FULL WAVEFORM INVERSION WITH REFLECTED WAVES FOR ACOUSTIC 2D VTI MEDIA

by

Sonali Pattnaik
A thesis submitted to the Faculty and the Board of Trustees of the Colorado School of Mines in partial fulfillment of the requirements for the degree of Master of Science (Geophysics).

Golden, Colorado
Date __________________________

Signed: _________________________
Sonali Pattnaik

Signed: _________________________
Dr. Ilya Tsvankin
Thesis Advisor

Golden, Colorado
Date __________________________

Signed: _________________________
Dr. Roel Snieder
Professor and Head
Department of Geophysics
ABSTRACT

With the recent advances in seismic data acquisition, such as wide-azimuth, long-offset surveys and low-frequency sources, full-waveform inversion (FWI) has become an efficient tool in building high-resolution subsurface models. Conventional FWI relies mainly on diving waves to update the low-wavenumber components of the background model. However, such FWI algorithms may fail to provide a satisfactory model update for regions probed primarily by reflected waves. This typically occurs for deep target zones where the conventional FWI updates mostly the high-wavenumber model components due to the absence of diving waves. Reflection waveform inversion (RWI) has been developed to retrieve the intermediate-to-long wavelength model components in those deeper regions from reflection energy.

In this thesis, I highlight the limitations of conventional waveform inversion when applied to reflections-dominated seismic data and propose a new implementation of RWI for acoustic VTI (transversely isotropic with a vertical symmetry axis) media. I extend the idea of scale separation between the background and perturbation models to VTI media and use an optimized parameterization to mitigate parameter trade-offs in RWI. The proposed workflow repeatedly alternates between updating the long-wavelength model components by fixing the perturbation model and the shorter-wavelength, migration-based reflectivity update. I develop an hierarchical two-stage approach that operates with the P-wave zero-dip normal-moveout velocity $V_{nmo}$ and anisotropy coefficients $\delta$ and $\eta$. At the first stage, $V_{nmo}$ is estimated by applying the Born approximation to a perturbation model in $\delta$ to compute the corresponding reflection data. Although the algorithm does not invert for $\delta$, this parameter helps improve the amplitude fit for the employed acoustic model that ignores the elastic nature of the subsurface. At the second stage, the parameter $\eta$, which can be constrained by far-offset data, is estimated from the obtained perturbation model in $V_{nmo}$. The proposed 2D algorithm is tested on a horizontally layered VTI medium and the VTI Marmousi...
model. Application of a temporal correlation-based objective function significantly improves recovery of the long-wavelength $\eta$-component, as demonstrated on the Marmousi model.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ABSTRACT</td>
<td>iii</td>
</tr>
<tr>
<td></td>
<td>LIST OF FIGURES</td>
<td>vii</td>
</tr>
<tr>
<td></td>
<td>ACKNOWLEDGMENTS</td>
<td>ix</td>
</tr>
<tr>
<td></td>
<td>DEDICATION</td>
<td>xi</td>
</tr>
<tr>
<td></td>
<td>CHAPTER 1 INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>CHAPTER 2 PSEUDOACOUSTIC WAVEFIELD SIMULATOR</td>
<td>5</td>
</tr>
<tr>
<td>2.1</td>
<td>Forward modeling</td>
<td>6</td>
</tr>
<tr>
<td>2.2</td>
<td>Adjoint modeling</td>
<td>7</td>
</tr>
<tr>
<td>2.3</td>
<td>Modeling example</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>CHAPTER 3 RADIATION PATTERNS AND SENSITIVITY KERNELS</td>
<td>12</td>
</tr>
<tr>
<td>3.1</td>
<td>Radiation patterns</td>
<td>13</td>
</tr>
<tr>
<td>3.2</td>
<td>Anisotropic sensitivity kernels and trade-off analysis</td>
<td>15</td>
</tr>
<tr>
<td>3.3</td>
<td>Numerical Examples</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>CHAPTER 4 REFLECTION WAVEFORM INVERSION</td>
<td>19</td>
</tr>
<tr>
<td>4.1</td>
<td>Introduction</td>
<td>19</td>
</tr>
<tr>
<td>4.2</td>
<td>FWI and RWI sensitivity kernels</td>
<td>21</td>
</tr>
<tr>
<td>4.3</td>
<td>Two-stage inversion approach</td>
<td>24</td>
</tr>
<tr>
<td>4.4</td>
<td>Born approximation using parameter $\delta$</td>
<td>27</td>
</tr>
<tr>
<td>4.5</td>
<td>Correlation objective function</td>
<td>27</td>
</tr>
<tr>
<td>4.6</td>
<td>Adjoint-state method</td>
<td>28</td>
</tr>
</tbody>
</table>
4.7 SYNTHETIC EXAMPLE ...................................................... 30
4.8 Temporal-correlation objective function ................................. 36
4.9 Application to the Marmousi Model .................................... 37

CHAPTER 5 CONCLUSIONS AND RECOMMENDATIONS .................. 38

5.1 Conclusions .................................................................. 38
5.2 Recommendations .......................................................... 39

REFERENCES CITED ................................................................ 41

APPENDIX A RADIATION PATTERNS ...................................... 49
APPENDIX B BORN APPROXIMATION FOR A PERTURBATION IN δ ................................................. 52
APPENDIX C RWI GRADIENTS ................................................. 54
LIST OF FIGURES

Figure 2.1 VTI model with parameters (a) $V_{\text{nmo}}$, (b) $\delta$, and (c) $\epsilon$. The source (red dot) is at $(z, x) = (0.03, 1)$ km and the receiver (blue dot) at $(z, x) = (0.03, 4.5)$ km. The shot and receiver are in the thin subsurface elliptic layer added to suppress the S-wave artifacts.

Figure 2.2 (a) Ricker wavelet of peak frequency 15 Hz is used as the source function for the forward-simulated wavefield. (b) The data recorded at the receiver (used as the adjoint-source term).

Figure 2.3 Snapshots of the forward-simulated wavefield (left column) and the adjoint wavefield (right column) in the order of increasing time.

Figure 3.1 Radiation patterns for $V_{\text{nmo}}$ (red), $\eta$ (green), and $\delta$ (blue) plotted as a function of the scattering angle $\vartheta$.

Figure 3.2 Sensitivity kernels for perturbations in (a) $V_{\text{nmo}}$, (b) $\eta$, and (c) $\delta$ for a source (yellow dot) located $x = 1$ km, receiver (pink dot) at $x = 2$ km, and a scatterer (white dot) at $x = 1.5$ km and $z = 0.4$ km.

Figure 4.1 Wavenumber vectors $\mathbf{k}_s$ and $\mathbf{k}_r$ associated with the rays connecting the source and receiver (respectively) to a subsurface diffractor. The scattering angle is denoted by $\theta$; $n$ is the reflector normal.

Figure 4.2 Schematic representation of the experiment. Both the transmission and reflection wavepaths contribute to the full FWI sensitivity kernel.

Figure 4.3 Schematic representation of the full FWI sensitivity kernel.

Figure 4.4 Sensitivity kernels for the model in Figure 2.1. (a) The full FWI sensitivity kernel formed by both diving and reflected waves. (b) The migration ellipse, which provides high-wavenumber updates. (c) The conventional FWI sensitivity kernel formed by diving waves. (d) The source-side reflection sensitivity kernel. (e) The receiver-side reflection sensitivity kernel. (f) The RWI sensitivity kernel, which provides smooth updates for deep targets.

Figure 4.5 Parameters (a) $V_{\text{nmo}}$, (b) $\eta$, and $\delta$ of a layered VTI model.

Figure 4.6 Gradients of the objective function (equation 4.10) with respect to (b) $V_{\text{nmo}}$, (c) $\eta$, and (d) $\delta$ for the model in Figure 4.5.
Figure 4.7  First stage of the inversion: (a) the actual $V_{nmo}$-field, (b) the initial $V_{nmo}$, and (c) the inverted $V_{nmo}$. ................................. 32

Figure 4.8  Second stage of the inversion: (a) the actual $\eta$-field, (b) the initial $\eta$, and (c) the inverted $\eta$. ................................. 32

Figure 4.9  Test for the VTI Marmousi model. (a) The actual $V_{nmo}$-field, (b) the initial $V_{nmo}$, and (c) $V_{nmo}$ obtained after the first stage of RWI. ................................. 34

Figure 4.10  (a) Actual $\eta$-field for the Marmousi model and (b) $\eta$ obtained after the second stage of RWI. ................................. 35

Figure 4.11  Parameter $\eta$ estimated by RWI with (a) the conventional correlation objective function, and (b) the temporal-correlation objective function. ................................. 37
I would first like to thank my parents for always believing in me and allowing me to realize my own potential. They have provided me endless support throughout my career to achieve my dreams. Without them, I may never have gotten to where I am today.

Next, I would like to express my sincere gratitude to my advisor, Ilya Tsvankin, for his continuous guidance, support, patience and insightful research advice throughout my M.S. study. The three seismology courses taught by Ilya were very helpful to my research and are the best graduate courses in the department. It is not very easy to find an advisor who always finds time to listen and understand the problems of students. Apart from that, Ilya stays connected with his students’ research progress through A-team seminars and the constructive comments in these seminars have helped me a lot in my thesis work.

It is with immense pleasure that I acknowledge the support and help of Tariq Alkhalifah who suggested this project in my first year. He has always remained very patient with me and provided keen scientific insights to solve seemingly practical difficulties associated with this project. I greatly benefitted from our interactions during the Skype meetings, his knack to provide practical solutions to difficult problems, never ending enthusiastic supervision, valuable suggestions, and critical comments.

The Seismic Imaging class taught by Paul Sava was very insightful and helped me a lot in understanding both basic and advanced imaging and inversion techniques. The Art of Science course taught by Roel Snieder was extremely dear to me as I could learn many important ways to productively handle my graduate studies at Mines. Paul Martin and Whitney Guitton were very supportive and kind enough to be a part of my committee and provided valuable feedbacks related to my thesis. Also, special thanks to CWP for providing the necessary financial support throughout my M.S study.
My thanks also goes to Pamela Kraus, Shingo Sean Ishida, Dawn Umpleby, and Diane Witters for all their support during my time in CWP. Pam made all the paperworks and non-technical stuffs very simple for me. Along with compiling my research reports and posters for meetings, Shingo was very helpful and encouraging. Diane is one of the nicest and kindest person who has not only helped me immensely with my technical papers but also helped me adjust to the new environment. I would also like to thank John Stockwell for his help with the softwares used here. He was always there even for small hardware and software related problems.

Finally, I would also like to acknowledge the help of my fellow CWP colleagues, particularly Hui Wang, Vladimir Li, Yogesh Arora, Yuting Duan and Nishant Kamath for all their technical advice, help and taking time out to solve complex problems with me. My colleagues were one of the biggest support during my times at CWP.
For those that shall follow after.
CHAPTER 1
INTRODUCTION

Conventional seismic imaging relies on a scale separation between the macro-velocity model (long-wavelength components) built by migration velocity analysis or reflection tomography and reflectivity (short-wavelength components) obtained by migration (Gazdag, 1978; Stolt, 1978; Baysal et al., 1983; Yilmaz, 2001). FWI attempts to fill this scaling gap by integrating the processes of macro-velocity model building and migration into a single workflow when low frequency, long-offset, wide-azimuth data are available.

Over the past decade, full-waveform inversion (FWI) has emerged as an efficient data-fitting technique capable of building high-resolution velocity models (Lailly 1983; Tarantola 1984; Pratt et al., 1998; Virieux and Operto, 2009). FWI seeks to exploit the information in the waveforms provided by diving waves, precritical and, sometimes, postcritical reflections. Conventional FWI is mainly driven by the contributions of diving waves that help reconstruct the low-to-intermediate wavenumber part of the velocity model. The penetration depth of diving waves is controlled by the velocity field, the maximum source-receiver offset, and the frequency of the source. For most conventional streamer surveys, diving waves illuminate only shallow layers. At these depths, the gradient of the conventional FWI obtained by correlating the forward and adjoint wavefields produces a high-wavenumber update, as in least-squares migration. However, low-to-intermediate wavenumbers do not get updated for deep targets without including reflections.

Several image-domain methods designed to retrieve the macro-velocity model from reflection data have been proposed (Symes and Carazzone, 1991; Sava and Biondi, 2004; Yang and Sava 2011; Almomin and Biondi, 2012; Biondi and Almomin, 2012). These approaches operate in the extended domain and improve the focusing of reflections. However, image-domain inversion produces relatively low-resolution models and are computationally
expensive. These limitations can be overcome by mixed image/data domain approaches that yield better images that contain higher-resolution features (Sun and Symes, 2012; Diaz et al., 2013).

FWI is a highly nonlinear inversion technique and generally requires low-frequency data and a sufficiently accurate initial model, so that the modeled events are not shifted by more than half a period with respect to the observed events (i.e., data are not cycle-skipped). Most existing FWI algorithms use a least-squares objective function, which may have multiple local minima, especially when reflected waves dominate the data. An initial model not prone to cycle-skipping can be obtained by kinematic methods, such as migration-based traveltome tomography (Chavent et al., 1994; Clément et al., 2001; Plessix et al., 1995) and reflection tomography (Wang and Tsvankin, 2013a,b). Ma and Hale (2013) apply waveform tomography to reflection data for long-wavelength velocity recovery using dynamic warping.

In the framework of FWI, Xu et al. (2012) develop an algorithm similar to migration-based traveltome tomography, which is called reflection-waveform inversion (RWI) and is designed to invert reflection data for the long-wavelength components of the velocity model. A migration/demigration approach (Zhou et al., 2012) can be employed to update the background model along the source and receiver wavepath. The main idea behind RWI is a scale separation between the macro-velocity model and reflectivity. Such a separation leads to a two-step workflow which repeatedly alternates between updating the long-wavelength component by fixing the reflectivity (or the perturbation) model and the reflectivity update obtained by migration. S. Wang et al. (2013) implement RWI in the frequency domain and demonstrate that low-frequency data are as essential for RWI as they are for conventional FWI.

A correlation-based objective function that can handle phase delays larger than half a period may alleviate the nonlinearity of the inverse problem (Brossier et al., 2015; H. Wang et al., 2015; Chi et al., 2015). Another way to build the long-wavelength components of the model involves the wavefield decomposition method (Tang et al., 2013; F. Wang et al., 2013;
Alkhalifah, 2014; Wu and Alkhalifah, 2014a), in which the forward- and backward-scattering components are separated in the sensitivity kernel. Integrating MVA and FWI using a hybrid objective function helps estimate both the background and perturbed models simultaneously (Biondi and Almomin, 2013a; Alkhalifah and Wu, 2015). In addition, combining diving and reflected waves helps overcome the cycle-skipping problem and build a reliable macromodel for both shallow and deep targets (Zhou et al., 2015; Wu and Alkhalifah, 2015, 2016).

However, most existing RWI implementations are limited to isotropic models, which are often incapable of matching even the kinematics of seismic arrivals, especially at far offsets where anisotropy tends to have the largest impact on wave propagation. Incorporating anisotropy into the inversion scheme leads to improved data fitting, which in turn results in better convergence towards the global minimum of the objective function (Plessix 2010; Lee et al., 2010; Vigh et al., 2010). A key issue in anisotropic inversion is its multiparameter nature, which leads to trade-offs. In the framework of RWI, an appropriate choice of the medium parameters is essential in improving the resolution and reducing the inherent parameter trade-offs.

A practical implementation of anisotropic RWI should take into account the sensitivity of data to different parameters as a function of the scattering angle. Radiation patterns are an important tool to study the amplitude variation of waves scattered by parameter perturbations. Gholami et al. (2013) perform numerical analysis of the radiation patterns for acoustic VTI media. Alkhalifah and Plessix (2014) investigate radiation patterns analytically and show that a set that includes $V_{nmo}$, $\eta$, and $\delta$ is suitable for inversion of high-quality reflections.

This thesis is divided into four parts and organized as follows. In Chapter 2, I start by describing the pseudoacoustic wave equation for VTI media. Because the pseudoacoustic wave operator is not self-adjoint, I discuss both the forward and adjoint modeling operators. The modeling procedure and wavefield properties are illustrated using a 2D horizontally layered model with a VTI background. The modeling results help understand the basic
steps involved in reflection waveform inversion.

In Chapter 3, I discuss the radiation patterns for VTI media based on the model vector \( \mathbf{m} = \{V_{nmo}, \eta, \delta\} \). The radiation patterns are constructed by linearizing the constant-density pseudoacoustic wave equation using the single-scattering Born approximation and employing a high-frequency approximation of the Green’s function for smooth media. As \( V_{nmo} \) and \( \eta \) largely control the kinematics of reflected P-waves in VTI media, the discussion of radiation patterns is focused primarily on these two parameters. I also present the radiation pattern for \( \delta \), which is used as an extra parameter to compensate for inadequate amplitude fitting due to the acoustic assumption. The numerical examples illustrate the radiation patterns for a horizontal reflector. The reflection-type sensitivity kernel for the chosen parameterization is also computed to corroborate the conclusions derived from the radiation patterns.

Chapter 4 first highlights the limitations of the conventional FWI when seismic data are dominated by reflections. A sensitivity kernel decomposition is employed to understand the FWI and RWI sensitivity kernels. Then I present a new implementation of anisotropic RWI based on a two-stage algorithm to invert reflection data for the parameters of VTI media. The methodology is tested on a horizontally layered medium and the VTI Marmousi model. The results illustrate how the outcome of the inversion depends on model assumptions and the choice of the objective function.

In chapter 5, I summarize the thesis results and provide recommendations for future work.
Using the elastic anisotropic wave equation in wavefield-based inversion methods remains challenging due to the high computational cost and the complexities involved in wave-mode separation. Therefore, anisotropic FWI is typically implemented using the pseudoacoustic approximation of the elastic wave equation. Employing the P-wave dispersion relation, Alkhalifah (2000) introduced the pseudoacoustic wave equation for transversely isotropic (TI) media by setting the shear-wave velocity along the symmetry axis to zero. The P-wave dispersion relation for a constant-density VTI medium is given by:

\[ v_n^2 \tilde{k}_z^2(2v_n^2\eta(k_x^2 + k_y^2) - \omega^2) - v_n^2(2\eta + 1)\omega^2(k_x^2 + k_y^2) + \omega^4 = 0, \tag{2.1} \]

where \( \tilde{k}_z = \frac{k_z}{1 + 2\delta} \). This dispersion relation leads to a complicated fourth-order partial differential equation (PDE) involving wavefield derivatives in space and time, including the mixed space-time derivatives. Therefore, various systems of second-order coupled equations for acoustic TI media have been developed (Fletcher et al., 2008, 2009; Fowler et al., 2010; Duveneck and Bakker, 2011; Zhang et al., 2011). Pseudoacoustic wave equations have been widely used in anisotropic waveform inversion because they significantly reduce the computational cost, while accurately describing the P-wave kinematic signatures.

Here, the VTI medium is parameterized by \( V_{\text{hor}} \) (the P-wave horizontal velocity), \( V_{\text{nmo}} \) (the P-wave normal-moveout velocity) and \( V_{P0} \) (the P-wave vertical velocity). This parameterization is related to the Thomsen parameters \( \epsilon, \delta \), and the anellipticity parameter \( \eta \) \([\eta = (\epsilon - \delta)/(1 + 2\delta)]\) introduced by Alkhalifah and Tsvankin (1995) as:

\[
V_{\text{nmo}} = V_{P0}\sqrt{1 + 2\delta}, \quad V_{\text{hor}} = V_{P0}\sqrt{1 + 2\epsilon} = V_{\text{nmo}}\sqrt{1 + 2\eta}. \tag{2.2}
\]
According to Fletcher et al. (2009), pseudoacoustic wave propagation in 2D TI media can be described by the following coupled equations:

\[
\begin{align*}
\frac{\partial^2 u^p}{\partial t^2} &= V^2_{\text{hor}} H_2 u^p + V^2_{P_0} H_1 u^q, \\
\frac{\partial^2 u^q}{\partial t^2} &= V^2_{\text{nmo}} H_2 u^p + V^2_{P_0} H_1 u^q,
\end{align*}
\] (2.3)

where \(u^p\) and \(u^q\) are the solutions of the fourth-order PDE proposed by Alkhalifah (2000), and \(H_1\) and \(H_2\) are differential operators:

\[
\begin{align*}
H_1 &= \sin^2 \theta \cos^2 \phi \frac{\partial^2}{\partial x^2} + \cos^2 \theta \frac{\partial^2}{\partial z^2} + \sin 2\theta \cos \phi \frac{\partial^2}{\partial x \partial z}, \\
H_2 &= \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial z^2} - H_1;
\end{align*}
\] (2.4)

\(\theta\) is the angle between the symmetry axis and the vertical and \(\phi\) is the azimuth of the symmetry axis. For simplicity, I use the VTI form \((\theta = 0^\circ)\) of equation 2.3:

\[
\begin{align*}
\frac{\partial^2 u^p}{\partial t^2} &= V^2_{\text{hor}} \frac{\partial^2 u^p}{\partial x^2} + V^2_{P_0} \frac{\partial^2 u^q}{\partial z^2}, \\
\frac{\partial^2 u^q}{\partial t^2} &= V^2_{\text{nmo}} \frac{\partial^2 u^p}{\partial x^2} + V^2_{P_0} \frac{\partial^2 u^q}{\partial z^2}.
\end{align*}
\] (2.5)

The fields \(u^p\) and \(u^q\) include a shear-wave “artifact” caused by setting \(V_{S0}\) to zero. S-wave artifacts can be eliminated by placing sources and receivers in a purely isotropic or elliptic \((\eta = 0)\) medium (Alkhalifah, 2000; Duveneck et al., 2008).

### 2.1 Forward modeling

The coupled system 2.5 can be written in a more compact form:

\[
L \begin{bmatrix} u^p_s \\ u^q_s \end{bmatrix} - \begin{bmatrix} f^p_s \\ f^q_s \end{bmatrix} = 0,
\] (2.6)

where the subscript ‘s’ stands for the state variables, \(u^p_s\) and \(u^q_s\) are the forward-propagated wavefields, \(f^p_s\) and \(f^q_s\) are the source terms, and the operator \(L\) is defined as
\[ L = \begin{bmatrix} V_{\text{hor}}^2 \partial_{xx} - \partial_{tt} & V_{P0}^2 \partial_{zz} \\ V_{\text{nmo}}^2 \partial_{xx} & V_{P0}^2 \partial_{zz} - \partial_{tt} \end{bmatrix}. \] (2.7)

2.2 Adjoint modeling

Since the pseudoacoustic operator \( L \) is not self-adjoint, the adjoint wave equations do not retain the same form and are solved differently from those used for forward modeling. Here, the source terms and the output wavefield in the adjoint system have to be adjoint to those in the forward system. The adjoint wavefield can be evaluated from:

\[ L^T \begin{bmatrix} u_p^r \\ u_q^r \end{bmatrix} - \begin{bmatrix} d_p^r \\ d_q^r \end{bmatrix} = 0, \] (2.8)

where ‘\( r \)’ stands for the adjoint variables, \( u_p^r \) and \( u_q^r \) are the p and q adjoint wavefields, \( d_p^r \) and \( d_q^r \) are the source terms for the adjoint equations, respectively, and the adjoint operator \( L^T \) is defined as:

\[ L^T = \begin{bmatrix} \partial_{xx} V_{\text{hor}}^2 - \partial_{tt} & \partial_{xx} V_{\text{nmo}}^2 \\ \partial_{zz} V_{P0}^2 & \partial_{zz} V_{P0}^2 - \partial_{tt} \end{bmatrix}. \] (2.9)

2.3 Modeling example

Here, I apply both the forward and adjoint modeling algorithms to a simple 2D horizontally layered VTI model (Figure 2.1). The finite-difference modeling is second-order in time and fourth-order in space and employs a regular grid with the horizontal and vertical grid spacing equal to 10 m. There is an elliptic layer (\( \eta = 0 \)) at the top of the model to suppress the SV artifacts for both the forward-propagated source wavefield and back-propagated adjoint wavefield. The model parameters satisfy the pseudoacoustic stability condition (\( \epsilon \geq \delta \) or \( V_{\text{hor}} \geq V_{\text{nmo}} \)). A Ricker wavelet with a peak frequency of 15-Hz (Figure 2.2(a)) is used to
model the source wavefield.

The adjoint wavefield is generated by propagating back in time the data recorded at the receiver shown in Figure 2.2(b) (used as the adjoint source function). The adjoint wavefield interacts with the forward-simulated source wavefield (Figure 2.3) along both the transmission and reflection wavepaths. These interactions, represented in the form of transmission and reflection sensitivity kernels in section 4.2, are helpful in understanding the spatial distribution of long- and short- wavenumber model updates in FWI.
Figure 2.1: VTI model with parameters (a) $V_{nmo}$, (b) $\delta$, and (c) $\epsilon$. The source (red dot) is at $(z, x) = (0.03, 1)$ km and the receiver (blue dot) at $(z, x) = (0.03, 4.5)$ km. The shot and receiver are in the thin subsurface elliptic layer added to suppress the S-wave artifacts.
Figure 2.2: (a) Ricker wavelet of peak frequency 15 Hz is used as the source function for the forward-simulated wavefield. (b) The data recorded at the receiver (used as the adjoint-source term).
Figure 2.3: Snapshots of the forward-simulated wavefield (left column) and the adjoint wavefield (right column) in the order of increasing time.
CHAPTER 3
RADIATION PATTERNS AND SENSITIVITY KERNELS

In multiparameter FWI, the resolution is greatly influenced by the parameterization used in the inversion algorithm. A major issue with anisotropic inversion is the trade-off or crosstalk between different parameters. Trade-offs occur when different parameter combinations can explain the same data, which renders the multiparameter inversion nonunique.

To assess the sensitivity of seismic data to different parameterizations in the framework of anisotropic FWI, one can use the angle-dependent energy (“radiation pattern”) scattered by a perturbation in a particular parameter (Wu and Aki, 1985; Tarantola, 1986; Forgues and Lambaré, 1997; Virieux and Operto, 2009; Plessix and Cao, 2011; Prieux et al., 2011; Gholami et al., 2013a, 2013b). Such patterns can help in choosing the optimal parameterization, reduce the null space of the inversion and increase the parameter resolution. Plessix and Cao (2011) investigate parameter trade-offs for VTI media by using a numerical eigenvalue and eigenvector decomposition of the Hessian matrix of the objective function. They show that $\delta$ is poorly constrained by surface seismic data because of the inherent trade-off between the vertical velocity and depth. They conclude that the kinematics of the diving and reflected waves in a VTI medium is controlled just by $V_{nmo}$ and $\eta$, as was previously shown by Alkhalifah and Tsvankin (1995).

Gholami et al. (2013) perform numerical analysis of the radiation patterns for VTI media and propose a “monoparameter strategy” for data dominated by reflections, where the parameter which has the most influence is updated, while the other two parameters are kept constant. Alkhalifah and Plessix (2014) investigate radiation patterns analytically and show that a set that includes $V_{hor}$, $\eta$ and $\epsilon$ is suitable for an hierarchical implementation of FWI starting with diving waves, whereas the parameters $V_{nmo}$, $\eta$, and $\delta$ are preferable for joint inversion of reflections and diving waves.
The sensitivity of waveform inversion to anisotropy parameters can also be studied in a more qualitative way with the help of finite-frequency sensitivity kernels (Sieminski and Zhou, 2009; Zhou and Greenhalgh, 2009, 2011; Djebbi and Alkhalifah, 2013). Sensitivity kernels for the acoustic VTI equation describe the influence of parameter perturbations when the data residuals are smeared along the wavepath between a source and receiver.

In this chapter, I discuss the radiation patterns for the parameterization \( \mathbf{m} = \{V_{nmo}, \eta, \delta\} \). The patterns are constructed by linearizing the constant-density pseudoacoustic wave equation using the single-scattering Born approximation and a high-frequency approximation of the Green’s function for smooth media. The discussion is focused on the parameters \( V_{nmo} \) and \( \eta \) that largely control P-wave reflection traveltimes in VTI media. I also present the radiation pattern for \( \delta \) – the parameter used to compensate for inadequate amplitude fitting due to the acoustic assumption. This discussion helps justify parameterizing the medium in terms of \( V_{nmo}, \eta, \) and \( \delta \) for purposes of RWI (see section 4.4). Finally, I analyze parameter trade-offs using reflection-type sensitivity kernels, where the wavepath is formed by two segments: one from the source to a subsurface scatterer and the other from the scatterer to the receiver.

### 3.1 Radiation patterns

An overlap of the scattering patterns corresponding to perturbations in different parameters indicates the trade-off between these parameters for this range of angles. Insignificant overlap of scattering patterns means that the corresponding parameters can be resolved by the data. The derivation of radiation patterns is generally based on the slowness and polarization vectors obtained by replacing the Green’s functions with their asymptotic representation in the Born approximation (Wu and Aki, 1985; Gholami et al., 2013; Alkhalifah and Plessix, 2014; Kamath and Tsvankin, 2016).

By parameterizing the medium in terms of \( V_{nmo}, \eta, \) and \( \delta \), the pseudoacoustic wave equation 2.3 can be written as:
\[
\frac{1}{V_{nmo}^2} \partial_{tt} u^p - (1 + 2\eta)\partial_{xx} u^p - \frac{1}{\sqrt{1 + 2\delta}} \partial_{zz} u^p = f^p, \\
\frac{1}{V_{nmo}^2} \partial_{tt} u^q - \frac{1}{\sqrt{1 + 2\delta}} \partial_{xx} u^q - \frac{1}{(1 + 2\delta)} \partial_{zz} u^q = f^q.
\]

(3.1)

According to Zhou et al. (2006), Plessix and Cao (2011), and Alkhalifah and Plessix (2014), equation 3.1 can be transformed into an equivalent expression in terms of the parameter \( p_n \), where \( p_n = u^q \sqrt{1 + 2\delta} \), and its deviation from \( u^p \) \( (p_r = u^p - p_n) \), leading to:

\[
\frac{1}{V_{nmo}^2} p_n - \partial_{xx}(p_n + p_r) - \frac{1}{\sqrt{1 + 2\delta}} \partial_{zz}(\frac{1}{\sqrt{1 + 2\delta}} p_n) = f, \\
\frac{1}{V_{nmo}^2} p_r - 2\eta \partial_{xx}(p_n + p_r) = 0.
\]

(3.2)

The radiation patterns are derived from the Born approximation of equation 3.2 for perturbations in the parameters \( V_{nmo}, \eta, \) and \( \delta \). The perturbed parameters along with the wavefield components can be expressed as the sum of background value (subscript ‘o’) and a perturbation (subscript ‘d’) in the following way

\[
\frac{1}{V_{nmo}^2} = (1 + v^d_d); \quad \eta = \eta_o + \eta_d; \quad \delta = \delta_o + \delta_d; \\
p_n = p_no + p_{nd}; \quad p_r = p_ro + p_{rd}.
\]

(3.3)

Assuming a purely isotropic background \( (\eta_o = 0, \delta_o = 0) \) and neglecting the second-order terms, the Born approximation of equation 3.2 in the frequency domain is given by (see Appendix A):

\[
-\frac{1}{v_o^2} \omega^2 p_{nd} - (\partial_{xx}p_{nd}) = -\omega^2 \frac{1}{v_o^2} (v^d_d) p_{no} - 2 \frac{v^2_o}{\omega^2} \left[ \partial_{xx} \eta_d (\partial_{xx} p_{no}) \right] - (\delta_d \partial_{zz} p_{no} + \partial_{zz} \delta_d p_{no}).
\]

(3.4)

Using the asymptotic Green’s function, the perturbation \( p_{nd} \) can be written as:

\[
p_{nd}(x_s, x_r, \omega) = -\omega^2 \int_{V(x)} \frac{1}{v^2_o} G(x_s, x, \omega) G(x_r, x, \omega) \frac{r \cdot k}{v^2_o} dV(x),
\]

(3.5)

where \( x_s \) and \( x_r \) are the source and receiver locations, respectively, \( V(x) \) is the volume containing all scatterers, and
The elements of $k$ represent the radiation patterns (Aki and Richards, 1980; Alkhalifah and Plessix, 2014) due to the corresponding perturbed parameters given in $r$. Here, $n_s$ and $n_r$ are the unit slowness vectors for the source and receiver side, respectively, which for a horizontal reflector can expressed through the scattering angle $\theta_s$:

$$n_s = \left( \frac{n_{sx}}{n_{sz}} \right) = \left( \sin \theta_s \cos \theta_s \right) ; \quad n_r = \left( \frac{n_{rx}}{n_{rz}} \right) = \left( \sin \theta_s \cos \theta_s \right).$$

Substituting equation (3.7) into equation (3.6), the radiation patterns can be expressed as:

$$r = \left( \begin{array}{c} v_d \\ \eta_d \\ \delta_d \end{array} \right) ; \quad k = \left( \begin{array}{c} 1 \\ \frac{2n_{sx}^2 n_{rx}^2}{(n_{sz}^2 + n_{rz}^2)} \end{array} \right).$$

### 3.2 Anisotropic sensitivity kernels and trade-off analysis

The sensitivity kernel for a particular model parameter is the response in the model space to data perturbations for a single source and a single receiver. Here, I extend the trade-off analysis based on radiation patterns by using reflection-type sensitivity kernels for the pseudoacoustic VTI wave equation. In the case of waveform inversion with reflected waves, the sensitivity kernel includes two parts. One is the source-side kernel, which is the cross-correlation of the source wavefield with the adjoint-source wavefield, and the other is the receiver-side kernel which represents the cross-correlation of the residual receiver wavefield with the adjoint-receiver wavefield.

The sensitivity kernel can be computed from the pseudoacoustic VTI wave equation 2.5 using the least-squares misfit function and the adjoint state method (Tromp et al., 2005; Plessix, 2006). The source-side and receiver-side gradients are described in Appendix C.
### 3.3 Numerical Examples

For a perturbation in the NMO velocity the radiation pattern is independent of the scattering angle and, therefore, has a significant trade-off with both \( \eta \) and \( \delta \) (Figure 3.1). A perturbation in \( \delta \) produces most energy at small scattering angles, which implies that \( \delta \) is mainly associated with the vertical wavenumber for the source and receiver wavefields. At small scattering angles, as in the case of short-spread reflections, a trade-off exists between \( V_{nmo} \) and \( \delta \). On the other hand, the radiation pattern for \( \eta \) shifted toward large scattering angles and is associated with the horizontal wavenumber.

Therefore, there is limited coupling between the parameters \( \delta \) and \( \eta \) at either small or large scattering angles. This analysis suggests an hierarchical approach for reflection waveform inversion, where \( V_{nmo} \) is reconstructed first, followed by \( \eta \). In this case, \( \delta \) serves as an additional parameter used for amplitude fitting, while density is kept constant. The amplitudes of the radiation patterns for the parameters \( \eta \) and \( \delta \) are much smaller than those for \( V_{nmo} \), so the inversion gradient has to be properly scaled to ensure accurate updates in \( \eta \) and \( \delta \).

To complement the analysis of radiation patterns, I present reflection-type sensitivity kernels for the parameters \( V_{nmo} \), \( \eta \), and \( \delta \). In all three cases, the parameter perturbation is generated for a monochromatic 12-Hz wavefield. As expected, the energy distribution in the sensitivity kernel for the NMO velocity (Figure 3.1) is relatively uniform. However, for the perturbations in both \( \eta \) and \( \delta \), the sensitivity kernels vary with direction of wave propagation. The energy with \( \delta \)-sensitivity kernel is concentrated mostly near the vertical direction. In contrast, the sensitivity kernel for \( \eta \) has a higher amplitude for near-horizontal wave propagation. This analysis confirms that the NMO velocity has a trade-off with \( \eta \) near the horizontal direction and with \( \delta \) near the vertical direction. If \( \delta \) is known from additional information (e.g., check shots) and data have a sufficiently wide angular coverage, reflection waveform inversion should be able to estimate both \( V_{nmo} \) and \( \eta \).
Figure 3.1: Radiation patterns for $V_{\text{mm}}$ (red), $\eta$ (green), and $\delta$ (blue) plotted as a function of the scattering angle $\theta$. 
Figure 3.2: Sensitivity kernels for perturbations in (a) $V_{nmo}$, (b) $\eta$, and (c) $\delta$ for a source (yellow dot) located $x = 1$ km, receiver (pink dot) at $x = 2$ km, and a scatterer (white dot) at $x = 1.5$ km and $z = 0.4$ km.
CHAPTER 4
REFLECTION WAVEFORM INVERSION

Reflection waveform inversion aims to retrieve long-wavelength components of the velocity model, mostly in the deeper horizons illuminated primarily by reflections. The key components of anisotropic RWI are an accurate and efficient wavefield simulator, proper medium parameterization, a robust misfit function that prevents cycle skipping, and an efficient optimization algorithm. In this chapter, I highlight the limitations of conventional FWI in handling reflections and then discuss RWI for acoustic VTI media (the wavefield simulators were described in chapter 2). The gradient is computed from the adjoint-state method (see section 4.6). Synthetic tests illustrate how the inversion results depend on model assumptions and the choice of the objective function.

4.1 Introduction

When surface seismic data contain intensive diving waves, FWI has been successfully used to employ them in building the long-wavelength components of the subsurface model. In FWI, model updates at a subsurface point are governed by the basic principles of diffraction tomography (Devaney, 1982; Miller et al., 1987, Wu and Toksöz, 1987). In isotropic media, the wavenumber vector \( \mathbf{k} \) at a potential diffractor is related to scattering angle \( \theta \), the velocity \( v \) and the angular frequency \( \omega \) by

\[
\mathbf{k} = \mathbf{k}_s + \mathbf{k}_r = \frac{\omega}{v} \cos \left( \frac{\theta}{2} \right) \mathbf{n},
\]

where \( \mathbf{k}_s \) and \( \mathbf{k}_r \) are the source and receiver wave vectors, respectively, and \( \mathbf{n} \) is a unit vector normal to the reflector. Here, the model is assumed to be locally homogeneous with respect to the dominant wavelength \( \lambda \) at the subsurface diffractor, and the approximation is based on a plane-wave representation of the source and receiver wavefields. For anisotropic media \( \mathbf{k} \).
is not the midpoint wavenumber vector (as is the case for isotropy), even though the formula \( \mathbf{k} = \mathbf{k}_s + \mathbf{k}_r \) still holds (Figure 4.1). For 2D TI media, the wavenumber \( \mathbf{k} \) and scattering angle \( \theta \) are related by (Sava and Alkhalifah, 2013):

\[
\mathbf{k} = \mathbf{k}_s + \mathbf{k}_r = K(\theta) \mathbf{n},
\]

where \( K(\theta) \) is a certain function of the scattering angle and \( \mathbf{n} \) is a unit reflector normal. If the symmetry axis is perpendicular to the reflector, this relationship can be simplified to

\[
\mathbf{k} = \frac{\omega}{v_p(\frac{\theta}{2})} \cos \left( \frac{\theta}{2} \right) \mathbf{n},
\]

where \( v_p \) is the phase velocity.

Figure 4.1: Wavenumber vectors \( \mathbf{k}_s \) and \( \mathbf{k}_r \) associated with the rays connecting the source and receiver (respectively) to a subsurface diffractor. The scattering angle is denoted by \( \theta \); \( \mathbf{n} \) is the reflector normal.

This relationship between the wavenumber vector and scattering angle shows that for wide-azimuth acquisition, large scattering angles associated with diving waves facilitate the reconstruction of low-wavenumber components of the subsurface model. In contrast, for narrow-azimuth acquisition, small scattering angles associated with reflections contribute to high-wavenumber model updates.
FWI starts with a smooth initial velocity model free from reflections by applying only low-wavenumber updates to the background model. Because the low-wavenumber information is associated mostly with diving waves, this update is typically restricted to the shallow part of the model. In the deeper regions, FWI produces a high-wavenumber update due to the absence of diving waves but fails to properly update the low-wavenumber components. To improve the resolution at depth, longer-offset data and lower frequencies are required. Unfortunately, for most offshore surveys the maximum offset-to-depth ratios are insufficient for FWI to recover the deeper segments of the model. To achieve meaningful background model updates at depth, it is necessary to incorporate reflection data. Hence, it is important to understand the relationship between the sensitivity kernels of the conventional FWI and RWI.

4.2 FWI and RWI sensitivity kernels

The sensitivity kernel identifies the model areas that can be updated for a particular receiver location (Woodward, 1992; Snieder and Lomax, 1996; Dahlen et al., 2000). In the framework of FWI, dividing the full sensitivity kernel into its sub-kernels helps in separating the contributions from diving and reflected waves. To understand the performance of both FWI and RWI, it is convenient to study the sensitivity kernels numerically.

![Figure 4.2: Schematic representation of the experiment. Both the transmission and reflection wavepaths contribute to the full FWI sensitivity kernel.](image-url)
First, I consider the full FWI sensitivity kernel that includes the Fresnel zones associated with both diving and reflected waves (Figure 4.2). FWI starts with a smooth initial model, which is updated iteratively by minimizing an objective function, generally the least-squares norm of the difference between the modeled and recorded seismic data (Tarantola, 1987):

\[
C_1(m) = \frac{1}{2}\|d_{\text{obs}}(x_s, x_r, t) - d_{\text{cal}}(x_s, x_r, t)\|^2,
\]  

(4.3)

where \(x_s\) and \(x_r\) are the source and receiver location, respectively, \(d_{\text{obs}}(x_s, x_r, t)\) is the observed data, \(d_{\text{cal}}(x_s, x_r, t) = u_s(x, z = 0, t)\) is the modeled (calculated) data and \(m\) is the model vector. The modeled wavefield \(u_s\) can be computed from the wave equation:

\[
L(m) u_s(x, t; x_s) = f(t) \delta(x - x_s),
\]  

(4.4)

where \(L\) is the pseudoacoustic wave operator (equation 2.7) and \(f(t)\) is the source function. For a single source (\(x_s\))-receiver (\(x_r\)) pair, the gradient \(K\) corresponding to the objective function in equation 4.3 can be computed by the adjoint-state method (Plessix, 2006):

\[
K(x; x_s, x_r) = u_s(x, t; x_s) \ast u_r(x, t; x_r),
\]  

(4.5)

where \(\ast\) denotes the cross-correlation in the time domain, \(u_s(x, t; x_s)\) is the forward-modeled wavefield and \(u_r(x, t; x_r)\) is the adjoint-wavefield for the receiver. The adjoint wavefield is computed by using the data residuals for selected waves (could be direct, diving, reflected, and/or backscattered waves) as the source term. The adjoint-state method is discussed in more detail in section 4.6.

The full FWI sensitivity kernel for a homogenous VTI background model is shown in Figure 4.3. The initial model is smooth with no reflections, so that the data residual contains both diving and reflected waves. These data residuals are back-propagated along the sensitivity-kernel path to update both the low-and high-wavenumber model components. The full FWI sensitivity kernel can be written as the sum of the primary kernel \(K_p\) and
secondary kernel $K_s$:

$$K(x; x_s, x_r) = K_p(x; x_s, x_r) + K_s(x; x_s, x_r),$$

(4.6)

where

$$K_p = u^t_s(x, t; x_s) \ast u^t_r(x, t; x_r),$$

$$K_s = k_1 + k_2,$$

$$k_1 = u^t_s(x, t; x_s) \ast u^r_r(x, t; x_r) + u^t_r(x, t; x_s) \ast u^r_r(x, t; x_r),$$

$$k_2 = u^r_s(x, t; x_s) \ast u^r_r(x, t; x_r).$$

(4.7)

Here, $u^t$ is the transmitted wavefield and $u^r$ is the wavefield generated from the seismic response of reflectors. The primary kernel $K_p(x; x_s, x_r)$, formed by the correlation of the transmitted source and receiver wavefields, produces the Fresnel zone associated with diving waves and represents the main part of the conventional FWI sensitivity kernel. The secondary kernel $K_p(x; x_s, x_r)$ consists of two components, $k_1$ and $k_2$. The component $k_1$ includes the correlations of the downgoing source-side wavefield and the upgoing scattered receiver wavefield and of the upgoing scattered source wavefield and the downgoing receiver wavefield. The component $k_2$ is formed by the correlation of the scattered source and receiver wavefields and usually is negligibly small (Brossier et al., 2015, Chi et al., 2015).

The component $k_1$ produces a wide Fresnel zone (rabbit-ears) between the reflector and the surface and is the dominant part of the RWI sensitivity kernel. Because $k_1$ is formed by
the correlation of the transmission wavepaths between the surface and the reflector, it allows for the reconstruction of long-to-intermediate wavelengths at deep targets. In practice, a single-scattering Born approximation is used in computing the RWI gradient to generate the scattered wavefield for both the source and receiver sides (Xu et al., 2012b; S. Wang et al., 2013); for more details, see section 4.6.

The contributions of the FWI and RWI sensitivity kernels are illustrated in Figure 4.4. The full FWI sensitivity kernel in Figure 4.4(a) can be divided into three parts. The conventional sensitivity kernel in Figure 4.4(c), largely based on diving waves, has the highest amplitude in the shallow regions. For the deeper regions, the main contribution comes from the migration ellipse (Figure 4.4(b)), which has a higher amplitude than the RWI sensitivity kernel (Figure 4.4(f)). Therefore, removing the migration ellipse from the full sensitivity kernel highlights the RWI sensitivity kernel (rabbit-ear wavepath), which is responsible for updating the low-wavenumber components in the deeper regions of the model.

4.3 Two-stage inversion approach

The model parameters in the FWI gradient have different units and magnitude, which may prevent FWI from successful convergence. An appropriate choice of parameterization and proper scaling of the gradient accelerates the rate of convergence of the inversion algorithm. Optimal parameterization is supposed to mitigate the trade-offs between model parameters for a given type of input data. I have analyzed these trade-offs with the help of radiation patterns and reflection-type sensitivity kernels presented in Chapter 3. Based on that discussion, the model vector is defined as \( \mathbf{m} = \{V_{\text{nmo}}, \eta, \delta\} \), which helps reduce the trade-offs and improve model resolution.

The RWI approach introduced here separates the subsurface model into long-wavelength components updated via a correlation-based objective function and short-wavelength components estimated by perturbing model parameters. The algorithm is designed to estimate just the NMO velocity and parameter \( \eta \) because typically \( \delta \) is poorly constrained by P-wave reflection data. Stable inversion for \( \delta \) cannot be performed without additional (e.g.,
Figure 4.4: Sensitivity kernels for the model in Figure 2.1. (a) The full FWI sensitivity kernel formed by both diving and reflected waves. (b) The migration ellipse, which provides high-wavenumber updates. (c) The conventional FWI sensitivity kernel formed by diving waves. (d) The source-side reflection sensitivity kernel. (e) The receiver-side reflection sensitivity kernel. (f) The RWI sensitivity kernel, which provides smooth updates for deep targets.
borehole) information (Wang and Tsvankin, 2013a,b).

The method operates with P-wave reflection data and includes the following steps:

1. At small scattering angles (i.e., for small offsets), there is a trade-off between $V_{nmo}$ and $\delta$. Hence, the parameter $\delta$ is perturbed first, while keeping $V_{nmo}$ and $\eta$ fixed. At each iteration, the reference NMO velocity is updated using a demigration process for the $\delta$-perturbation model to generate the simulated data. The main goal of the first inversion stage is to eliminate the data residuals at the near offsets caused by inaccurate initial values of $\delta$ and $V_{nmo}$. Because this procedure is limited to near-offset data, it seldom suffers from cycle skipping.

2. According to the analysis of the radiation patterns and sensitivity kernels for $\eta$, that parameter is mainly associated with the horizontal wavenumber and there exists a trade-off between $V_{nmo}$ and $\eta$ in near-horizontal directions. Therefore, at the second stage, the previously updated NMO velocity is used to generate the perturbation model. At each iteration, $\eta$ is updated by applying a demigration to the $V_{nmo}$-perturbation model to generate the simulated data. The goal of the second stage is to invert for $\eta$ by eliminating the data residuals at the far offsets. With sufficiently wide angular coverage, the two-stage RWI is expected to be able to estimate both $V_{nmo}$ and $\eta$.

A correlation-based objective function, which evaluates the similarity between the observed and predicted data, is employed at both stages. The RWI gradient is computed from the adjoint-state method discussed in section 4.6. Employing $\delta$ in this approach helps compensate (at least, to a certain extent) for inadequate amplitude fitting of reflection data in the acoustic approximation. Therefore, $\delta$ is perturbed at the first inversion stage and the Born approximation for the pseudoacoustic wave equation is used to generate the scattered wavefield; this is described in the next section.
4.4 Born approximation using parameter $\delta$

Representing the parameter $\delta$ as the sum of the background value ($\delta_o$) and a perturbation ($\delta_d$), the wavefields from equation 2.6 can be expressed as $u^p = u^{po} + u^{p1}$ and $u^q = u^{qo} + u^{q1}$, where $u^{p1}$ and $u^{q1}$ are the perturbations and $u^{po}$ and $u^{qo}$ are the wavefields for the background medium.

For a small $\delta$-perturbation, equation 2.4 can be written in matrix form as (see Appendix B):

$$
L \begin{bmatrix} u^{p1} \\ u^{q1} \end{bmatrix} = \begin{bmatrix} 2\delta_d V_{P0}^2 \frac{\partial^2 u^{qo}}{\partial z^2} \\ 2\delta_d V_{P0}^2 \frac{\partial^2 u^{qo}}{\partial z^2} \end{bmatrix},
$$

(4.8)

where $L = \begin{bmatrix} V_{hor}^2 \partial_{xx} - \partial_{tt} & V_{P0}^2 \partial_{zz} \\ V_{nmo}^2 \partial_{xx} & V_{P0}^2 \partial_{zz} - \partial_{tt} \end{bmatrix}$.

Here, we refer to the term $2\delta_d V_{P0}^2$ as the “$\delta$-image.” The dot product of the $\delta$-image with the squared double-derivative of the $q$-component of the source wavefield produces secondary sources in the model space. The Born-scattered data (i.e., predicted data) are computed by forward modeling with these secondary sources using equation 4.9. Next, this $\delta$-image is used as the perturbation model to invert only for $V_{nmo}$ and get an updated NMO-velocity field. This update, based primarily on short-offset data, completes the first stage of the inversion. In contrast, existing RWI algorithms obtain the image from near-offset least-squares migration. At the second inversion stage, the estimated $V_{nmo}$ is used to generate the perturbation model and invert for $\eta$ by eliminating the far-offset data residuals.

4.5 Correlation objective function

We employ the constant-density pseudoacoustic wave equation, which does not properly model reflection amplitudes. In addition, because the actual reflectivity cannot be obtained by cross-correlating the source and receiver wavefields, amplitude matching of the observed and predicted data using a least-squares objective function may be problematic. The require-
ment of amplitude matching can be relaxed by using a normalized cross-correlation objective function $C$ that evaluates the similarity between the observed ($d_{\text{obs}}$) and Born-modeled ($d_m$) data (Routh et al., 2011a,b; Choi and Alkhalifah, 2012; Xu et al., 2012a; Liu et al., 2016):

$$C = -\sum_s \sum_r \frac{d_m}{\|d_m\|} \cdot \frac{d_{\text{obs}}}{\|d_{\text{obs}}\|}. \quad (4.9)$$

The similarity between the observed and simulated data is maximized in the process of model updating. The correlation value goes to zero when the observed and predicted data are completely out of phase, and is equal to unity when their kinematics completely match.

4.6 Adjoint-state method

The adjoint-state method (Tromp et al., 2005; Plessix, 2006) provides an efficient way of computing the derivatives of the objective function with respect to the model parameters. The method operates with the objective function, state equations, and adjoint equations. The objective function depends on the model parameters through the state equations.

The adjoint-state method involves four main steps:

(i) Computation of the state variables (forward wavefield) by solving the state equations.
(ii) Computation of the adjoint source functions.
(iii) Computation of the adjoint-state variable (adjoint wavefield) by solving the adjoint equations.
(iv) Computation of the gradient of the objective function.

The problem in hand includes two state equations: the pseudoacoustic wave equation that generates the forward wavefield and the Born approximation of that equation for a small perturbation in a certain model parameter. RWI can be posed as the following optimization problem:

$$\begin{align*}
\min_{\mathbf{m}} & \quad C(\mathbf{m}, \mathbf{u}^p, \mathbf{u}^q, \mathbf{u}^{p1}, \mathbf{u}^{q1}), \\
\text{subject to} & \quad F(\mathbf{m}, \mathbf{u}^p, \mathbf{u}^q) = 0 \& F_1(\mathbf{m}, \mathbf{u}^{p1}, \mathbf{u}^{q1}) = 0,
\end{align*} \quad (4.10)$$

28
where $u^p$ and $u^q$ are the forward-modeled wavefields and $u^{p1}$ and $u^{q1}$ are the Born-scattered wavefields. As discussed above, the functions $F$ and $F_1$ are the state equations for the forward-modeled and Born-simulated data, respectively, and $m = \{V_{\text{nmo}}, \eta, \delta\}$ is the model vector.

Using equation 4.9 and 4.10, the Lagrangian can be formulated as:

$$\Lambda = C(x, t) - \langle Lu_s - f, \lambda_{u_s} \rangle - \langle Lu_1 - Iu_s, \lambda_{u_1} \rangle,$$

where $C(x, t) = -\frac{d_m}{\|d_m\|} \cdot \frac{d_{\text{obs}}}{\|d_{\text{obs}}\|}$ is the correlation-based objective function (equation 4.9), $L$ is the wave operator, $I$ is the perturbation model, $f$ is the source term, $u_s$ and $u_1$ are the state variables, and $\lambda_{u_s}$ and $\lambda_{u_1}$ are the adjoint variables. The source wavefield $u_s$ and the Born-scattered wavefield $u_1$ are generated by solving the following equations:

$$Lu_s - f = 0,$$

$$Lu_1 - Iu_s = 0.$$

Next, the adjoint variables for both the source and receiver sides are obtained from:

$$L^T \lambda_{u_s} - I \lambda_{u_1} = 0,$$

$$L^T \lambda_{u_1} = \frac{1}{\|d_m\|} \left[ \hat{d}_m (\hat{\hat{d}}_m \cdot \hat{d}_{\text{obs}}) - \hat{d}_{\text{obs}} \right] = 0.$$

Here, $\hat{d}_m = \frac{d_m}{\|d_m\|}$, $\hat{d}_{\text{obs}} = \frac{d_{\text{obs}}}{\|d_{\text{obs}}\|}$, and the adjoint sources are:

$$r_d = \frac{1}{\|d_m\|} \left[ \hat{d}_m (\hat{\hat{d}}_m \cdot \hat{d}_{\text{obs}}) - \hat{d}_{\text{obs}} \right].$$

Equation 4.16 shows that the Born-simulated data are scaled by the dot product of the observed data and the Born-simulated data. This ties the amplitude matching to the similarity
between these two data sets. Finally, the gradient for the model parameters \( \mathbf{m} \) is computed from \( \partial \Lambda / \partial \mathbf{m} = \partial J / \partial \mathbf{m} \). The expressions for the source- and receiver-side gradients can be found in Appendix C.

4.7 SYNTHETIC EXAMPLE

The proposed two-stage inversion approach is applied to two gridded VTI models. The forward and adjoint wavefield extrapolation is carried out with a finite difference algorithm developed within the framework of MADAGASCAR. For both tests, I use the zero-lag normalized correlation-based objective function (equation 4.9).

Test 1

The algorithm is first tested on the layered VTI model in Figure 4.5. The horizontal and vertical grid spacing is 25 m. The data are excited by 16 sources positioned at 500 m intervals on the surface with a maximum source-receiver offset of 7.5 km. The receivers are also located on the surface at every grid point. The source signal is a Ricker wavelet with a central frequency of 5 Hz. The sources and receivers are embedded in a thin isotropic layer to suppress the shear-wave artifacts produced by the pseudoacoustic wave equation. The modeled wavefield is computed for a smooth background medium free from reflections, and the adjoint sources are injected back into the medium to generate the adjoint wavefield.

The gradients in Figure 4.6 are obtained by perturbing the parameter \( \delta \) and using the adjoint-state method discussed in section 4.6. Note that the gradient spreads even to the third layer without applying any illumination compensation. The initial model is isotropic \( (\delta = \eta = 0) \), with the NMO velocity equal to the velocity in the first layer. As described above, the first stage involves using the \( \delta \)-perturbation model to generate the near-offset demigrated data, which are inverted for \( V_{\text{nmo}} \). The algorithm was able to recover \( V_{\text{nmo}} \) in the second and the third layers, although the velocity in the deepest layer is somewhat overestimated (Figure 4.7).

At the second stage, the demigrated data are obtained by using the inverted \( V_{\text{nmo}} \) to generate the perturbation model. The algorithm inverts just for \( \eta \) by fitting the data at the
far offsets. The estimated \( \eta \) in the first layer is accurate because the offset-to-depth ratio for its bottom is uncommonly large \((x/z = 7)\), whereas the inverted \( \eta \)-values in the second and third layer are slightly distorted (Figure 4.8). Even better estimates of \( \eta \) could be obtained if the model contained dipping reflectors (Tsvankin, 2012). Clearly, the hierarchical inversion approach makes it possible to handle the nonlinearity of the objective function and mitigate the trade-offs between the parameters \( V_{\text{nmo}} \) and \( \eta \).

Figure 4.5: Parameters (a) \( V_{\text{nmo}} \), (b) \( \eta \), and \( \delta \) of a layered VTI model.

Figure 4.6: Gradients of the objective function (equation 4.10) with respect to (b) \( V_{\text{nmo}} \), (c) \( \eta \), and (d) \( \delta \) for the model in Figure 4.5.
Figure 4.7: First stage of the inversion: (a) the actual $V_{nmo}$-field, (b) the initial $V_{nmo}$, and (c) the inverted $V_{nmo}$.

Figure 4.8: Second stage of the inversion: (a) the actual $\eta$-field, (b) the initial $\eta$, and (c) the inverted $\eta$. 
The algorithm was also applied to the VTI Marmousi model with the acquisition geometry consisting of 20 shots and 151 receivers. The shots and receivers are distributed evenly in the horizontal direction at a depth of 30 m. The grid spacing is 10 m in both the $x$- and $z$- directions, and the time step size for the modeling is 1 ms. The synthetic data set is generated with a a finite-difference pseudoacoustic simulator and a 5-Hz Ricker wavelet. A water layer is added on top of the model to suppress the shear-wave artifact.

The initial model for RWI is elliptic ($\eta = 0$) with a linearly increasing NMO velocity in the vertical direction (Figure 4.9(a)) and a smooth $\delta$-field, which is assumed to be known (e.g., from check shots). The conventional FWI for this initial model is likely to get trapped in local minima. In contrast, after 25 iterations of RWI, the background $V_{nmo}$-field (Figure 4.9(c)) converges towards the actual model. Note that the NMO velocity has been updated even at depth, although the wavefield is reflection-dominated. This is because at this stage the algorithm fits the near-offset data free from cycle-skipping.

Inverting for the $\eta$-parameter at the second stage requires fitting the far-offset data distorted by the inaccurate initial $\eta$-values ($\eta = 0$). The $V_{nmo}$ perturbation model is obtained by using the inverted background $V_{nmo}$ (Figure 4.9 (c)). Despite the initial $\eta = 0$ being far from the actual model, the algorithm was successful in obtaining a smooth $\eta$-update (Figure 4.10 (c)) towards the actual background model. However, the accuracy of the update is reduced by cycle-skipping caused by the low quality of the initial model and the need to fit the far-offset data.
Figure 4.9: Test for the VTI Marmousi model. (a) The actual $V_nmo$-field, (b) the initial $V_nmo$, and (c) $V_nmo$ obtained after the first stage of RWI.
Figure 4.10: (a) Actual \( \eta \)-field for the Marmousi model and (b) \( \eta \) obtained after the second stage of RWI.
4.8 Temporal-correlation objective function

To alleviate the problems of cycle-skipping during the inversion for $\eta$, I use an alternative, correlation-based objective function that can handle time delays larger than half a period. A better match between the predicted and observed data can be achieved through a time shift or an extension in the time domain. A time-lag normalized cross-correlation function can be defined as (Brossier et al., 2014; Chi et al., 2015):

$$X(\tau, h) = -\frac{\int d_m(t, h)d_{obs}(t + \tau, h) \, dt}{\|d_m(t, h)\| \|d_{obs}(t + \tau, h)\|}, \quad (4.17)$$

where $\tau$ is the time lag, $d_m(t, h)$ and $d_{obs}(t, h)$ are the predicted and observed data, respectively, recorded at the source-receiver offset $h$ and time $t$. The corresponding objective function can be written as (Leeuwen and Mulder, 2008, Brossier et al., 2014):

$$C = \sum_h \sum_\tau \frac{1}{2} \left( X(\tau, h) \right)^2. \quad (4.18)$$

The adjoint sources for the function in equation 4.18 can be derived using the adjoint state method:

$$r_d = J(\tau, h) \left( \frac{d_{obs}(t + \tau, h)}{\|d_m(t, h)\| \|d_{obs}(t + \tau, h)\|} - \frac{d_m(t, h)J(\tau, h)}{\|d_m(t, h)\|^3 \|d_{obs}(t + \tau, h)\|} \right), \quad (4.19)$$

where

$$J(\tau, h) = \sum_t d_m(t, h) d_{obs}(t + \tau, h). \quad (4.20)$$

The computational cost of inversion stays the same as that for the zero-lag normalized correlation-based objective function (equation 4.9). The only change in the gradient involves a different adjoint-source term.
4.9 Application to the Marmousi Model

The goal of this test is to demonstrate that the temporal-correlation objective function (equation 4.18) helps invert for $\eta$ by minimizing the influence of cycle-skipping at the far offsets. RWI is applied with the same initial $\eta$-model that was used in the previous section where the inversion was based on the zero-lag normalized correlation-based objective function. The addition of the time lag to the objective function leads to a significant improvement in the background $\eta$-update, mostly for the deeper part of the model. Indeed, the time-lag cross-correlation objective function can properly handle delays greater than half a period, which relaxes the requirements for the high accuracy of the initial model.

Figure 4.11: Parameter $\eta$ estimated by RWI with (a) the conventional correlation objective function, and (b) the temporal-correlation objective function.
In this chapter I summarize the thesis results and provide some recommendations for future work.

5.1 Conclusions

I developed a two-stage RWI (reflection waveform inversion) algorithm for VTI media designed to incorporate P-wave reflections into acoustic waveform inversion. To reduce the nonlinearity of the inverse problem, fitting of near-offset data is performed first, followed by inversion of long-offset reflections. To analyze the trade-offs inherent in such multiparameter inversion, I studied the radiation patterns and reflection sensitivity kernels for perturbations in the VTI parameters $V_{nmo}$, $\eta$, and $\delta$.

These trade-offs are mitigated by the proposed hierarchical implementation of RWI where the inversion for the normal-moveout velocity $V_{nmo}$ is followed by $\eta$-estimation, with the coefficient $\delta$ used to improve amplitude fitting. At the first stage of RWI, a $\delta$-perturbation model is employed to estimate $V_{nmo}$ from near-offset data, which makes the inversion less sensitive to cycle-skipping. Then the updated NMO velocity is used to generate a perturbation model, which includes far-offset velocity information needed to constrain $\eta$ at the second inversion stage. At both stages, the parameter fields are decomposed into the background and perturbation models and data are generated by applying Born modeling for the corresponding parameter perturbations. This strategy has proved to be effective in recovering low-wavenumber model components from reflection data.

The employed correlation-based objective function reduces the sensitivity to amplitude errors because it evaluates the general similarity between the observed and Born-modeled data. Synthetic examples demonstrate that the hierarchical inversion approach mitigates the multimodal nature of the objective function and the trade-offs between the model parameters.
Testing for layered VTI media proves that the algorithm can resolve the interval parameters $V_{nmo}$ and $\eta$ from P-wave reflection waveforms. As is the case in moveout analysis, either dipping interfaces or large offset-to-depth ratios are required to estimate $\eta$ with sufficient accuracy.

A synthetic test for the more complicated VTI Marmousi model shows that the background velocity $V_{nmo}$ can be recovered even when the initial model is relatively simple and significantly deviates from the actual velocity field. This is an important advantage of the developed algorithm that employs a $\delta$-perturbation model to mitigate cycle-skipping in fitting the near-offset data. The recovery of the long-wavelength component of $\eta$ by fitting the far-offset data is more sensitive to cycle-skipping problem if the initial $\eta$-model is inaccurate. This issue could be partially addressed by using a temporal correlation-based objective function designed to handle cycle-skipped data. Overall, the proposed method has shown the potential to recover the long-wavelength components of $V_{nmo}$ and $\eta$, while mitigating the trade-offs and nonlinearity-related issues during model updating.

5.2 Recommendations

The conventional pseudoacoustic wave equations become unstable for $\eta < 0$ and for highly complex, heterogeneous TI media with varying interface dips and azimuths. Moreover, the pseudoacoustic approximation produces shear-wave artifacts that contaminate modeled data and inversion gradients. These artifacts can be eliminated by placing the sources and receivers in an isotropic or elliptic layer or by using the generalized pseudospectral methods, but P-waves may still be converted into shear modes while propagating through complex media. To incorporate shear-wave information, it is necessary to use the more complicated elastic wave equation. However, employing elastic wave equation is computationally expensive and requires mode separation and development of advanced imaging conditions.

The proposed RWI method inverts reflection data for $V_{nmo}$ and $\eta$ in two stages to mitigate the nonlinearity of the inverse problem. Employing scattering-angle filters to condition the parameter gradients can also reduce the nonlinearity and address trade-off-related issues.
(Alkhalifah et al., 2015). Additionally, incorporating information from both diving and reflected waves into a single objective function to invert for $V_{nmo}$ and $\eta$ can increase the efficiency of model building. A parameter set that includes $V_{\text{hor}}$, $\eta$, and $\epsilon$ may be worthwhile to explore in the RWI framework because it does not require an hierarchical implementation (Alkhalifah and Plessix, 2014; Alkhalifah and Wu, 2016). Other parameter combinations can also be analyzed for different types of input data using radiation patterns and sensitivity kernel analysis (Plessix and Cao, 2011; Gholami et al., 2013; Alkhalifah and Plessix, 2014).

The efficiency of RWI implementation and convergence of the optimization algorithm depends on proper use of core-memory resources available on modern computational platforms. For example, RWI can be carried out for a few selected frequencies (Sirgue and Pratt, 2004) with a parallelized frequency-domain algorithm that uses a direct matrix solver implemented on distributed-memory platforms (S. Wang et al., 2011; 2013).

To incorporate multiply-scattered waves into the RWI optimization scheme, it is necessary to estimate the inverse Hessian matrix via the Gauss-Newton or exact Newton algorithms. Recent research shows that implementation of the truncated Newton framework can substantially increase the rate of convergence of waveform inversion (Métivier et al., 2014). Also, application of preconditioners and regularization-based misfit functions can make the algorithm more robust, while improving its convergence rate.

Finally, extending the proposed RWI technique to 3D is highly challenging due to its computational cost, the need to account for 3D parameter variations, and parameter trade-offs. Also, application of this method to field data will require careful quality control, accurate estimation of the initial model, and case-by-case parameterization analysis.
REFERENCES CITED


———, 2015a, Conditioning the full-waveform inversion gradient to welcome anisotropy: Geophysics, 80, R111–R122.


——–, 2013a, Tomographic full-waveform inversion (tfwi) by combining fwi and wave-equation migration velocity analysis: The Leading Edge, 32, 1074–1080.

——–, 2013b, Tomographic full waveform inversion (tfwi) by extending the velocity model along the time-lag axis, in SEG Technical Program Expanded Abstracts 2013: Society of Exploration Geophysicists, 1031–1036.


Claerbout, J., 1985, Imaging the earth’s interior: Blackwell scientific publications.


——–, 2013b, Ray-based gridded tomography for tilted transversely isotropic media: Geophysics, 78, C11–C23.


——–, 2014b, Spectral implementation of full waveform inversion based on reflections: Presented at the 76th EAGE Conference and Exhibition 2014.


APPENDIX A
RADIATION PATTERNS

The constant-density pseudoacoustic wave equation (Duveneck et al., 2008) can be written in the following way by parameterizing the medium in terms of $V_{\text{nmo}}$, $\eta$, and $\delta$:

\[
\frac{1}{V_{\text{nmo}}^2} \partial_{tt} u^p - (1 + 2\eta) \partial_{xx} u^p - \frac{1}{\sqrt{1 + 2\delta}} \partial_{zz} u^q = f^p
\]
\[
\frac{1}{V_{\text{nmo}}^2} \partial_{tt} u^q - \frac{1}{\sqrt{1 + 2\delta}} \partial_{xx} u^p - \frac{1}{(1 + 2\delta)} \partial_{zz} u^q = f^q.
\]

(A.1)

According to Zhou et al. (2006), Plessix and Cao (2011), and Alkhalifah and Plessix (2014), equation A.1 can be transformed into an equivalent expression in terms of $p_n$ (where $p_n = u^q\sqrt{1 + 2\delta}$) and its deviation from $u^p$ ($p_r = u^p - p_n$), leading to

\[
\frac{1}{V_{\text{nmo}}^2} p_n - \partial_{xx} (p_n + p_r) - \frac{1}{\sqrt{1 + 2\delta}} \partial_{zz} (\frac{1}{\sqrt{1 + 2\delta}} p_n) = f;
\]
\[
\frac{1}{V_{\text{nmo}}^2} p_r - 2\eta \partial_{xx} (p_n + p_r) = 0.
\]

(A.2)

The radiation patterns are derived from the first-order Born approximation of equation A.2 due to perturbations in the parameters $V_{\text{nmo}}$, $\eta$, and $\delta$. The perturbed parameters along with the wavefield components can be expressed as the sum of their background values (subscript ‘o’) and a perturbation (subscript ‘d’):

\[
\frac{1}{V_{\text{nmo}}^2} = \frac{(1 + v_d)}{v_o^2}, \quad \eta = \eta_o + \eta_d, \quad \delta = \delta_o + \delta_d;
\]
\[
p_n = p_{no} + p_{nd}, \quad p_r = p_{ro} + p_{rd}.
\]

(A.3)

Assuming a purely isotropic background ($\eta_o = 0, \delta_o = 0$) and neglecting the second-order terms, the background values $p_{no}$ and $p_{ro}$ in the frequency domain can be found as (Alkhalifah and Plessix, 2014):
\[-\frac{1}{\nu_o^2}\omega^2 p_{\text{no}} - (\partial_{xx} p_{\text{no}} + \partial_{zz} p_{\text{no}}) = f, \]  
\[p_{\text{ro}} = 0. \tag{A.4}\]

Since \(p_{\text{ro}} = 0\), the scattered wavefield \(p_{\text{rd}}\) depends only on \(p_{\text{no}}\) and can be obtained from:

\[-\frac{1}{\nu_o^2}\omega^2 p_{\text{rd}} = 2\eta_d(\partial_{xx} p_{\text{no}}) \tag{A.5}\]

Replacing the parameter \(p_{\text{rd}}\) in the Born equation yields an equation for the scattered wavefield \(p_{\text{nd}}\):

\[-\frac{1}{\nu_o^2}\omega^2 p_{\text{nd}} - (\partial_{xx} p_{\text{nd}}) = -\omega^2 \frac{1}{\nu_o^2}(v_d) p_{\text{no}} - 2\frac{v_o^2}{\omega^2} \left[ \partial_{xx} \eta_d(\partial_{xx} p_{\text{no}}) \right] - (\delta_d \partial_{zz} p_{\text{no}} + \partial_{zz} \delta_d p_{\text{no}}). \tag{A.6}\]

Equation A.2 can be solved in the frequency domain using the representation theorem:

\[p_{\text{n}}(x', x_r, \omega) = \int_{V(x')} k(x', \omega) G(x_r, x', \omega) \, dV(x'), \tag{A.7}\]

where \(k\) is the force density, \(V(x')\) is the volume that includes all sources, and \(G(x_r, x', \omega)\) is the Green’s function for the source at \(x'\) and receiver at \(x_r\) given by

\[\left[ -\frac{1}{\nu_o^2}\omega^2 - (\partial_{xx} + \partial_{zz}) \right] G(x, x', \omega) = \delta(x - x'). \tag{A.8}\]

The background wavefield component \(p_{\text{no}}\) can be expressed through the force applied at the source location \(x_s\) and the Green’s function:

\[p_{\text{no}}(x_s, x_r, \omega) = f(x_s, \omega) G(x', x_s, \omega). \tag{A.9}\]

Using equations A.9 and A.7 and employing the reciprocity theorem, i.e., \(G(x', x_s, \omega) = G(x_s, x', \omega)\), the scattered wavefield \(p_{\text{nd}}\) can be expressed as:

\[p_{\text{nd}}(x_s, x_r, \omega) = -\int_{V(x)} f(x_s, \omega) \frac{\omega^2}{\nu_o^2}(v_d) G(x_s, x, \omega) G(x_r, x, \omega) \, dV(x) \]
\[ -\int_{V(x)} \frac{2\nu_o^2}{\omega^2} \eta_d G_{xx}(x_s, x, \omega) G_{xx}(x_r, x, \omega) \, dV(x) \]
\[ -\int_{V(x)} \delta_d (G(x_s, x, \omega) G_{zz}(x_r, x, \omega) + G_{zz}(x_s, x, \omega)) G(x_r, x, \omega)). \tag{A.10}\]
Under a high-frequency approximation, the Green’s functions in equation A.10 can be replaced by their asymptotic versions (Vavryčuk, 2007):

\[
G(x_s, x, \omega) = K_s \exp \left[ \frac{i \pi}{2} \sigma_0 + i \omega \frac{n_s}{v_s} \cdot (x_s - x') \right],
\]
\[
G(x_r, x, \omega) = K_r \exp \left[ \frac{i \pi}{2} \sigma_0 + i \omega \frac{n_r}{v_r} \cdot (x_r - x') \right],
\]

(A.11)

where \( K \) is a function of the group velocity and distance along the ray, \( \sigma_0 \) is a function of the Gaussian curvature of the slowness surface, \( n_s \) and \( n_r \) are the unit slowness vectors for the source and receiver side, respectively, and \( v \) is the phase velocity. Employing the WKBJ approximation, the derivatives in equation A.10 are computed only for the exponentials in the Green’s functions (Aki and Richards, 2002). Equation A.10 can be rewritten as:

\[
p_{nd}(x_s, x_r, \omega) = -\omega^2 \int_{V(x)} f(x_s, \omega) \frac{G(x_s, x, \omega)G(x_r, x, \omega)}{v_o^2} \mathbf{r} \cdot \mathbf{k} \, dV(x),
\]

(A.12)

where

\[
\mathbf{r} = \begin{pmatrix} v_d \\ \eta_d \\ \delta_d \end{pmatrix}; \quad \mathbf{k} = \begin{pmatrix} 1 \\ \frac{2n^2_{sx}n_{rz}}{(n^2_{sx} + n^2_{sz})} \end{pmatrix}.
\]

(A.13)

Here, \( n_s \) and \( n_r \) for a horizontal reflector can be given as functions of the scattering angle \( \theta_s \):

\[
n_s = \begin{pmatrix} n_{sx} \\ n_{sz} \end{pmatrix} = \begin{pmatrix} \sin \theta_s \\ \cos \theta_s \end{pmatrix}; \quad n_r = \begin{pmatrix} n_{rx} \\ n_{rz} \end{pmatrix} = \begin{pmatrix} \sin \theta_s \\ -\cos \theta_s \end{pmatrix}.
\]

(A.14)

Substituting equation A.14 into equation A.13, the radiation patterns for the parameterization \( \{V_{nmo}, \eta, \delta\} \) can be found as:

\[
r = \begin{pmatrix} v_d \\ \eta_d \\ \delta_d \end{pmatrix}; \quad k = \begin{pmatrix} 1 \\ 2 \sin^4 \theta \\ 2 \cos^2 \theta \end{pmatrix}.
\]

(A.15)
APPENDIX B

BORN APPROXIMATION FOR A PERTURBATION IN $\delta$

The coupled system of the pseudoacoustic wave equations for VTI media can be written as (equation 2.6):

$$\mathbf{L} \begin{bmatrix} u_p^s \\ u_q^s \end{bmatrix} - \begin{bmatrix} f_p^s \\ f_q^s \end{bmatrix} = 0,$$

(B.1)

where $u_p^s$ and $u_q^s$ are the forward-propagated wavefield variables, $f_p^s$ and $f_q^s$ are the corresponding source terms, and the operator $\mathbf{L}$ is defined as:

$$\mathbf{L} = \begin{bmatrix} V_{\text{hor}}^2 \partial_{xx} - \partial_{tt} & V_{\text{nmo}}^2 \partial_{zz} \\ V_{\text{nmo}}^2 (\partial_{xx}) & V_{\text{hor}}^2 \partial_{zz} - \partial_{tt} \end{bmatrix}.$$

(B.2)

The adjoint wavefield can be evaluated by computing the adjoint of equation B.2:

$$\mathbf{L}^T \begin{bmatrix} u_p^r \\ u_q^r \end{bmatrix} - \begin{bmatrix} d_p^r \\ d_q^r \end{bmatrix} = 0,$$

(B.3)

where $u_p^r$ and $u_q^r$ are the adjoint wavefield variables, $d_p^r$ and $d_q^r$ are the source terms for the adjoint equations, and the adjoint operator $\mathbf{L}^T$ is:

$$\mathbf{L}^T = \begin{bmatrix} (\partial_{xx})V_{\text{hor}}^2 - \partial_{tt} & (\partial_{xx})V_{\text{nmo}}^2 \\ \partial_{zz}V_{\text{hor}}^2 & \partial_{zz}V_{\text{nmo}}^2 - \partial_{tt} \end{bmatrix}.$$

(B.4)

The parameter $\delta$ can be represented as the sum of the background value ($\delta_o$) and a perturbation ($\delta_d$) (see equation A.3):

$$\delta = \delta_o + \delta_d,$$

$$1 + 2\delta = 1 + 2\delta_o + 2\delta_d \approx (1 + 2\delta_o)(1 + 2\delta_d).$$

(B.5)

(B.6)

The wavefields can be expressed as $u_p^s = u_{p_o}^s + u_{p_1}^s$ and $u_q^s = u_{q_o}^s + u_{q_1}^s$, where $u_{p_o}^s$ and $u_{q_o}^s$ are computed for the background medium and $u_{p_1}^s$ and $u_{q_1}^s$ are the perturbations. Then the
pseudoacoustic wave equation B.1 can be rewritten as:

\[
\begin{bmatrix}
V_{\text{hor}}^2 \partial_{xx} - \partial_{tt} & V_{\text{nmo}}^2 \\
V_{\text{nmo}}^2 (\partial_{xx}) & (1 + 2\delta_d)/(1 + 2\delta_d) \partial_{zz}
\end{bmatrix}
\begin{bmatrix}
u_{sp} + \nu_{s1}^{p} \\
u_{s}^{q} + \nu_{s1}^{q}
\end{bmatrix}
- \begin{bmatrix}
f_{sp} \\
f_{s}^{q}
\end{bmatrix} = 0.
\]

(B.7)

Because \(\delta_d\) is assumed to be small, we can use a linear Taylor-series expansion:

\[
\begin{bmatrix}
V_{\text{hor}}^2 \partial_{xx} - \partial_{tt} & V_{\text{nmo}}^2 (1 - 2\delta_d)/(1 + 2\delta_d) \partial_{zz} \\
V_{\text{nmo}}^2 (\partial_{xx}) & (1 + 2\delta_d)/(1 + 2\delta_d) \partial_{zz} - \partial_{tt}
\end{bmatrix}
\begin{bmatrix}
u_{sp} + \nu_{s1}^{p} \\
u_{s}^{q} + \nu_{s1}^{q}
\end{bmatrix}
- \begin{bmatrix}
f_{sp} \\
f_{s}^{q}
\end{bmatrix} = 0.
\]

(B.8)

Then the Born approximation of the pseudoacoustic wave equation takes the form:

\[
\begin{bmatrix}
V_{\text{hor}}^2 \partial_{xx} - \partial_{tt} & V_{\text{nmo}}^2 P_{0}^2 \partial_{zz} \\
V_{\text{nmo}}^2 (\partial_{xx}) & V_{\text{nmo}}^2 P_{0}^2 (\partial_{zz} - \partial_{tt})
\end{bmatrix}
\begin{bmatrix}
u_{sp}^{1} \\
u_{s}^{q1}
\end{bmatrix}
= \begin{bmatrix}
2\delta_{d}V_{0}^2 \frac{\partial^2 u_{s}^{qo}}{\partial z^2} \\
2\delta_{d}V_{0}^2 \frac{\partial^2 u_{s}^{qo}}{\partial z^2}
\end{bmatrix}.
\]

(B.9)
APPENDIX C

RWI GRADIENTS

The source-side gradients for $V_{nmo}$, $\eta$, and $\delta$ can be found from equation 4.11 as:

$$\frac{\partial J_s}{\partial V_{nmo}} = \int \int 2 u_{sx}^p V_{nmo}(1 + 2\eta) u_{rx}^p \, ds \, dt + \int \int 2 u_{sz}^q \frac{V_{nmo}}{(1 + 2\delta)} u_{r2}^q \, ds \, dt$$

$$+ \int \int 2 u_{sx}^p V_{nmo} u_{rx}^q \, ds \, dt + \int \int 2 u_{sz}^q \frac{V_{nmo}}{(1 + 2\delta)} u_{r2}^q \, ds \, dt,$$

(C.1)

$$\frac{\partial J_s}{\partial \eta} = \int \int 2 u_{sx}^p V_{nmo}^2 u_{rx}^p \, ds \, dt,$$

(C.2)

$$\frac{\partial J_s}{\partial \delta} = \int \int -2 u_{sx}^q \frac{V_{nmo}^2}{(1 + 2\delta)^2} u_{r2}^q \, ds \, dt + \int \int -2 u_{sz}^q \frac{V_{nmo}^2}{(1 + 2\delta)^2} u_{r2}^q \, ds \, dt.$$

(C.3)

The receiver-side gradients for $V_{nmo}$, $\eta$ and $\delta$ are:

$$\frac{\partial J_r}{\partial V_{nmo}} = \int \int 2 u_{sx}^{p1} V_{nmo}(1 + 2\eta) u_{rx}^{p1} \, ds \, dt + \int \int 2 u_{sz}^{q1} \frac{V_{nmo}}{(1 + 2\delta)} u_{r2}^{p1} \, ds \, dt$$

$$+ \int \int 2 u_{sx}^{p1} V_{nmo} u_{rx}^{q1} \, ds \, dt + \int \int 2 u_{sz}^{q1} \frac{V_{nmo}}{(1 + 2\delta)} u_{r2}^{q1} \, ds \, dt,$$

(C.4)

$$\frac{\partial J}{\partial \eta} = \int \int 2 u_{sx}^{p1} V_{nmo}^2 2 u_{rx}^{p1} \, ds \, dt,$$

(C.5)

$$\frac{\partial J}{\partial \delta} = \int \int -2 u_{sx}^{q1} \frac{V_{nmo}^2}{(1 + 2\delta)^2} u_{r2}^{p1} \, ds \, dt + \int \int -2 u_{sz}^{q1} \frac{V_{nmo}^2}{(1 + 2\delta)^2} u_{r2}^{p1} \, ds \, dt.$$

(C.6)

Here, the integration for both the source- and receiver-side gradients is performed over the sources and time.