 Inversion. For all source and velocity inversions, a constant attenuation model (Q = 200) was used.

FWI Stage	Inversion Frequencies [Hz]
1	400, 500, 600 600, 700, 800
2	1000, 1100, 1200 1200, 1300, 1400 1400, 1500, 1600 1600, 1700, 1800

Table 4.4: Inversion frequencies for anisotropic and elliptical FWI of the Western Canada crosshole dataset.

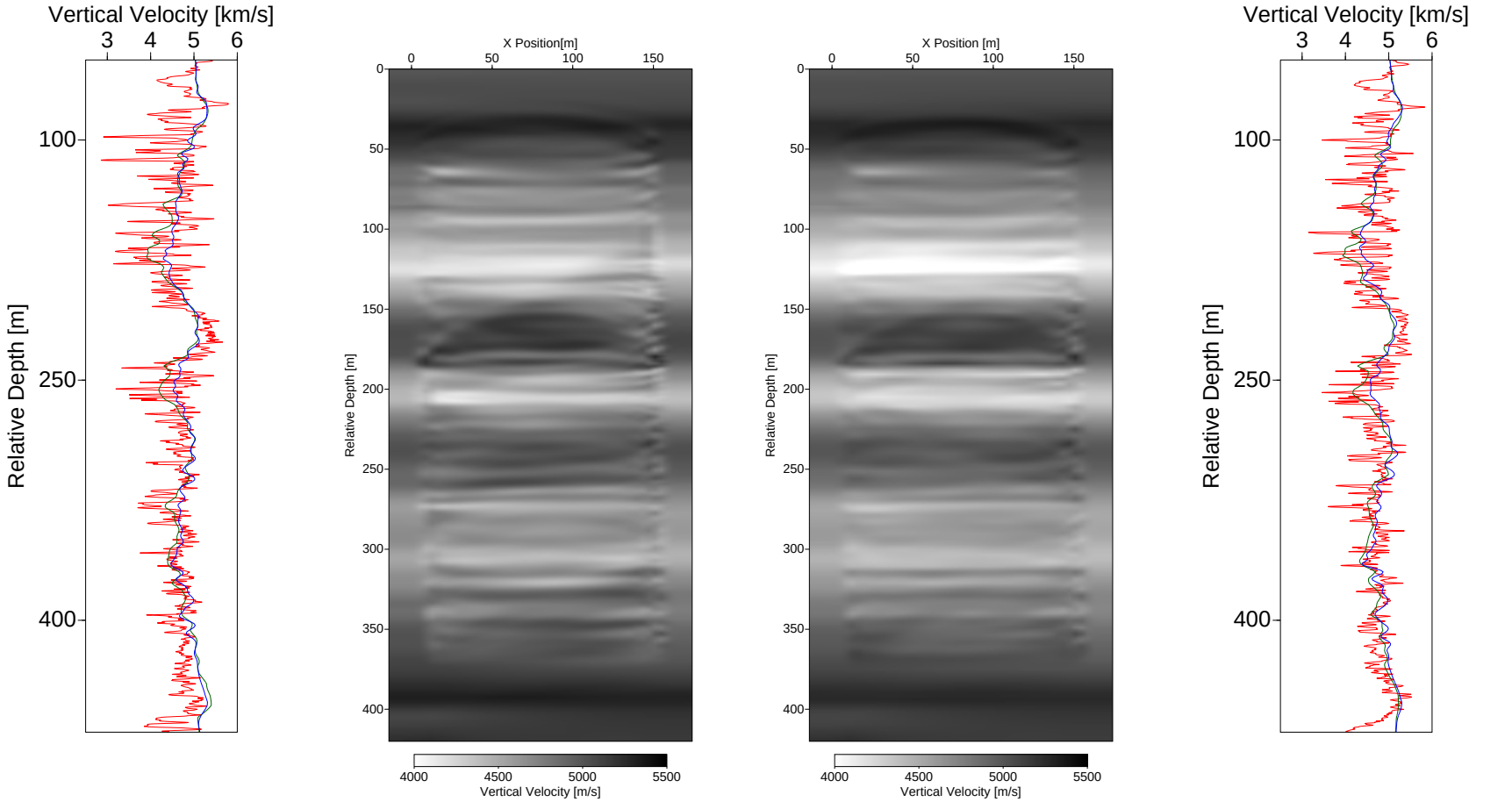
Figure 4.6 shows the starting model obtained from traveltime tomography plotted alongside the sonic logs from the source and receiver boreholes. 1-D vertical profiles have been extracted from the starting model that are proximal to each borehole and plotted with the sonic logs for a direct comparison. The x-locations for these were carefully chosen to be outside the tapered

regions described above. Figures 4.7 and 4.8 show the FWI results for both the elliptical and anelliptical cases for Stages 1 and 2 respectively. For a comparison of computational efficiency, the elliptical result in Figure 4.8 was obtained after 3 hours, whereas the anelliptical result required 60 hours for the same number of FWI iterations running on a workstation equipped with an Intel Xeon CPU E5-2650 with 16 active processors. Figure 4.9 shows the results of the source inversions performed before Stage 1 Inversion, after Stage 1 Inversion and after Stage 2 Inversion for the anelliptical case.



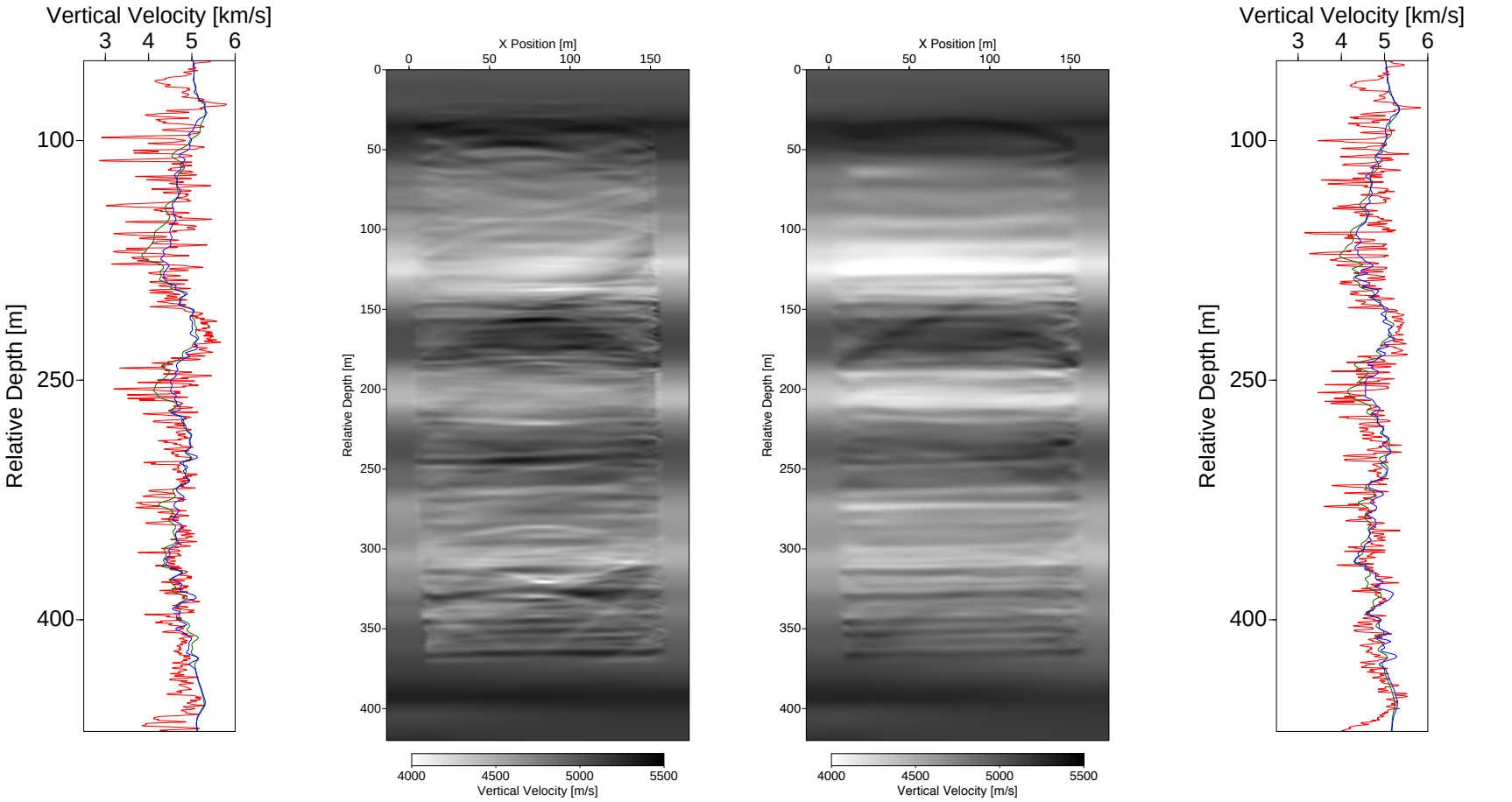
(a) Sonic log from source borehole (b) Starting model for FWI (c) Sonic log from receiver borehole

Figure 4.6: Left and Right: Vp Sonic logs (red) from the source and receiver boreholes with 1-D traces from the elliptical (Green) and anelliptical (Blue) models overlaid for comparison. Middle: Starting model for FWI.



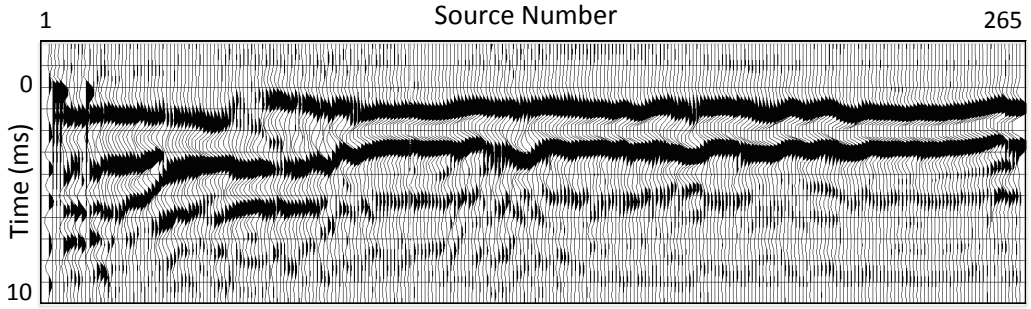
(a) Sonic log from source borehole (b) Anelliptical FWI (c) Elliptical FWI (d) Sonic log from receiver borehole

Figure 4.7: Left and Right: Vp Sonic logs (red) from the source and receiver boreholes with 1-D traces from the elliptical (Green) and anelliptical (Blue) models overlaid for comparison. Middle: FWI results for vertical velocity after Stage 1 Inversion.

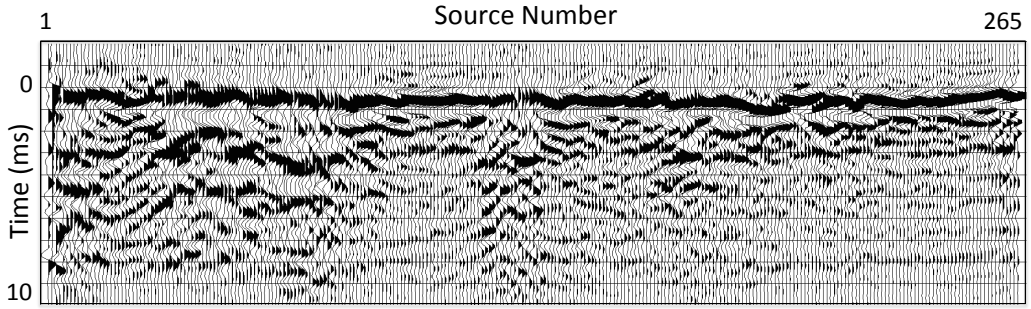


(a) Sonic log from source borehole (b) Anelliptical FWI (c) Elliptical FWI (d) Sonic log from receiver borehole

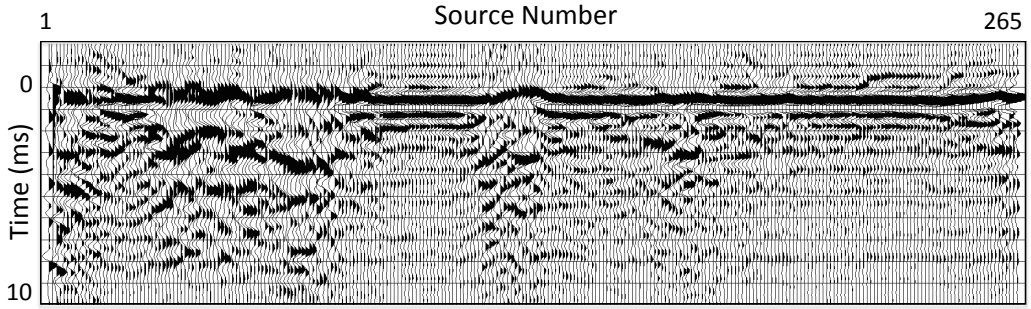
Figure 4.8: Left and Right: Vp Sonic logs (red) from the source and receiver boreholes with 1-D traces from the elliptical (Green) and anelliptical (Blue) models overlaid for comparison. Middle: FWI results for vertical velocity after Stage 2 Inversion.



(a) Initial Source Inversion



(b) Source Inversion After Stage 1 of Anelliptical FWI



(c) Source Inversion After Stage 2 of Anelliptical FWI

Figure 4.9: Results for the source inversions conducted before Stage 1 Inversion (a), after Stage 1 Inversion (b) and after Stage 2 Inversion (c) for the anelliptical case. Note the increase in the frequency content of the data between panels (a) and (b). Overall, the FWI results appear to have improved the coherency in the recovered source signals.

After Stage 1 inversion, we see that both models show a more complex layering pattern, with some lower velocity layers being introduced in the suspected reservoir (Formation A) which could indicate the presence of interbedded shale layers within the predominantly sandstone formation. The inversion also appears to have further refined the top and bottom boundaries of the reservoir layer (beginning at the depth of approximately 150 m and extending to a depth of 185 m), as well as the lower velocity layers directly above and below it within the sedimentary sequence. Furthermore, the Stage 1 result seems to suggest a second high velocity layer exists beginning at a depth of approximately 320 m, however these deeper features are still poorly resolved at this frequency.

It is apparent from the Stage 2 results that the introduction of the higher frequency data further resolved some of the finer (≈3-10m) layers in the models. Figure 4.10 is a zoomed-in version of Figure 4.8, shown to illustrate the finer differences between the two results in the region of the suspected reservoir. Comparing the results between the two subcases, it is apparent that the anelliptical result was more successful in resolving these finer structures. This is especially obvious near the bottom of the reservoir region (≈180m depth), and the results are also consistent with the structure observed in the sonic logs at this depth. Comparing the long wavelength features of the 1-D profiles of both models with the sonic logs in Figure 4.8, it is apparent that both models match this trend reasonably well. It is difficult to compare the values of velocity, recalling that the expected tradeoffs between the velocity and ϵ parameters makes it impossible to recover the true velocities exactly.

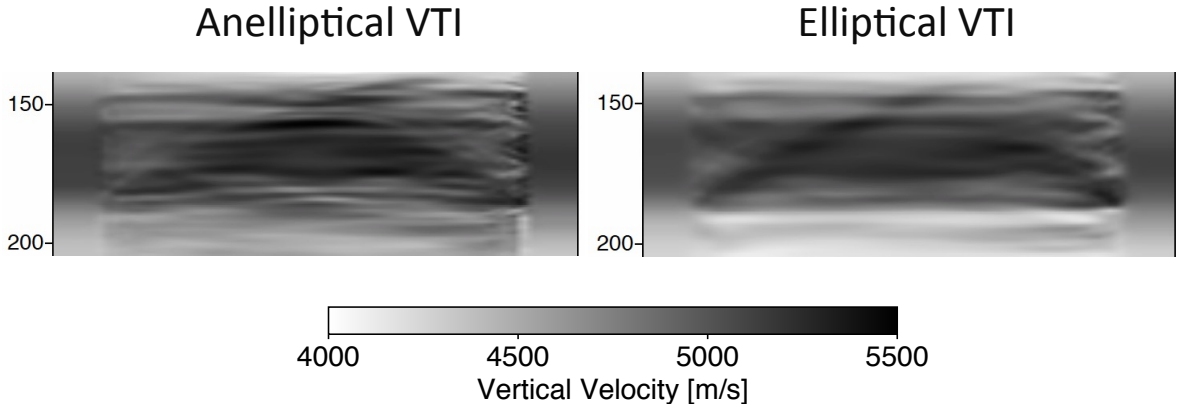
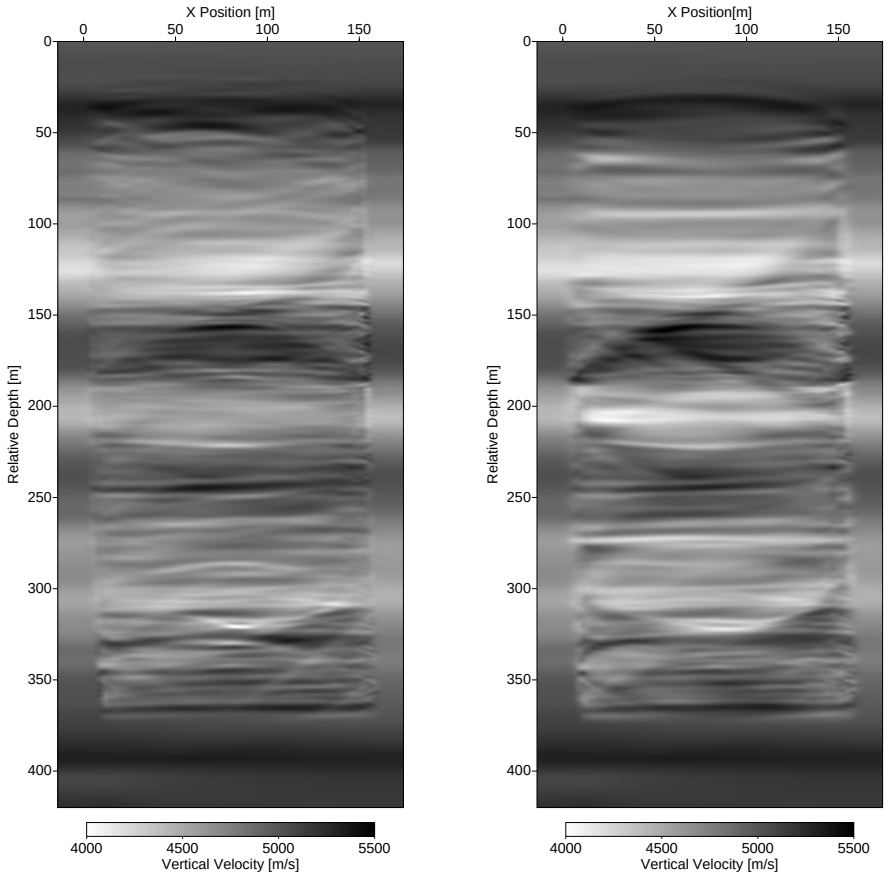


Figure 4.10: Stage 2 FWI results for vertical velocity zoomed in to focus on the features within the reservoir for the anelliptical (left) and elliptical (right) cases.

Regarding the source inversions, we see some incoherencies in the first ≈100 sources of the initial inversion (panel (a) of Figure 4.9). These can be attributed to the lack of data coverage in this region, recalling the decreased SNR observed for the shallower parts of the region. One might also notice an increase in the sharpness of the source signatures recovered before

and after Stages 1 and 2 of FWI (panels (b) and (c) of Figure 4.9). This is because these inversions were conducted using the full frequency spectrum (up to 2000 Hz). In the initial source inversion, we can see that the shallowest sources are slightly delayed in time, which suggests a mismatch in velocity (the inversion algorithm is compensating for these velocity mismatches by shifting the signatures in time). For the source inversion conducted after Stage 1 of anelliptical FWI, we see that the source signatures have shifted closer to zero time, and some of the incoherencies in the first 100 sources have been resolved. For the source inversion conducted after Stage 2 of anelliptical FWI, we see that the overall coherency of the signals has been further improved, most notably from the 130th source onwards. No *a priori* information regarding the source characteristics was provided, so I assume that the desired result is a suite of sources which are coherent and are as close to zero time as possible. Therefore, the results shown in Figure 4.9 support the hypothesis that FWI has improved the velocity model for this region, based on the improvement in the source signals recovered after each successive inversion.

It is also important to acknowledge the undesirable features in the FWI results as well, the most noteworthy being the presence of steeply-dipping artifacts that appear to originate from the boreholes. These artifacts are less drastic in the elliptical results which suggests that they are directly related to the accuracy of the anelliptical anisotropy models. Afanasiev et al. (2014) showed crosshole FWI results that also contain these steeply dipping artifacts and they suggest that these are related to mismatches between the observed and predicted data at the largest offsets. This is consistent with what I observe as well. Figures 4.11 and 4.12 shows the results for Stage 2 inversion compared against a similar result obtained without using offset weighting for the anelliptical and elliptical cases respectively. It is evident that the steeply dipping artifacts are more prominent in the results without offset weighting, which suggests that gradually increasing the range of data offsets reduces the magnitude of these artifacts. Still, the reduced presence of these artifacts in the elliptical results draws the accuracy of the anisotropy models obtained from traveltime tomography into question. As I have highlighted through this thesis, tradeoffs in parameter resolution is an unfortunate consequence of multiparameter inversions. I therefore conclude that these artifacts are a result of these tradeoffs and can only be completely removed by obtaining more accurate starting models for ϵ and δ, or by updating these models during FWI.



(a) Anelliptical FWI with offset weighting (b) Anelliptical FWI without offset weighting

Figure 4.11: Left: Anelliptical FWI results for vertical velocity after Stage 2 Inversion using the offset weighting parameters described in the text. Right: Anelliptical FWI results for vertical velocity after Stage 2 Inversion without using offset weighting of the data.

(a) Elliptical FWI with offset weighting (b) Elliptical FWI without offset weighting

Figure 4.12: Left: Elliptical FWI results for vertical velocity after Stage 2 Inversion using the offset weighting parameters described in the text. Right: Elliptical FWI results for vertical velocity after Stage 2 Inversion without using offset weighting of the data.

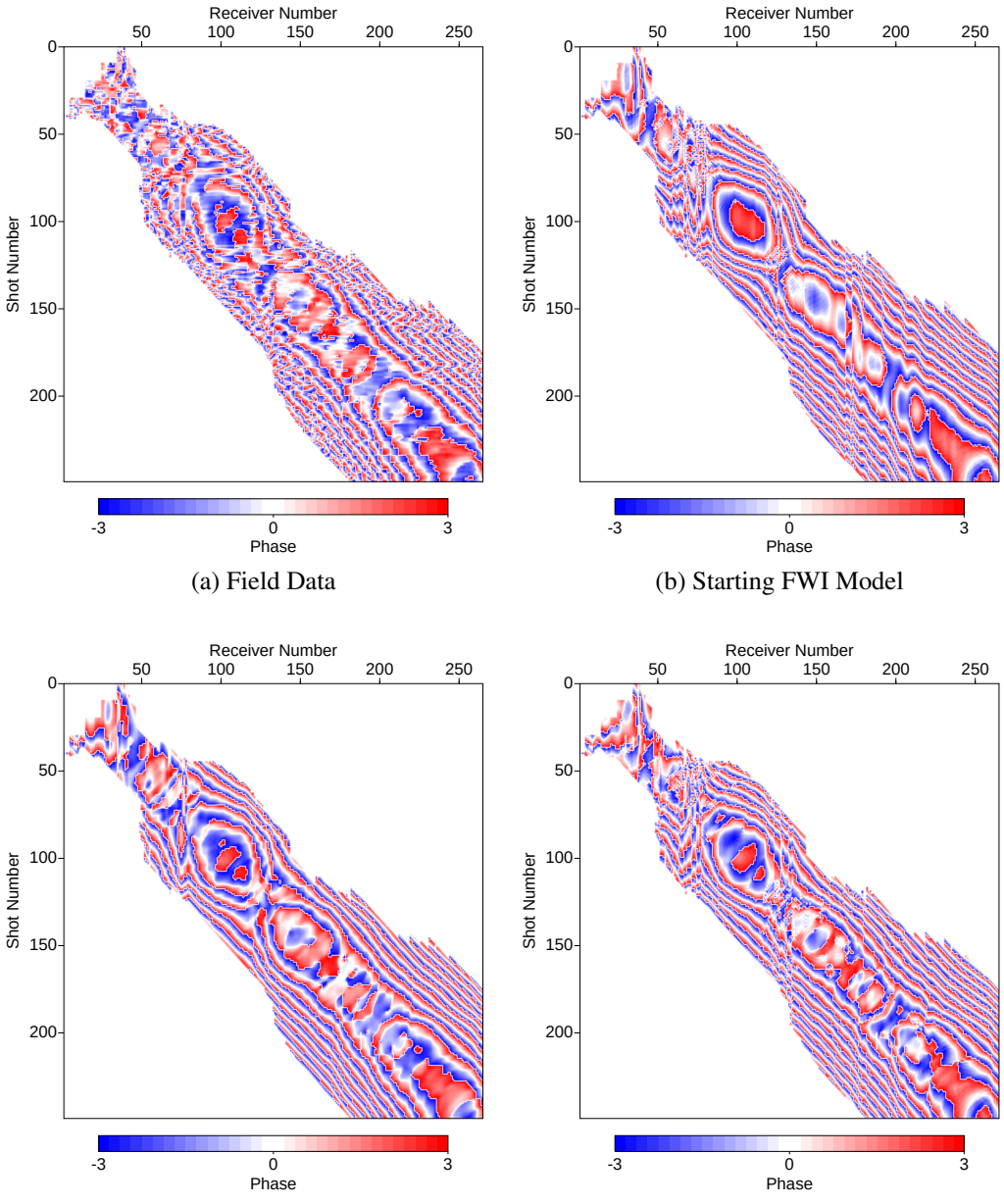
4.4 Data Fit

In Chapter 3, I was able to verify my traveltime tomography and FWI results by direct comparison to the true models. For field data cases, the true models are often not available in any capacity. Therefore, in order to asses the data fit, I compare the forward modelling results in the final FWI models (Figure 4.8) in both the frequency- and time-domains to the true data.

Figure 4.13 shows the phase component of the 800 Hz wavefield extracted from the field data (a), the best-fitting velocity model obtained from traveltime tomography (b), and the final FWI models obtained from Stage 2 inversion for the elliptical (c) and anelliptical (d) cases. Here I chose to analyze the phase mismatch in the data as this was the only component of the waveforms that was updated during FWI. Comparing the field data to the starting model, it is apparent that traveltime tomography was successful in recovering the low wavenumber features of the model, however some of the finer structure is not present in these results, which is as expected. Comparing the elliptical and anelliptical wavefields to the field data, the differences are subtle, although the elliptical result appears to better match the field data at the near offsets (the elements nearest to the diagonal) in the shallower parts of the model. I suggest that the inaccuracies in the anelliptical result could be related to the mismatches between the long wavelength features of the velocity and anisotropy models (similar to the steeply-dipping artifacts, as these are first introduced in the lower frequencies of FWI), however these differences are minor when you take into consideration the SNR of the data. There is also some high wavenumber 'ringing' near the centre of the anelliptical result which I believe is related to the spurious S-wave energy, as these artifacts do not appear in the elliptical results. Furthermore, these artifacts are most apparent where the anellipticity is greatest, which is consistent with my observations in Section 2.3. This would suggest that the Laplace transform was not successful in completely damping the S-waves, although it appears to have damped them enough to allow the FWI algorithm to succeed.

Figures 4.14 and 4.15 show the fields gathers for sources located at depths of 125 m and 325 m respectively, as well as my interpretation of the first arrivals in each case. Figures 4.16 and 4.17 shows the forward modelled shot gathers for the elliptical and anelliptical FWI models with the first arrivals overlaid. It is evident that, while both models predict the near offsets reasonably well, the anelliptical model better predicts the far offsets in both cases. Although the improvements are minor, we can see that the elliptical model is at risk of cycle-skipping at the deepest receivers ($\approx$130 m and below in panel (a) of Figure 4.16, and $\approx$340 m and below in panel (a) of Figure 4.17) in both cases. These observations are consistent with the frequency-domain results for the elliptical case, as in panel (c) of Figure 4.13, the phase mismatch is most pronounced at the furthest offsets in the deeper part of the model.

Combining the observations in the final FWI models, as well as the differences between the field data and the predicted data in both the time- and frequency-domains, I conclude that the application of anisotropic FWI to these data was successful in improving the FWI results when compared to those obtained using the elliptical isotropy assumption. These results validate the hypothesis that properly accounting for anisotropy can lead to an improvement in the models produced by FWI. The anelliptical result shows more complex structures within the reservoir region, which could have an impact on reservoir simulations, as well as on the overall geological interpretation of the area. Nevertheless, these results do contain some undesirable features which suggests that further improvement is possible. I summarize my thoughts on the overall success of this work, as well as my suggestions for future research in the fifth and final chapter of this thesis.



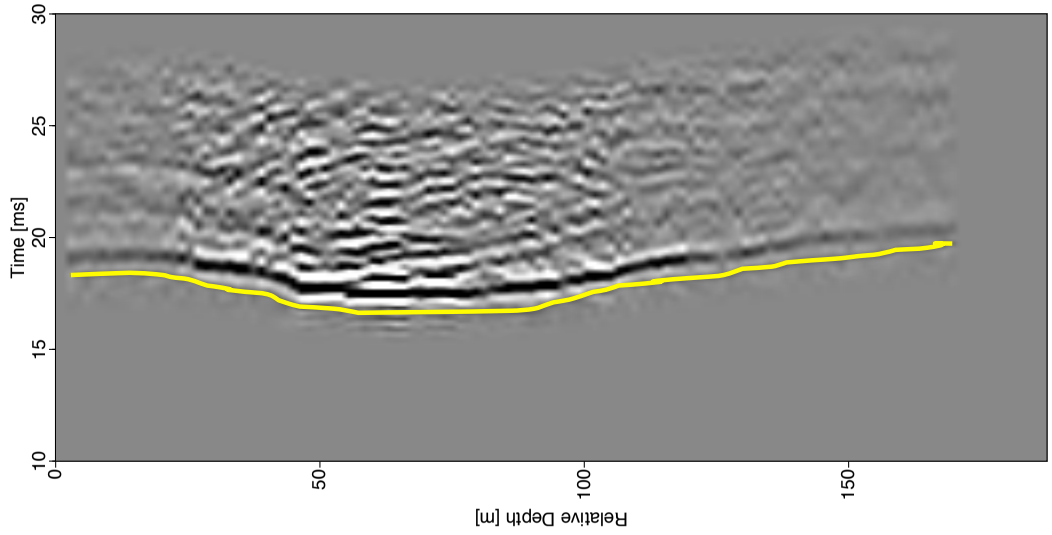
(a) Field Data

(b) Starting FWI Model

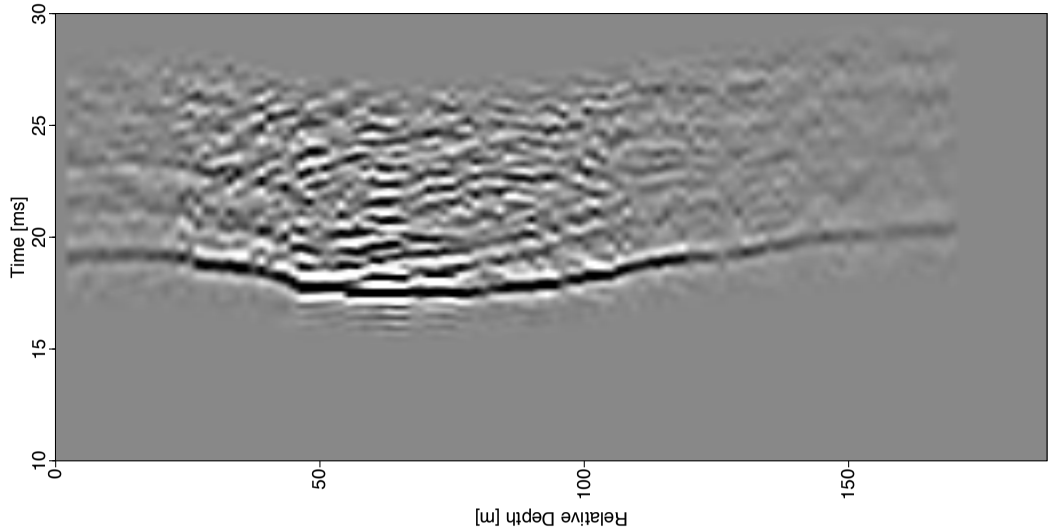
(c) Stage 2 Elliptical FWI Model

(d) Stage 2 Anelliptical FWI Model

Figure 4.13: Phase component of the 800 Hz complex-valued frequency-domain wavefield for the field data compared to the 800 Hz wavefield in the starting FWI model, as well as in the models produced by Stage 2 Inversion for the elliptical and anelliptical cases. Note here that the axes are source and receiver number.

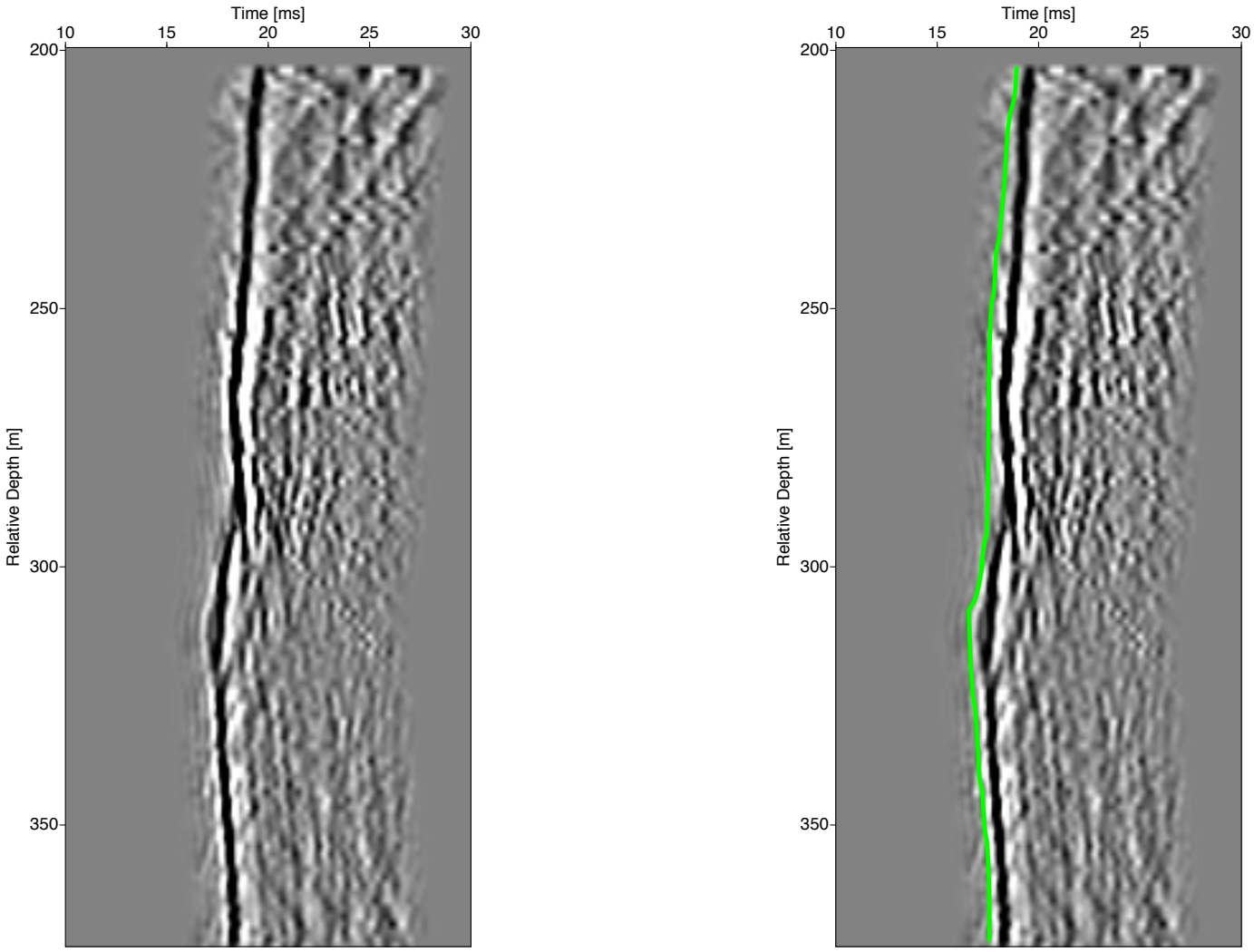


(b) Raw Shot Gather with First Arrivals



(a) Raw Shot Gather

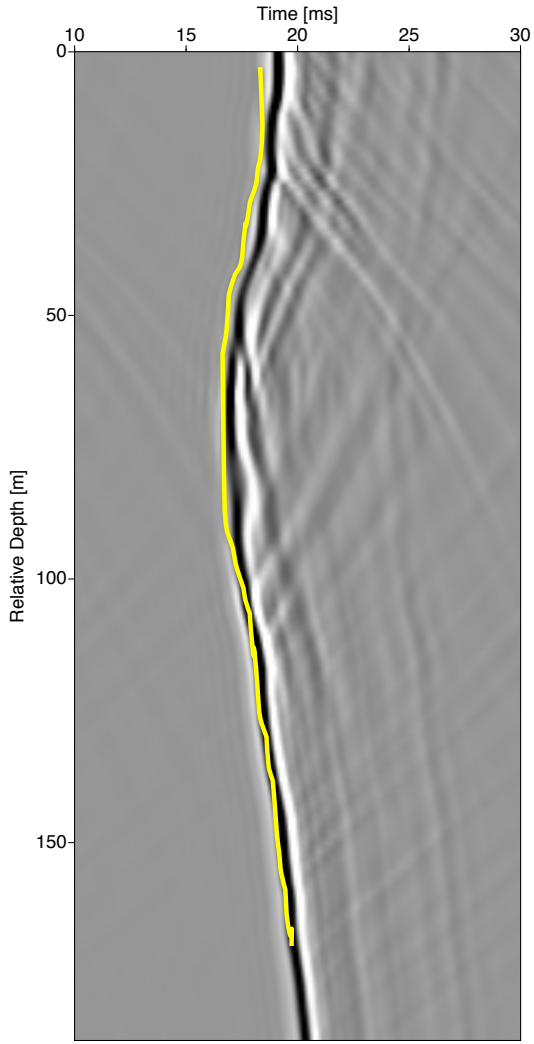
Figure 4.14: A selected field gather for a source located at a depth of 125 m with first arrivals overlaid in yellow. The data are modelled in reduced time with a reduction velocity of 4000 m/s.



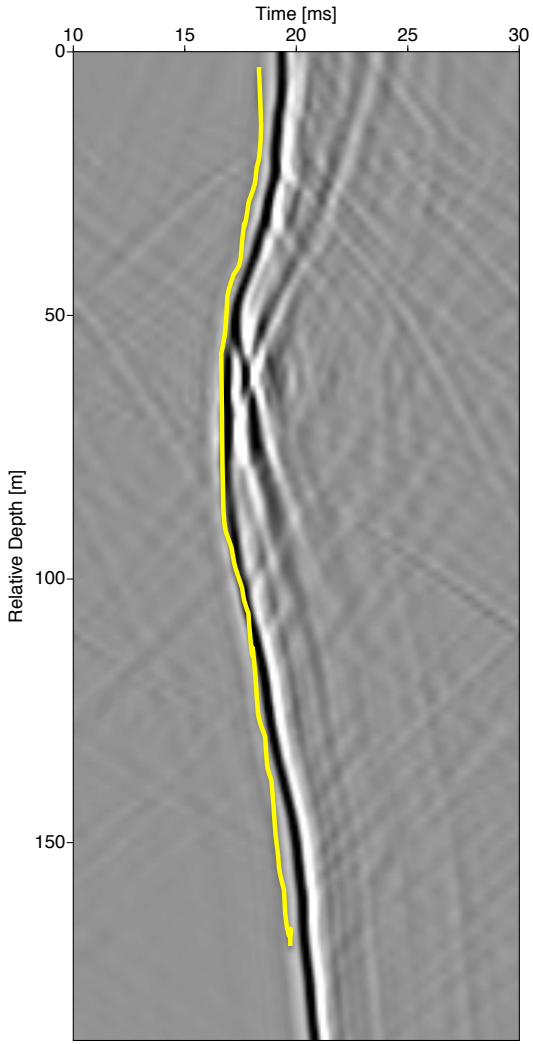
(a) Raw Shot Gather

(b) Raw Shot Gather with First Arrivals

Figure 4.15: A selected field gather for a source located at a depth of 325 m with first arrivals overlaid in green. The data are modelled in reduced time with a reduction velocity of 4000 m/s.

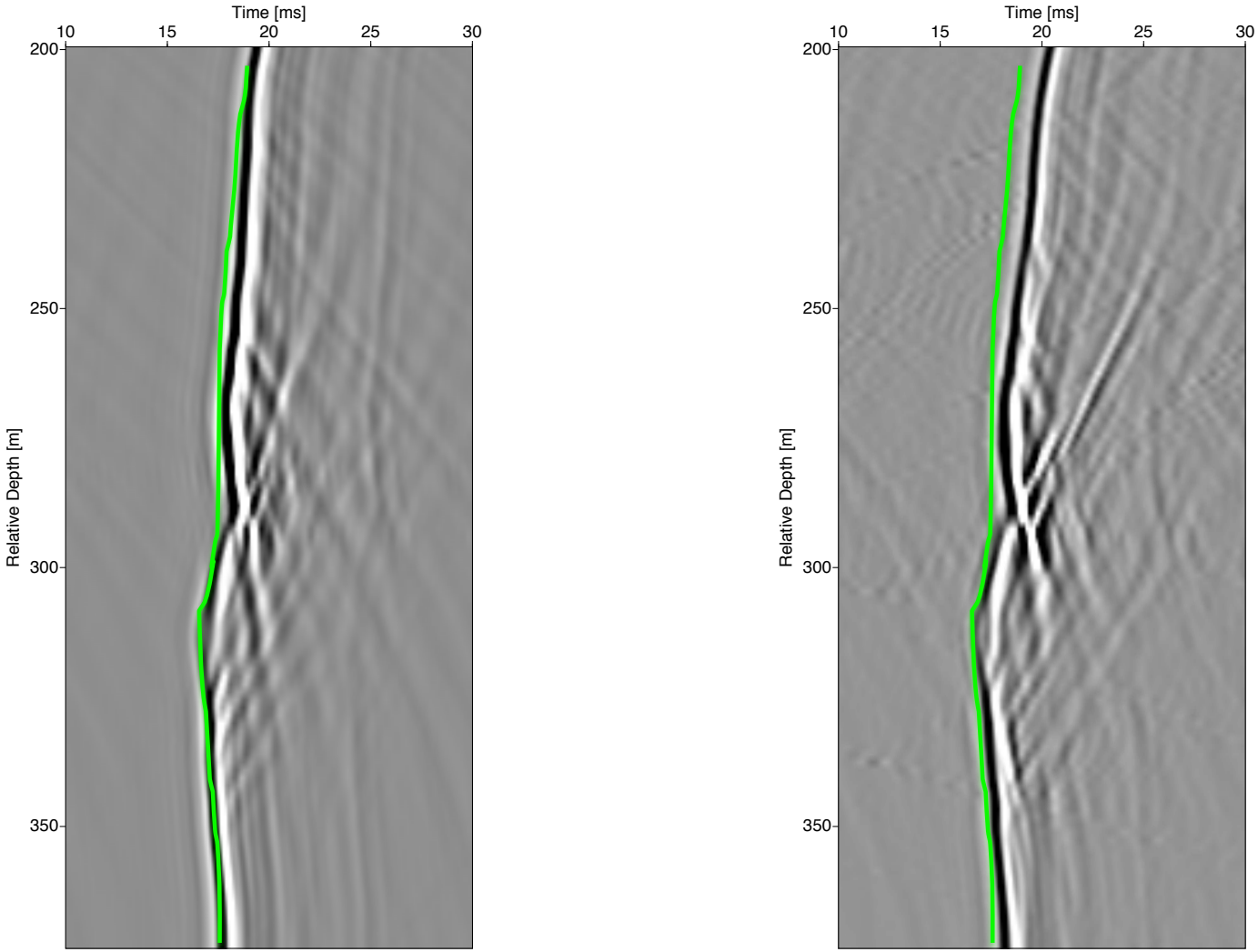


(a) Elliptical FWI Shot Gather



(b) Anelliptical FWI Shot Gather

Figure 4.16: Forward modelled shot gather for a source located at a depth of 125 m for the Elliptical and Anelliptical FWI results with the field arrivals overlaid in yellow. The data are modelled in reduced time with a reduction velocity of 4000 m/s. The source used for this simulation is the best-fitting source obtained from Stage 2 inversion. The damping factor τ is set to $0.4*t_{max}$. When directly comparing the results (in particular, the gathers from ≈130 m to 155 m), it is evident that the anelliptical result better predicts the true arrivals.



(a) Elliptical FWI Shot Gather

(b) Anelliptical FWI Shot Gather

Figure 4.17: Forward modelled shot gather for a source located at a depth of 325 m for the Elliptical and Anelliptical FWI results with the field arrivals overlaid in green. The data are modelled in reduced time with a reduction velocity of 4000 m/s. The source used for this simulation is the best-fitting source obtained from Stage 2 inversion. The damping factor τ is set to $0.4*t_{max}$. When directly comparing the results (in particular, the gathers from ≈200 m to 220 m, as well as ≈340 m to 380 m,), it is evident that the anelliptical result better predicts the true arrivals.

Chapter 5

Conclusions

In this thesis, I have applied the joint approach of anisotropic traveltime tomography and Full Waveform Inversion to recover high resolution models of velocity and Thomsen's anisotropy parameters for both real and synthetic data. In order to do this, I implemented a wave modelling technique for FWI that better accounts for anisotropy for applications where significant anisotropy is known to exist.

I began in Chapter 1 by introducing crosshole seismic surveys as well as their applications within the geophysical industry. Next, I discussed anisotropy in the context of seismic waves, the challenges that anisotropy introduces and I described several anisotropy models that are commonly encountered in realistic media. Following this, I gave a brief literature review of traveltime tomography and FWI, including their inception, development and improvement over time, strengths and pitfalls, as well as the current industry standard regarding implementations of these techniques.

In Chapter 2, I derived the elastic wave equation from stress-strain relations and showed how the internal stress tensor can be modified for different anisotropy systems, as well as how the acoustic wave equation for isotropic media can be approximated from the elastic formulation. Next, I introduced the fundamental concepts of forward modelling and inversion which I then applied to the underlying methodologies of traveltime tomography and FWI. Derivations of the governing system of equations for each technique were complimented by discussions regarding the justification of assumptions made, modelling limitations and strategies that can be applied to aid in obtaining successful results. Next, I provided a full history of the anisotropic acoustic wave equation that I have implemented into the FWI codebase maintained by Dr. Pratt and his research colleagues. Beginning with the original derivation from phase velocity relations for VTI media, I show how the equation evolved over time through the contributions of several authors, as well as how it compared to its isotropic counterpart. Finally, I presented the FDFD modelling technique that I have adopted for accurate forward modelling of the wavefield

in VTI media. To validate that I implemented this technique correctly, I replicated some of the published results for this implementation.

In Chapter 3, I tested the joint approach with a synthetic VTI model. Several preliminary traveltime tomography inversions were conducted to determine the optimal combination of regularization parameters that produced traveltime results that were suitable starting models for FWI. Several inversions were conducted using different approaches for handling anisotropy based on modern anisotropic FWI studies. Overall, these results were encouraging as the targeted velocity anomaly was well resolved in all cases, which suggests that applying anisotropic traveltime tomography to obtain starting models for monoparameter FWI is a valid approach for transversely isotropic media. These results also highlighted the expected tradeoffs in resolution between the two parameters classes (velocity and anisotropy), which is a commonly encountered problem in multiparameter inversion studies.

In Chapter 4, I applied anisotropic waveform tomography to field data from a crosshole survey in Western Canada. Best-fitting models of velocity, ϵ and δ obtained from traveltime tomography were used as the starting models for two separate applications of FWI: the first was the anisotropic devlopment that I have implemented and the second was an isotropic modelling approach with a coordinate transform applied to simulate elliptical isotropy. An efficient two-stage FWI strategy was developed that encompassed several regularization techniques including spatial weighting of the gradient, wavenumber filtering, as well as gradually increasing the inversion frequencies and data offset range. Comparison of the final models obtained for vertical velocity in each case showed that the anisotropic modelling technique was superior in imaging the finer structure. Furthermore, comparison of the predicted data in each model showed that the anisotropic result better matched the true data. These results suggest that the more generalized approach for handling FWI led to an improvement of the results for this case study. There were, however, some artifacts present in the anisotropic results that suggest the anisotropy models were not fully representative of the furthest offsets. These results have implications for future works related to transversely isotropic media, as well as multiparameter FWI studies.

5.1 Final Thoughts and Future Directions

The primary conclusion of this thesis is that the extra computational costs and modelling complexities introduced by properly accounting for anisotropy within the FWI process is justifiable on the basis of the improvements that it provides to the velocity images, as well as the overall accuracy of the model. This is evident in the results of Chapter 4 when comparing the data fit for the elliptical and anelliptical results. Furthermore, the anisotropic modelling technique is

fully compatible with the models produced from traveltime tomography, whereas the elliptical approach requires the ϵ and δ models to be simplified significantly. From my results, I believe that local optimization methods are capable of producing acceptable results from these starting models in the anisotropic case.

There are, however, several outstanding issues related to this work that I would like to acknowledge. Without effectively removing the S-wave artifacts that are generated at the source when $\epsilon > \delta$, it is impossible to tell how these artifacts manifest themselves within the FWI results, or whether they have any noticeable effect at all. As I have previously alluded to, there are several techniques presented in the literature that attempt to remove these artifacts (Fletcher et al., 2009; Operto et al., 2009; Métivier et al., 2014). I believe that the application of the Laplace transform aided in the attenuation of these late arriving artifacts, although some S-wave energy is still present in the frequency-domain results. Future FWI studies that adopt this acoustic wave equation for TI media should investigate one or more of these techniques to see whether or not the FWI images are improved by the effective removal of the spurious S-waves. Another criticism is the simplicity in the synthetic model presented in Chapter 3. A synthetic model that is more consistent with the layered lithology encountered in Chapter 4 would have allowed a better understanding of how the tradeoffs in resolution manifest themselves within this type of environment. This would have allowed for a more direct comparison to be made between the traveltime results in each study as it is evident that the recovered anisotropy models in Chapter 3 are contaminated by the velocity anomaly and vice-versa.

One might also question the decision to restrict the FWI analyses to monoparameter inversions only. Plessix and Cao (2011), Gholami et al. (2013a) and Alkhalifah and Plessix (2014) all presented strategies for multiparameter FWI for anisotropic media that are based on the data dependancies and radiations patterns. The primary reason for this decision is that, presently, the inversion methods of Pratt et al. (1998) support inversions for velocity and/or attenuation only. Therefore in order to invert for the Thomsen's parameters as well, the inversion methods of Pratt et al. (1998) will need to be updated significantly. Nevertheless, the objectives of the field survey presented in Chapter 4 were all concerned with the interpretation of the velocity model only. As anisotropic studies are still relatively new in the context of seismic waveform modelling, there is less understanding of how these parameters are related to different lithologies/minerologies than its velocity (and to some extent, attenuation) counterparts.

I would also like to address the case in which the symmetry axis might be tilted from the vertical ($\theta_0 \neq 0$). All of the data examples presented in this thesis were restricted to VTI symmetry. While this model is appropriate for some regions (such as the one presented in Chapter 4), more complex environments such as overthrusted regions or tilted orebody intrusions are better represented by TTI symmetry. The sedimentary formations below the unconformity in

Chapter 4 do appear to be dipping slightly, therefore it would be interesting to re-perform these analyses with a TTI modelling approach for direct comparison. Operto et al. (2009) attributed the instabilities observed in the tilted case to the slow-moving S-wave artifacts and show that they can be effectively removed by matching the grid interval to the P-wavelength. This is a potential solution for multifrequency datasets as the grid interval is often chosen to accommodate the shortest wavelengths (highest frequencies).

This has been a topic of discussion within our research group for some time, and an idea was put forth called the 'multi-grid' technique. In this approach, the wavefield at each frequency is interpolated onto a separate grid where the interval is matched specifically to the wavelength at that frequency. Some preliminary work has been done on this method, however is remains incomplete and untested. Therefore, I propose that future research could be centred around the further development of this technique for its application to TTI FWI. For the implementation described in this thesis to be extended to TTI media, there are two approaches that could be taken. The first would be to reformulate the boundary conditions entirely as these instabilities are directly related to geometric limitations in the PML layer when the axis of symmetry is tilted (Bécache et al., 2003). One such alternative was proposed by Métivier et al. (2014) who replaced the PML layer with a new formulation called a SMART layer which embeds the source in a small region in which only the S-wave energy is damped. The second alternative would be to develop a novel technique that successfully attenuates the S-wave for PML absorbing boundary conditions, such as the one described above.

I would like to leave the readers with one final thought. Some might be surprised that the works described in this thesis are based on an unrealizable physical phenomena (acoustic anisotropy). This can be related to the longstanding philosophical argument between the 'Engineering' approach and the 'Scientific' approach. As my academic background involves both of these fields, this has certainly weighed on my thoughts over the course of this thesis. I have come to the conclusion that, with today's modelling and computation capabilities, completely honoring the 'science' of the problem is often not feasible, as even in the fully elastic case the science must be simplified significantly. As such, I employed several 'engineering' assumptions, including the restriction to acoustic physics, the restriction to 2-D models, and the imposition of regularization during the inversion processess. I believe that the results presented in this thesis lie somewhere between the two end members of this debate. While one could argue that the acoustic framework (consideration of P-waves only) is suitable for answering all of the crosshole survey objectives presented in Chapter 4 (as the client was not interested in the S-wave models, and the P-wave models represent the data reasonably well), the presence of artifacts in the final FWI models, as well as in the predicted frequency-domain data suggest there is still the potential to 'engineer' a better solution.

To quote the world renowned physicist Albert Einstein:

"It can scarcely be denied that the supreme goal of all theory is to make the irreducible basic elements as simple and as few as possible without having to surrender the adequate representation of a single datum of experience. " (Einstein, 1934).

From this influential philosophy, the following famous paraphrase has been extracted and applied across several scientific disciplines:

"Everything should be made as simple as possible, but no simpler."

Perhaps the answer to these artifacts lies within one of the suggestions outlined above, or it could be the case that the 'true' solution lies outside the realm of acoustic anisotropy. Nevertheless, if there is one thing I would encourage the reader to take away from the works presented in this thesis, it is that FWI is indeed a state-of-the-art technique for resolving fine velocity structure within complex, geological environments. I would like to end on that note, but I look forward to what the future holds for both FWI and seismic modelling studies, in the context of meeting the challenges of seismic anisotropy.

Bibliography

Afanasiev, M., Pratt, R.G., Kamei, R. and McDowell, G. [2014] Waveform-based simulated annealing of crosshole transmission data: a semi-global method for estimating seismic anisotropy. *Geophysical Journal International*, **199**(3), 1586–1607.

Airhart, T.P. [1989] Vibratory seismic source.

Aki, K., Christoffersson, A. and Husebye, E.S. [1977] Determination of the threedimensional seismic structure of the lithosphere. *Journal of Geophysical Research*, **82**(2), 277–296.

Aki, K. and Richards, P. [2002] *Quantitative Seismology*. University Science Books.

Alkhalifah, T. [1998] Acoustic approximations for processing in transversely isotropic media. *Geophysics*, **63**(2), 623–631.

Alkhalifah, T. [2000] An acoustic wave equation for anisotropic media. *Geophysics*, **65**(4), 1239–1250.

Alkhalifah, T. and Plessix, R.é. [2014] A recipe for practical full-waveform inversion in anisotropic media : An analytical parameter resolution study. *Geophysics*, **79**(3), R91–R101.

Allen, R.V. [1978] Automatic earthquake recognition and timing from single traces. *Bulletin of the Seismological Society of America*, **68**(5), 1521–1532.

Bécache, E., Fauqueux, S. and Joly, P. [2003] Stability of perfectly matched layers , group velocities and anisotropic waves. *Journal of Computational Physics*, **188**(1), 399–433.

Bois, P., La Porte, M., Lavergne, M. and Thomas, G. [1972] Well-to-Well Seismic Measurements. *Geophysics*, **37**(3), 471–480.

Brenders, A. [2011] *Strategies for Waveform Tomography of Long-Offset, 2-D Exploration Seismic Data*. PhD, University of Western Ontario.

Brenders, A.J. and Pratt, R.G. [2007] Full waveform tomography for lithospheric imaging: Results from a blind test in a realistic crustal model. *Geophysical Journal International*, **168**(1), 133–151.

Brossier, R. [2011] Two-dimensional frequency-domain visco-elastic full waveform inversion: Parallel algorithms, optimization and performance. *Computers & Geosciences*, **37**(4), 444–455.

Brossier, R., Operto, S. and Virieux, J. [2009] Seismic imaging of complex onshore structures by 2D elastic frequency-domain full-waveform inversion. *Geophysics*, **74**(6), WCC105–WCC118.

Bunks, C., Saleck, F.M., Zaleski, S. and Chavent, G. [1995] Multiscale seismic waveform inversion. *Geophysics*, **60**(5), 1457–1473.

Cervený, V. [1972] Seismic rays and ray intensities in inhomogeneous anisotropic media. *Geophysical Journal International*, **29**(1), 1–13.

Cervený, V. [1982] Direct and inverse kinematic problems for inhomogeneous anisotropic media-linearization approach. *Contrib. Geophys. Inst. Slov. Acad. Sci*, **13**, 127–133.

Cervený, V. [2005] *Seismic ray theory*. Cambridge University Press.

Cervený, V. and Jech, J. [1982] Linearized solutions of kinematic problems of seismic body waves in inhomogeneous slightly anisotropic media. *Journal of Geophysics*, **51**, 96–104.

Červený, V., Molotkov, I.A., Molotkov, I.A. and Pšenčík, I. [1977] *Ray method in seismology*. Univerzita Karlova.

Chapman, C.H. and Pratt, R.G. [1992] Traveltime tomography in anisotropic media-I. Theory. *Geophysical Journal International*, **109**(1), 1–19.

Claerbout, J., Green, C. and Green, I. [1985] *Imaging the earth's interior*. Blackwell scientific publications, Oxford.

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

Crase, E., Pica, A., Noble, M., McDonald, J. and Tarantola, A. [1990] Robust elastic nonlinear waveform inversion: Application to real data. *Geophysics*, **55**(5), 527–538.

Davison, M. [1991] *Some Mathematical and Computational Aspects of Cross-Borehole Seismic Tomography*. B.a.sc., University of Toronto.

Dellinger, J. [1991] *Anisotropic seismic wave propagation*. PhD, Stanford University.

Drossaert, F.H. and Giannopoulos, A. [2007] A nonsplit complex frequency-shifted PML based on recursive integration for FDTD modeling of elastic waves. *Geophysics*, **72**(2), T9.

Einstein, A. [1934] On the method of theoretical physics. *Philosophy of science*, **1**(2), 163–169.

Fletcher, R.P., Du, X. and Fowler, P.J. [2009] Reverse time migration in tilted transversely isotropic (TTI) media. *Geophysics*, **74**(6), WCA179–WCA187.

Gardner, G., Gardner, L. and Gregory, A. [1974] Formation velocity and density- the diagnostic basics for stratigraphic traps. *Geophysics*, **39**(6), 770–780.

Gardner, L. [1949] Seismograph determination of salt dome boundary using well detection deep on dome flank. *Geophysics*, **14**(1), 29–38.

Gauthier, O., Virieux, J. and Tarantola, A. [1986] Two-dimensional non-linear inversion of seismic waveforms : Numerical results. *Geophysics*, **51**(7), 1387–1403.

Gholami, Y., Brossier, R., Operto, S., Prieux, V., Ribodetti, A. and Virieux, J. [2013a] Which parameterization is suitable for acoustic vertical transverse isotropic full waveform inversion ? Part 2 : Synthetic and real data case studies from Valhall. *Geophysics*, **78**(2), R107–R124.

Gholami, Y., Brossier, R., Operto, S., Ribodetti, A. and Virieux, J. [2013b] Which parameterization is suitable for acoustic vertical transverse isotropic full waveform inversion ? Part 1 : Sensitivity and trade-off analysis. *Geophysics*, **78**(2), R81–R105.

Grechka, V., Pech, A., Tsvankin, I. and Han, B. [2001] Velocity analysis for tilted transversely isotropic media: A physical modeling example. *Geophysics*, **66**(3), 904–910.

Grechka, V., Zhang, L. and Rector III, J.W. [2004] Shear waves in acoustic anisotropic media. *Geophysics*, **69**(2), 576–582.

Haber, E., Ascher, U., Aruliah, D. and Oldenburg, D. [2000] Fast simulation of 3D electromagnetic problems using potentials. *Journal of Computational Physics*, **163**(1), 150–171.

Han, D., Nur, A. and Morgan, D. [1986] Effects of porosity and clay content on wave velocities in sandstones. *Geophysics*, **51**(11), 2093–2107.

Harris, J.M., Nolen-Hoeksema, R.C., Langan, R.T., Schaack, M.V., Lazaratos, S.K. and Rector, J.W. [1995] High-resolution crosswell imaging of a west Texas carbonate reservoir : Part l-Project summary and interpretation. *Geophysics*, **60**(3), 667–681.

Hicks, G. [2002] Arbitrary source and receiver positioning in finitedifference schemes using Kaiser windowed sinc functions. *Geophysics*, **67**(1), 156–165.

Hill, D., Combee, C. and Bacon, J. [2006] Over/under acquisition and data processing: The next quantum leap in seismic technology? *First Break*, **24**(6).

Hole, J.A. [1992] Nonlinear High-Resolution Three-Dimendional Seismic Traveltime Tomography. *Journal of Geophysical Research*, **97**(5), 6553–6562.

Hustedt, B., Operto, S. and Virieux, J. [2004] Mixed-grid and staggered-grid finite-difference methods for frequency-domain acoustic wave modelling. *Geophysical Journal International*, **157**(3), 1269–1296.

Jo, C.h., Shin, C. and Suh, J.H. [1996] An optimal 9-point, finite-difference, frequency-space, 2-D scalar wave extrapolator. *Geophysics*, **61**(2), 529–537.

Kaipio, J. and Somersalo, E. [2007] Statistical inverse problems: discretization, model reduction and inverse crimes. *Journal of computational and applied mathematics*, **198**(2), 493–504.

Kamath, N. and Tsvankin, I. [2013] Full-waveform inversion of multicomponent data for horizontally layered VTI media. *Geophysics*, **78**(5), WC113–WC121.

Kamei, R., Pratt, R.G. and Tsuji, T. [2013] On acoustic waveform tomography of wide-angle OBS data-strategies for pre-conditioning and inversion. *Geophysical Journal International*, **194**(2), 1250–1280.

Kamei, R., Pratt, R.G. and Tsuji, T. [2014] Misfit functionals in Laplace-Fourier domain waveform inversion , with application to wide-angle ocean bottom seismograph data. *Geophysical Prospecting*, **62**(5), 1054–1074.

Komatitsch, D. and Martin, R. [2007] An unsplit convolutional perfectly matched layer improved at grazing incidence for the seismic wave equation. *Geophysics*, **72**(5), SM155–SM167.

Komatitsch, D., Vilotte, J.P., Vai, R., Castillo-Covarrubias, J.M. and Sanchez-Sesma, F.J. [1999] The spectral element method for elastic wave equations-application to 2-D and 3-D seismic problems. *International Journal for numerical methods in engineering*, **45**(9), 1139–1164.

Lailly, P. [1983] The seismic inverse problem as a sequence of before stack migrations. In: Bednar, J., Redner, R., Robinson, E. and Weglein, A. (Eds.) *Conference on Inverse Scattering: Theory and Applications*. SIAM Philadelphia, 206–220.

Lee, M.W., Balch, A.H. and Parrott, K.R. [1984] Radiation from a downhole air gun source. *Geophysics*, **49**(1), 27–36.

Leeuwen, T.V., Herrmann, F.J. and Herrmann, J. [2013] Mitigating local minima in full-waveform inversion by expanding the search space. *Geophysical Journal International*, **195**, 661–667.

Leeuwen, T.V. and Mulder, W.A. [2010] A correlation-based misfit criterion for wave-equation traveltime tomography. *Geophysical Journal International*, **182**(1), 1383–1394.

Lines, L.R. and LaFehr, E.D. [1989] Tomographic modeling of a cross-borehole data set. *Geophysics*, **54**(10), 1249–1257.

Luo, Y. and Schuster, G.T. [1991] Wave-equation traveltime inversion. *Geophysics*, **56**(5), 645–653.

Marfurt, K.J. [1984] Accuracy of finite-difference and finite-element modeling of the scalar and elastic wave equations. *Geophysics*, **49**(5), 533–549.

Métivier, L., Brossier, R., Labbé, S., Operto, S. and Virieux, J. [2014] A robust absorbing layer method for anisotropic seismic wave modeling. *Journal of Computational Physics*, **279**, 218–240.

Mora, P. [1987] Nonlinear Two-Dimensional Elastic Inversion of Multi Offset Seismic Data. *Geophysics*, **52**(9), 1211–1228.

Musgrave, M. [1970] *Crystal acoustics*. Holden-Day.

Nabighian, M.N., Ander, M.E., Grauch, V.J.S., Hansen, R.O., LaFehr, T.R., Li, Y., Pearson, W.C., Peirce, J.W., Phillips, J.D. and Ruder, M.E. [2005] Historical development of the gravity method in exploration. *Geophysics*, **70**(6), 63ND.

Nolet, G. [1987] Seismic wave propagation and seismic tomography. In: *Seismic Tomography*, Springer Netherlands, 1–23.

Operto, S., Brossier, R., Combe, L., Métivier, L., Ribodetti, A. and Virieux, J. [2014] Computationally efficient three-dimensional acoustic finite-difference frequency-domain seismic modeling in vertical transversely isotropic media with sparse direct solver isotropic media with sparse direct solver. *Geophysics*, **79**(5), T257–T275.

Operto, S., Virieux, J., Ribodetti, a. and Anderson, J.E. [2009] Finite-difference frequency-domain modeling of viscoacoustic wave propagation in 2D tilted transversely isotropic (TTI) media. *Geophysics*, **74**(5), T75–T95.

Owen, T.E., Balogh, W. and Peters, W.R. [1988] Arc discharge pulse source for borehole seismic applications. In: *58th Annual International Meeting of Society of Exploration Geophysicist, Expanded Abstracts*. 151–154.

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.

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.

Phinney, R.A. [1965] Theoretical calculation of the spectrum of first arrivals in layered elastic mediums. *Journal of Geophysical Research*, **70**(20), 5107–5123.

Pinches, G.M. and Thompson, R.P. [1990] Crosshole and downhole seismic surveys in the UK Trias and Lias. *Geological Society, London, Engineering Geology Special Publications*, **6**(1), 299–307.

Plessix, R.E. and Cao, Q. [2011] A parametrization study for surface seismic full waceform inversion in an acoustic vertical transversely isotropic medium. *Geophysical Journal International*, **185**(1), 539–556.

Pratt, R. [1990] Frequency-domain elastic wave modeling by finite differences: A tool for crosshole seismic imaging (short note). *Geophysics*, **55**(5), 626–632.

Pratt, R. [2007] Reprocessing Report: Part A - Traveltime Tomography. Tech. rep., BP Exploration and Technology Group.

Pratt, R. [2011] Waveform Tomography Course Notes. Tech. rep., University of Western Ontario.

Pratt, R. and Chapman, C. [1992] Traveltime tomography in anisotropic media-II . Application. *Geophysical Journal International*, **109**, 20–37.

Pratt, R., Hou, F., Bauer, K. and Weber, M. [2005] Waveform tomography images of velocity and inelastic attenuation from the Mallik 2002 crosshole seismic surveys. *Bulletin-Geological Survey of Canada*, (585), 122.

Pratt, R., Mcgaughey, W.J. and Chapman, C.H. [1993] Anisotropic velocity tomography : A case study in a near-surface rock mass. *Geophysics*, **58**(12), 1748–1763.

Pratt, R., Shin, C. and Hicks, G. [1998] Gauss-Newton and full Newton methods in frequency-space seismic waveform inversion. *Geophysical Journal International*, **133**, 341–362.

Pratt, R. and Shipp, R.M. [1999] Seismic waveform inversion in the frequency domain, Part 2: Fault dilineation in sediments using crosshole data. *Geophysics*, **64**(3), 902–914.

Pratt, R., Sirgue, L., Hornby, B. and Wolfe, J. [2008] Crosswell waveform tomography in fine-layered sediments: Meeting the challenges of anisotropy. In: *70th EAGE Conference & Exhibition, Rome, Italy*.

Pratt, R., Song, Z., Williamson, P. and Warner, M. [1996] Two-dimensional velocity models from wide-angle seismic data by wavefield inversion. *Geophysics*, **124**(1), 323–340.

Press, W.H., Teukolsky, S.A., Vetterling, W.T. and Flannery, B.P. [1992] *Numerical recipes in FORTRAN*. Cambridge University Press, Cambridge.

Ricker, N. [1953] The form and laws of propagation of seismic wavelets. *Geophysics*, **18**(1), 10–40.

Sadd, M. [2009] *Elasticity: theory, applications, and numerics*. Academic Press.

Sava, P.C. and Fomel, S. [2003] Angle-domain common-image gathers by wavefield continuation methods. *Geophysics*, **68**(3), 1065–1074.

Savazzi, S. and Spagnolini, U. [2008] Wireless geophone networks for high-density land acquisition: Technologies and future potential. *The Leading Edge*, **27**(7), 882–886.

Shin, C. and Cha, Y.H. [2009] Waveform inversion in the Laplace Fourier domain. *Geophysical Journal International*, **177**(1), 1067–1079.

Shin, C. and Min, D.J. [2006] Waveform inversion using a logarithmic wavefield. *Geophysics*, **71**(3), R31–R42.

Shipp, R.M. and Singh, S.C. [2002] Two-dimensional full wavefield inversion of wide-aperture marine seismic streamer data. *Geophysical Journal International*, **151**(1), 325–344.

Sirgue, L. and Pratt, R.G. [2004] Efficient waveform inversion and imaging: A strategy for selecting temporal frequencies. *Geophysics*, **69**(1), 231–248.

Smithyman, B., Pratt, R. and Hadden, S. [2015] Zephyr: Open-source Parallel Seismic Waveform Inversion in an Integrated Python-based Framework. In: *AGU Fall Meeting*.

Symes, W.W. [2008] Migration Velocity Analysis and Waveform Inversion. *Geophysical Prospecting*, **56**(6), 765–790.

Stekl, I. and Pratt, R.G. [1998] Accurate viscoelastic modeling by frequency-domain finite differences using rotated operators. *Geophysics*, **63**(5), 1779.

Tarantola, A. [1984] Inversion of seismic reflection data in the acoustic approximation. *Geophysics*, **49**(8), 1259–1266.

Tarantola, A. [1987] *Inverse problem theory: Methods for data fitting and parameter estimation*. Elsevier, Amsterdam.

Thomsen, L. [1986] Weak elastic anisotropy. *Geophysics*, **51**(10), 1954–1966.

Tikhonov, A.N. and Arsenin, V.Y. [1977] *Solutions of ill-posed problems*. Washington, DC.

Tsvankin, I. [1996] P-wave signatures and notation for transversely isotropic media: An overview. *Geophysics*, **61**(2), 467–483.

Tsvankin, I. [1997] Anisotropic parameters and P-wave velocity for orthorhombic media. *Geophysics*, **62**(4), 1292–1309.

Tsvankin, I. [2012] *Seismic signatures and analysis of reflection data in anisotropic media*. SEG Books, 3 edn.

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.

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.

Wesson, R.L. [1971] Travel-time inversion for laterally inhomogeneous crustal velocity 5iodels by. *Bulletin of the Seismological Society of America*, **61**(3), 729–746.

Williamson, P.R. [1991] A guide to the limits of resolution imposed by scattering in ray tomography. *Geophysics*, **56**(2), 202–207.

Winterstein, D.F. [1990] Velocity anisotropy terminology for geophysicists. *Geophysics*, **55**(8), 1070–1088.

Wong, J., Hurley, P. and West, G.F. [1983] Crosshole seismology and seismic imaging in crystalline rocks. *Geophysical Research Letters*, **10**(8), 686–689.

Wu, R.S. and Toksöz, M.N. [1987] Diffraction tomography and multisource holography applied to seismic imaging. *Geophysics*, **52**(1), 11–25.

Yilmaz, Ö. [2001] *Seismic data analysis*. Society of Exploration Geophysicists, Tulsa.

Zelt, C.A. and Barton, P.J. [1998] Three-dimensional seismic refraction tomography ' A comparison of two methods applied to data. *Journal of Geophysical Research*, **103**(B4), 7187–7210.

Zhang, L., Rector, J.W. and Hoversten, G.M. [2005] Finite-difference modelling of wave propagation in acoustic tilted TI media. *Geophysical Prospecting*, **53**(6), 843–852.

Zhou, C., Cai, W., Schuster, G.T. and Hassanzadeh, S. [1995] Acoustic wave-equation traveltime and waveform inversion of crosshole seismic data. *Geophysics*, **60**(3), 765–773.

Zhou, H., Zhang, G. and Bloor, R. [2006a] An anisotropic acoustic wave equation for modeling and migration in 2D TTI media. In: *76th Annual International Meeting, SEG*. 194–198.

Zhou, H., Zhang, G. and Bloor, R. [2006b] An anisotropic acoustic wave equation for VTI media. In: *68th EAGE Conference & Exhibition*, June 2006. 12–15.

Curriculum Vitae

Name:	Shaun Hadden
Post-Secondary Education and Degrees:	Queen's University Kingston, ON, Canada 2014 B. Sc. Applied Science in Geological Engineering University of Western Ontario London, ON 2016 M. Sc. Geophysics
Honours and Awards:	Dean's Scholar Queen's University 2011-2014
Related Work Experience:	Teaching Assistant The University of Western Ontario 2014 - 2016 Research Assistant The University of Western Ontario Summer 2014 Student to Research and Development ESG Solutions Summer 2013

Publications:

Hadden, S.M., Smithyman, B.R., and Pratt, R.G. [2016] Frequency-domain Acoustic Wave Modelling through TI Media in Western Canada - Implications for Anisotropic FWI. In: *78th EAGE Conference & Exhibition, Vienna, Austria.*