EXTENDED IMAGING AND TOMOGRAPHY UNDER TWO-WAY OPERATORS

by

Esteban F. Díaz Pantin

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 Doctorate degree (Geophysics).

Golden, Colorado

Date _____

Signed: _____
Esteban F. Díaz Pantin

Signed: _____
Dr. Paul C. Sava
Thesis Advisor

Golden, Colorado

Date _____

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

ABSTRACT

Seismic images reconstructed by migration from seismic data recorded at the surface represent the earth structure at every location in the subsurface. The data are remapped into the subsurface by using a wave equation and an estimate of the earth velocity. The velocity model is of great importance for obtaining accurate seismic images and for understanding the subsurface reservoir properties and conditions. In this thesis, I focus on structure and velocity estimation from surface seismic data. In order to invert for the velocity model, I seek to improve the synchronization and consistency of the extrapolated surface seismic data inside the subsurface. I use correlation functions to measure the synchronization of seismic wavefields with extended images in the image-domain, and with local correlations in the data-domain. Both correlation methods are conceptually similar and yield, after penalization, the estimate of synchronization. The penalty operator objective is to highlight areas of the correlation gathers that need to be improved through optimization of the model parameters. I incorporate, in both subsurface and data correlations, a penalty operator that takes into account the geometry of the data and its bandwidth. Another important aspect of the subsurface imaging process are the wavefields themselves. Most of the wavefields used in imaging and tomography disregard internal and surface-related multiples. Conventional methods fail to handle multiple scattered waves; hence, the multiples act as noise in the conventional and extended images producing an important bias towards low velocities as an outcome. I compute extended images using wavefields reconstructed with single scattering assumptions and wavefields that correctly handle multiple scattered waves. I show how the extended images with multiples contain similar sensitivity to the background velocity model as the conventional extended images. Additionally, the obtained extended images do not introduce a bias towards low velocities as is the case with conventional extended images.

TABLE OF CONTENTS

[1]CHAPTER SUBMITTED TO GEOPHYSICS
[2]IN PREPARATION FOR GEOPHYSICS

[3]IN PREPARATION FOR GEOPHYSICS

[4]IN PREPARATION FOR GEOPHYSICS

LIST OF FIGURES

LIST OF ABBREVIATIONS

Reverse Time Migration . . . . . RTM

Migration Velocity Analysis . . . . . MVA

Full Waveform Inversion . . . . . FWI

Center for Wave Phenomena . . . . . CWP

Conventional Imaging Condition CIC

Extended Imaging Condition . . . . . EIC

Surface Related Multiple Elimination . . . . . SRME

Common image gather . . . . . CIG

Common image point . . . . . CIP

ACKNOWLEDGMENTS

A little bit more than five years ago a received a Skype call from Paul Sava saying that I was accepted into CWP. My wife (she was accepted into RCP) and I could not be happier about going back to school, to Colorado School of Mines! My experience at Mines, and particularly at CWP did not prove me wrong. It has been an amazing journey of learning, thinking, friendship, and development.

I want to thank Paul Sava for all these years of his guidance, suggestions, and discussions. I appreciate the amount of freedom and team projects he encouraged us to take on. That made us build a technology infrastructure, based on team collaboration, that we use on a daily basis.

I would also like to thank the faculty at CWP from whom I learned a great deal. The courses I took and discussions I had with Dave Hale, Roel Snieder, Ilya Tsvankin, Ken Larner, and Paul Sava where extremely important for my development as student. On the same line, I thank the discussions, ideas, and coursework I learned from my committee: Yaoguo Li, Luis Tenorio, Kamini Singha, and Dave Hale.

I learned a great deal from my friends at CWP. Our open environment with white boards all around was a medium for great discussions. When I arrived, I had many discussions with senior students like Francesco Perrone, Clément Fleury, Tongning Yang, and Simon Luo. I want to acknowledge the great discussions and collaboration I had with my CWP friends Satyan Singh, Nishant Kamath, Yuting Duan, Oscar Jarillo, and Vladimir Li.

The friends I made in all these years are a big motor. The hikes, ski trips, discussions, beers, tennis, intramural, volleyball, football (the original), among many other experiences we shared made these past years a big highlight in my life! I hope our paths cross often in the future.

I would like to thank the staff of CWP and the Department. Pamela Kraus was extremely diligent in making our lives as students very smooth, so we could only focus on research. Shingo Ishida did a great job getting our name "out there" in the web. I appreciate the many discussions we had during lunch or coffee breaks in his office. John Stockwell is a fantastic source of anecdotes and mathematical (and geophysical) wisdom. Diane Witters is very important as a friend and writing coach, I really learned a lot from her.

Finally, I want to thank my family for their unconditional support. To my beloved wife, Carla Carvajal, for her laughs, love, and always being there with me in the ups and downs. To Lara, my little baby, for her generous smiles which have been like rocket-grade fuel during these last months.

To Carla and Lara

CHAPTER 1

INTRODUCTION

Two-way propagators are routinely used in reverse time migration (RTM) (Baysal et al., 1983; Whitmore, 1983; McMechan, 1983). In general, and especially in complex geological settings, the use of two-way wavefields produce better images than other methods such as Kirchhoff migration (Schneider, 1978; Bleistein, 1987) and one-way equation migration (Gazdag, 1978; Gazdag and Sguazzero, 1984). RTM's propagation engine, a two-way wave equation, makes this imaging method robust and accurate because it honors the kinematics of the wave phenomena by allowing waves to propagate in all directions regardless of the complexity in the velocity model. However, a drawback of the one-way migration, Kirchhoff, or RTM propagation engines is that they fail to properly handle multiple-scattered waves. Another type of propagator, called Marchenko (Broggini and Snieder, 2012; Behura et al., 2014; Wapenaar et al., 2014b; Singh et al., 2015a), is based on an integral solution to the inverse scattering problem first proposed by Rose (2002a,b). This modeling operator correctly handle primaries and multiple scattered waves; hence, it makes use of the full-scattered nature of the recorded seismic data by redatuming the surface observations inside the medium given the kinematics provided by the background model. In this thesis, I use the two-way equation and the Marchenko modeling frameworks to either assess the quality of the background model or to estimate updates in the background velocity model, which is the main driver for the quality of the seismic image.

In seismic imaging, the velocity model acts as the lens that focuses the seismic energy in the subsurface. As the lens gets more accurate, the retrieved seismic image becomes more focused. In conventional images, the focusing can be observed as diffractions are correctly colapsed into image points. In practice, the accuracy of the velocity model is more often measured by analyzing the consistency of seismic images as a function of different points of

view. When the velocity model is correct, the position of reflectors in the image remains invariant with respect to the observation point following the so called semblance-principle (Al-Yahya, 1989). This extra image dimension (extension) for measuring consistency can take different forms. In the context of two-way operators, the extension is usually formed in the image domain (Rickett and Sava, 2002; Sava and Fomel, 2006; Sava and Vasconcelos, 2011a), in the angle domain (Sava and Fomel, 2003; Yoon et al., 2004; Jin et al., 2014; Yoon et al., 2011; Vyas et al., 2011), or in the surface offset domain (Giboli et al., 2012). Once a suitable extension is chosen, the velocity error information can be extracted by measuring focusing in extended gathers (Shen and Symes, 2008; Weibull and Arntsen, 2013a; Yang et al., 2013; Shan and Wang, 2013; Biondi and Almomin, 2014; Díaz and Sava, 2015; Yang and Sava, 2015), by measuring moveout in the angle or surface offset gathers (Liu and Han, 2010; Shen et al., 2015; Fleury et al., 2014) or, alternatively, by measuring the consistency between individual shot record images (Perrone et al., 2015).

Another type of inversion looks for similarities between simulated and recorded data directly at the receiver locations. This problem, usually referred to as Full Waveform Inversion (FWI), is generally formulated by improving the consistency between modeled and observed data. Originally, Tarantola (1984) introduced the data difference as a similarity estimate in the time domain. Alternatively, the problem can be solved in the frequency domain (Pratt, 1999). Contrary to the image-domain formulation, data-domain inversion is highly non-linear, i.e., the objective function has many local minima. To overcome this non-linearity, a multi-scale separation approach is needed (Bunks et al., 1995). Within each scale (frequency or frequency band), the problem can be more linear if the initial model is closer to the one corresponding to the global minimum. To tackle the intrinsic non-linearity, a great deal of research has been proposed to adjust the objective function (Shin and Ha, 2008; Shin and Ho Cha, 2009; Sirgue and Pratt, 2004; Engquist and Froese, 2013; Luo and Schuster, 1991; Warner and Guasch, 2014; Alkhalifah and Choi, 2014). In theory, the resolution of FWI can approach that of a seismic image. Hence, as a byproduct of the inversion, the resolution of

the velocity model can be used for reservoir characterization, 4D analysis, and improving the seismic image itself (Sirgue et al., 2009). The great benefits of FWI make solving the issue of non-linearity an area of very active research.

The methods described here rely on Born modeling approximation, which means that they can accurately image/describe seismic events that scatter only once inside the earth. However, the seismic data can contain strong multi-scattering that appears in seismic images as coherent noise, and is characterized by defocusing (or moveout) in the common-image gathers (CIGs). To overcome this problem, the multiples (either internal or surface-related) are removed from the recorded data before imaging (or before tomography) (Verschuur et al., 1992; Weglein et al., 1997; Guitton, 2005; Lin and Herrmann, 2013) or from the migrated CIGs (Sava and Guitton, 2005; Wang et al., 2010; Weibull and Arntsen, 2013b). If the multiples are not correctly attenuated, they can contaminate the seismic image and can be misleading during tomographic inversion.

In Born-based imaging (Figure 1.1(a)), one assumes that the image at point $\mathbf{x}$ is constructed from the wavepath that originates at the source location $\mathbf{x}_s$, reflects at image point $\mathbf{x}$, and is recorded at the surface at point $\mathbf{x}_r$. The paths that connect the source and the receiver are only due to the reflection at image point $\mathbf{x}$. Alternatively, within the multiple scattering framework, all possible paths that connect the source with the receiver at the image point $\mathbf{x}$ are taken into account (Figure 1.1(c)). Hence, the image point is illuminated from many other angles. Illumination from multiple directions allows to better represent a scatterer at point $\mathbf{x}$ through the imaging condition.

For tomography, the Born approximation implies that the sensitivity kernel updates the velocity model along the direct paths between the source coordinate $\mathbf{x}_s$ to the image point $\mathbf{x}$, and from the receiver $\mathbf{x}_r$ to the image point $\mathbf{x}$. Hence, the information extracted at the image point $\mathbf{x}$ contains limited sensitivity to the background model. In contrast, within the multi-scattering framework, the sensitivity of the image point $\mathbf{x}$ to the background model is along all possible paths that connect the source point $\mathbf{x}_s$ with receiver point $\mathbf{x}_r$ at image

point $\mathbf{x}$, which implies higher sensitivity to the background model parameters.

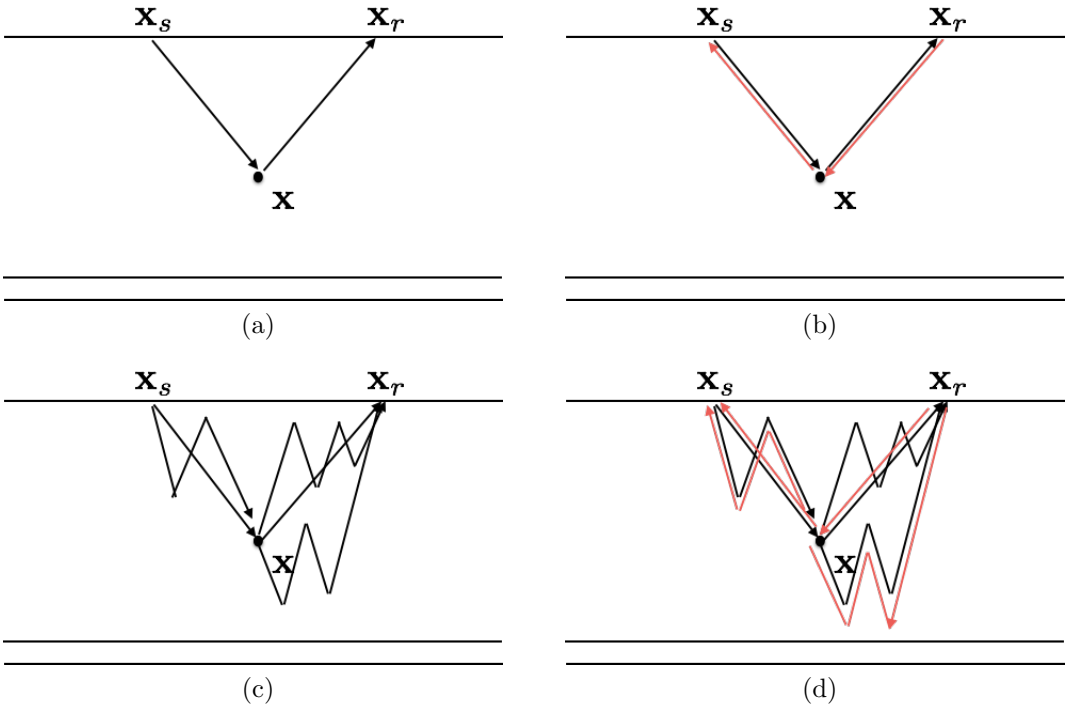


Figure 1.1: Depiction of Born and multiple scattering imaging and tomography: (a) shows the path used during Born imaging to construct the seismic image at point $\mathbf{x}$, (b) the sensitivity kernel (path) updated by Born-based tomographic updates, (c) the possible information that can be used in addition to the Born path during multiple scattered imaging, and (d) the sensitivity kernel from the multiple scattered waves.

1.0.1 Thesis outline

This thesis explores different ways to measure the velocity errors in the subsurface and offers insights into the effect of using different two-way operators (finite difference modeling and Marchenko reconstruction) in extended images.

In **Chapter 2**, I explore two different but complementary wavefield tomography methods for migration velocity analysis. The main idea is to build a smooth model with image-domain tomography which is later refined by data-domain tomography. I apply the cascaded strategy of image- and data-domain tomography to a 2D marine dataset. The image-domain tomography utilizes illumination constraints which are incorporated through the penalty

function applied to common image point (CIP) gathers at each iteration. Once the image-domain model is built, the inversion is continued using data-domain tomography. After the application of both methodologies, the migrated image and the kinematic match between observed and modeled data improves significantly.

In **Chapter 3**, I propose a data domain tomography misfit criterion based on local correlations. The local correlation share many similarities with extended images used in image-domain tomography. The main feature is that the misfit measurement is taken around seismic events in a similar fashion as the image-domain observations are taken at the reflector positions. Correlation-based inversions are less sensitive to local minima than difference-based inversions like full waveform inversion (FWI). Global correlations-based inversions, however, are often contaminated with coherent noise that is the result of comparing the misfit of events that should not be compared; in addition, the global correlations give just a general idea of the kinematic errors of the model because of they measure the delays along the entire time axis. Alternatively, local correlations are able to extract the local kinematic errors between modeled and observed data. Local correlations avoid the comparisson of events that are far from each other since the correlation is performed locally as a function of time. For extracting the misfit, I use a penalty operator that is consistent with the bandwidth of the data similarly to the procedure I use in **Chapter 2** for image-domain tomography.

In **Chapter 4**, I use two-way wavefields that explain all the scattered events present in the surface data as input to an inversion problem in which aims to derive the model that best explains the scattering present in the wavefield. To solve the proble, I insert the wavefields into the homogeneous wave equation and solve for the desired model parameter. The approach is an alternative to conventional imaging methodologies where the imaging objective is to obtain the contrasts of the subsurface properties. Here, I image the properties themselves. In conventional reservoir characterization the properties are obtained from seismic reflectivity in a process with many approximations. The strategy devised in this chapter avoids these post-imaging procedures.

In **Chapter 5**, I explore the problem of measuring focusing with wavefields that contain multiple scattered waves. Commonly, the multiples are regarded to as noise during the imaging and tomography. However, multiples contain redundant information that, if used correctly, can improve the illumination of the subsurface. To understand the differences, I compare common image gathers obtained with conventional wavefields that obey the single scattering assumption with gathers obtained by wavefields that correctly handle multiple scattered waves. Using all primaries and multiple-scattered waves improves the subsurface illumination and sensitivity of extended images to errors in the background model. I show that the CIGs obtained with wavefields that account for multiple scattered waves are sensitive to errors in ways similar to gathers computed with reverse-time wavefields. The main difference is that the wavefields that correctly model multi scattered waves image the multiples at the same location as the primaries, whereas the multiples present in the images computed with reverse-time are imaged at different locations than the primaries.

CHAPTER 2

CASCADED WAVEFIELD TOMOGRAPHY & INVERSION USING EXTENDED COMMON IMAGE POINT GATHERS: A CASE STUDY[5]

Esteban Díaz and Paul Sava

Center for Wave Phenomena, Colorado School of Mines

Image domain wavefield tomography exploits focusing characteristics of extended images for updating the velocity field. In order to make good use of this information, one must understand how such images behave if the migration velocity is accurate. This is not trivial since focusing depends on the acquisition setup, the data bandwidth, as well as on illumination variation caused by the geology separating the acquisition array from the imaged structure. We address this problem using a combination of migration/demigration to construct penalty functions that characterize focusing in extended images by incorporating acquisition parameters and data bandwidth. Moreover, instead of sampling the extended images at preset distances along the surface, we sample the image by constructing common image-point gathers, which are more economical from a computational point of view. Coupled with image residuals exploiting illumination-based penalty functions, we construct robust wavefield tomographic updates in areas of poor or uneven illumination. Models obtained with this type of methodology are a good starting point for more sensitive, but less robust waveform inversion methods. Starting with the velocity model obtained by illumination-based tomography, we perform an inversion using waveform tomography in the data-domain. We show an application of the method to synthetic examples and to a marine 2D dataset.

2.1 Introduction

An accurate velocity model is the main requirement for successful imaging. In order to be consistent with the typical band-limited seismic data, one ought to use wavefield-

[5]CHAPTER SUBMITTED TO GEOPHYSICS

based tomographic velocity estimation. The wavefield-based approach avoids shortcomings inherent in ray-based methods, such as limited model sensibility (a ray travels trough a infinitesimally narrow path inside the model) and instability around sharp boundaries in the velocity model.

Velocity analysis methods based on wavefield extrapolation are commonly referred to as Wavefield Tomography (WT) (Tarantola, 1984; Woodward, 1992; Pratt, 1999; Sava and Biondi, 2004a,b; Shen and Symes, 2008; Biondi and Symes, 2004; Symes, 2008); such tomographic approaches can be formulated either in the image domain, where one tries to improve image quality, or in the data domain, where one seeks consistency between modeled and observed data.

Image-domain wavefield tomography can be formulated in many ways. A common approach aims to improve the flatness of angle gathers, or equivalently one can improve the focusing of space-lag gathers (Sava and Biondi, 2004a; Shen and Calandra, 2005). Space-lag gathers (Rickett and Sava, 2002; Sava and Fomel, 2006), measure the spatial similarity between source and receiver wavefields. Hence, during tomography one seeks to increase the similarity of the spatial correlation for a collection of seismic experiments (Shen and Symes, 2008; Weibull and Arntsen, 2013a; Yang et al., 2013; Shan and Wang, 2013; Biondi and Almomin, 2014). The concept of focusing in extended images can be applied to space-lag gathers, to time-lag gathers, or to a combination of both space and time-lag gathers (Yang and Sava, 2015). In order to define the tomographic objective, one must first define an operator whose purpose is to highlight the defocused events in the gather that one needs to focus trough iterative tomography. In general, this operator is defined by a weight proportional to the lag axis, as introduced by Shen and Symes (2008). In contrast, Yang et al. (2013) advocate for using a more complicated penalty functions to account for illumination of the seismic data, which can improve the convergence rate of the problem. By incorporating illumination information into the penalty function the tomography goal is more realistic since it has the objective of improving the focus given the bandwidth and acquisition parametriza-

tion of the data. In contrast, conventional penalties seek to obtain extended images whose focus approaches a delta function which is an unrealistic goal given the limitations of seismic data.

Data-domain wavefield tomography is generally formulated by improving the consistency between modeled and observed data. Originally, Tarantola (1984) introduced the data difference as a similarity estimate in the time domain. Alternatively, the problem can be solved in the frequency domain (Pratt, 1999). Contrary to the image-domain formulation, data-domain wavefield tomography is highly non-linear , i.e., the objective function has many local minima. To overcome the non-linearity, a multi-scale separation approach is needed (Bunks et al., 1995). Within each scale (frequency or frequency band), the problem can be more linear if the initial model is closer to the one corresponding to the global minimum. Another loop of multi-scale can be added by introducing time damping, a method commonly referred to as Laplace-Fourier waveform inversion (Sirgue and Pratt, 2004; Shin and Ho Cha, 2009). The purpose of the time damping is to fit earlier arrivals first, and then to fit later arrivals progressively.

Data-domain and image-domain tomographic methods share similarities in their processes: both use the same extrapolation engine (the two-way wave equation), and build the gradient of the corresponding objective function through the adjoint state method (ASM) framework (Tarantola, 1984; Plessix, 2006; Symes, 2008). In this article, we interleave a time-domain application for image-domain tomography using common-image-point (CIP) gathers under illumination constraints with a frequency-domain application of data-domain tomography. We further analyze the tomographic term in the image-domain inversion, using different adjustments to the objective function like illumination and maximization to include the zero-lag gather contribution directly into the inversion (Zhang and Shan, 2013). We complement the low wavefield tomography resolution with a a high resolution data-domain inversion which has the potential to add resolution to the reconstructed model. We illustrate the tomographic implementation with a synthetic dataset, and with a field data

marine example.

2.2 Image domain tomography

Extended-image gathers (Rickett and Sava, 2002; Sava and Fomel, 2006; Sava and Vasconcelos, 2011b) highlight the spatial and temporal consistency between wavefields by exploring the focusing information in the image domain characterized with common image gathers. The moveout in the gather is sensitive to velocity perturbations, and hence can be optimized. A general extended image can be defined as follows:

$$R(\mathbf{x}, \boldsymbol{\lambda}, \tau) = \sum_e \sum_t u_s(e, \mathbf{x} - \boldsymbol{\lambda}, t - \tau) u_r(e, \mathbf{x} + \boldsymbol{\lambda}, t + \tau), \tag{2.1}$$

where $\boldsymbol{\lambda}$ is the space-lag vector, τ the time-lag, $\mathbf{x}$ the image location, e the experiment index, u_s the source wavefield, and u_r the receiver wavefield. Given the increase in dimensions for extensions in space and time, one can take advantage of a sparse sampling of the extended images in a subset of image locations $\mathbf{x}_c$ instead of the full image space $\mathbf{x}$. One efficient way to decide the locations $\mathbf{x}_c$ is by placing the observation points at the reflector locations (Cullison and Sava, 2011). Note that the process for computing the extended image is linear with respect to one of the wavefields. Hence, one can define the extended image in a matrix-vector form as

$$\mathbf{R} = \mathbf{I}_s \mathbf{u}_r = \mathbf{I}_r \mathbf{u}_s, \tag{2.2}$$

where $\mathbf{R}$ is the vector representation of the extended image, and $\mathbf{I}_s$ and $\mathbf{I}_r$ are the imaging operators for the source and receiver wavefield, respectively.

Even though the commutation of the forward mappings (equation 2.2) produces the same result ($\mathbf{R}$), the adjoint mapping satisfies different equations. The operator $\mathbf{I}_s^\top$ implements

$$\tilde{u_r}(e, \mathbf{x} + \boldsymbol{\lambda}, t + \tau) += \sum_{\tau, \boldsymbol{\lambda}} R(\mathbf{x}, \boldsymbol{\lambda}, \tau) u_s(e, \mathbf{x} - \boldsymbol{\lambda}, t - \tau), \tag{2.3}$$

whereas $\mathbf{I}_r^\top$ satisfies

$$\tilde{u_s}(e, \mathbf{x} - \boldsymbol{\lambda}, t - \tau) += \sum_{\tau, \boldsymbol{\lambda}} R(\mathbf{x}, \boldsymbol{\lambda}, \tau) u_r(e, \mathbf{x} + \boldsymbol{\lambda}, t + \tau). \tag{2.4}$$

The symbol $+=$ implies that the adjoint wavefield is accumulated in the output after summing all possible lags $\boldsymbol{\lambda}, \tau$. The output wavefields $\tilde{u_r}$ and $\tilde{u_s}$ are obtained after the application of the adjoint imaging operators to the input wavefields u_s and u_r, respectively.

The source and receiver wavefields follow the wave equation, which in this case is the acoustic wave equation:

$$m(\mathbf{x})\frac{\partial^2 u(e,\mathbf{x},t)}{\partial t^2} - \nabla^2 u(e,\mathbf{x},t) = f(e,\mathbf{x},t), \tag{2.5}$$

where $m = s^2(\mathbf{x})$ is the squared slowness. The source wavefield $u_s(e,\mathbf{x},t)$ is computed by forward propagation of the source function $f_s(e,\mathbf{x},t)$ for experiment index e, whereas the receiver wavefield $u_r(e,\mathbf{x},t)$ is created by backward (adjoint) propagation of the recorded data $f_r(e,\mathbf{x},t)$ as summarized by the following system of equations:

$$\begin{bmatrix} \mathcal{L}(m,t) & 0 \\ 0 & \mathcal{L}^\top(m,t) \end{bmatrix} \begin{bmatrix} u_s \\ u_r \end{bmatrix} = \begin{bmatrix} f_s \\ f_r \end{bmatrix}, \tag{2.6}$$

where $\mathcal{L}$ and $\mathcal{L}^\top$ are the matrix representation of the forward and adjoint (reverse) propagators, respectively.

A model with good kinematic agreement with the true background velocity exhibits a maximum correlation between source and receiver wavefields around $(\boldsymbol{\lambda}, \tau) = (\mathbf{0}, 0)$. For inaccurate velocity, extended gathers exhibit moveout indicative of model errors (Yang and Sava, 2010). Hence, the energy in the gathers can be used to formulate an inversion process to update the velocity model. A typical objective function (Shen and Symes, 2008; Weibull and Arntsen, 2013a; Yang and Sava, 2015) based on extended images is

$$J_1(m) = ||P(\boldsymbol{\lambda},\tau)R(\mathbf{x},\boldsymbol{\lambda},\tau)||^2, \tag{2.7}$$

where P is a penalty operator whose purpose is to highlight the events in the image that show velocity inaccuracies. The penalty operator is sometimes defined using the lags:

$$P(\boldsymbol{\lambda},\tau) = \sqrt{\boldsymbol{\lambda}^2 + (v\tau)^2}, \tag{2.8}$$

where v is the local velocity (Yang and Sava, 2015). This penalty function is highly idealized since it produces a zero residual if the extended image is a perfect spike (i.e., $R(\mathbf{x},\boldsymbol{\lambda},\tau) =$

$\delta(\mathbf{x}, \boldsymbol{\lambda}, \tau)$). In order to relax such requirements, Yang et al. (2013) propose a penalty function based on illumination, which depends on the acquisition setup, data bandwidth, and the velocity model itself. To compute the action of the illumination (blurring operator), one can cascade a demigration/migration process applied to spikes placed at the extended image locations $\mathbf{x}_c$:

$$\tilde{R}(\mathbf{x}_c, \boldsymbol{\lambda}, \tau) = \mathbf{M}\mathbf{M}^\top \sum_i \delta(\mathbf{x} - \mathbf{x}_c^i, \boldsymbol{\lambda}, \tau) \tag{2.9}$$

with $\mathbf{M}^\top$ being a demigration operator and $\mathbf{M}$ the extended imaging process. The index i corresponds to the CIP location $\mathbf{x}_i$. The synthetic image $\tilde{R}$ contains the action of the source function, the acquisition, and the model at the image locations $\mathbf{x}_c$. Once $\tilde{R}$ is obtained, one can proceed with the penalty

$$\tilde{P}(\mathbf{x}_c, \boldsymbol{\lambda}, \tau) = \frac{1}{Env(\tilde{R}(\mathbf{x}_c, \boldsymbol{\lambda}, \tau))}, \tag{2.10}$$

where Env is an envelope function that captures the energy of the synthetic image $\tilde{R}$. Effectively, $\tilde{R}$ contains the extended point spread functions (PSF) of the model, and is computed similarly to Valenciano et al. (2009) and Fletcher et al. (2012).

In order to include the zero-lag information, one can add an extra term in the objective function, similarly to Shen and Symes (2008), Weibull and Arntsen (2013a). Alternatively, we include the zero-lag term directly into the inversion by changing the goal of the objective function to a minimization problem (Zhang and Shan, 2013):

$$J_2(m) = -\left|\left|H(\boldsymbol{\lambda}, \tau)R(\mathbf{x}, \boldsymbol{\lambda}, \tau)\right|\right|^2. \tag{2.11}$$

The operator H has the opposite purpose from operator P; now the main task of H is to "highlight" the energy around zero lag. By using J_2 instead of J_1, we fuse all lags in a seamless manner into the inversion. Another advantage of J_2 is that it avoids the extra image power term in the inversion (to include zero lag). It also avoids assigning a weighting scalar value to balance the DSO and image power terms (it is scaled altogether by the operator H).

2.2.1 Tomography setup

We are interested in the gradient of equation 2.11 to update the background velocity model. The inverse problem is set up with constraints on the wave equation:

$$\begin{aligned} \underset{m}{\text{minimize}} \quad & J_2(m) \\ \text{subject to} \quad & \mathcal{L}u_s = f_s, \quad \mathcal{L}^\top u_r = f_r. \end{aligned} \tag{2.12}$$

In order to compute the gradient we follow the methodology described by Plessix (2006) for PDE constrained objective functions. First, we define the augmented functional

$$A(m, u_s, u_r, a_s, a_r) = -J_2(m) + < a_s, \mathcal{L}u_s - f_s > + < a_r, \mathcal{L}^\top u_r - f_r >, \tag{2.13}$$

where $< ., . >$ defines a dot product between vectors in the $(\mathbf{x}, t)$ space, a_s is the adjoint source wavefield, and a_r is the adjoint receiver wavefield. The augmented functional depends on the model m, the forward state variables u_s and u_r, and the adjoint state variables a_s and a_r. From the first optimality conditions, $\frac{\partial A}{\partial a_s} = 0$ and $\frac{\partial A}{\partial a_r} = 0$, we obtain the system in equation 2.6.

In order to find the adjoint source wavefield a_s, we rewrite equation 2.13 as

$$A(m, u_s, u_r, a_s, a_r) = - < HI_r u_s, HI_r u_s > + < a_s, \mathcal{L}u_s - f_s > + < a_r, \mathcal{L}^\top u_r - f_r > . \tag{2.14}$$

By setting $\frac{\partial A}{\partial u_s} = 0$, we obtain the adjoint state variable a_s:

$$\mathcal{L}^\top a_s = I_r^\top H^\top H I_r u_s = I_r^\top H^\top H R. \tag{2.15}$$

The operator H is diagonal and real, which means $H^\top = H$. This equation tells us to use the receiver wavefield u_r as an imaging operator together with the image residual $H^\top H R$. The wavefield a_s travels from the image towards the source.

To find the adjoint receiver wavefield a_r, we redefine equation 2.13 to the equivalent form

$$A(m, u_s, u_r, a_s, a_r) = - < HI_s u_r, HI_s u_r > + < a_s, \mathcal{L}u_s - f_s > + < a_r, \mathcal{L}^\top u_r - f_r > . \tag{2.16}$$

By setting $\frac{\partial A}{\partial u_r} = 0$, we obtain the adjoint variable a_r:

$$\mathcal{L}a_r = I_s^\top H^\top HR. \tag{2.17}$$

From the last optimality condition, we obtain the actual gradient

$$\frac{\partial J_2}{\partial m} = -\sum_{t,e} \frac{\partial^2 u_s}{\partial t^2} a_s + \frac{\partial^2 u_r}{\partial t^2} a_r. \tag{2.18}$$

2.3 Data domain Wavefield tomography

The construction of the tomography problem in the data domain amounts to measuring the error (or residual) at the receiver locations. For data domain wavefield tomography, one normally uses the data difference for the residual:

$$J_3 = \frac{1}{2} \left|\left| u_s(\mathbf{x}_r, \Omega) - f_r(\mathbf{x}_r, \Omega) \right|\right|^2 = \frac{1}{2} \left|\left| \Delta d \right|\right|^2, \tag{2.19}$$

where $\mathbf{x}_r$ are the receiver locations and Ω is the complex valued frequency whose purpose we will explain later. Note that $f_r(\mathbf{x}_r, \Omega) = u_r(\mathbf{x}_r, \Omega)$. Since for building the residual, we only need to forward propagate the source function

$$\mathcal{L}(m, \Omega) u_s = f_s(\mathbf{x}_s, \Omega), \tag{2.20}$$

we have to compute one adjoint wavefield $a_s(\mathbf{x}, \Omega)$. For data-domain wavefield tomography, computing a_s involves backpropagating the data residual:

$$\mathcal{L}^\top(m, \Omega) a_s = (\Delta d)^*. \tag{2.21}$$

Here $(\Delta d)^*$ is the complex conjugate residual and $\mathcal{L}(m, \Omega)$ is the acoustic wave equation in the frequency domain, defined as follows:

$$\mathcal{L}(m) = -\nabla \cdot \frac{1}{\rho(\mathbf{x})} \nabla - \frac{m(\mathbf{x})}{\rho(\mathbf{x})} \Omega^2, \tag{2.22}$$

where $\rho(\mathbf{x})$ is the density of the medium. In this article, we do not invert for $\rho(\mathbf{x})$; instead we parametrize it as a function of the velocity following Gardner et al. (1974).

Once we obtain $u_s(\mathbf{x},\Omega)$ and $a_s(\mathbf{x},\Omega)$, we can proceed to compute the gradient:

$$\nabla J_3(\mathbf{x}) = \mathbb{R}\left\{\sum_{t,e}\Omega^2 u_s(e,\mathbf{x},\Omega)a_s^*(e,\mathbf{x},\Omega)\right\}, \tag{2.23}$$

where * denotes complex conjugate and $\mathbb{R}\{\}$ denotes the real part.

The data domain wavefield tomography objective function, equation 2.19, is highly non-linear if we operate with signals in the normal frequency band. Hence, in order to increase the chances of convergence to the global minimum, it is customary to implement data-domain wavefield tomography in a multi-scale fashion. Bunks et al. (1995) propose to first invert lower frequencies and then move gradually to higher frequencies. The idea is that, within each scale, the problem looks more linear than when one inverts all the bandwidth at once.

An additional outer loop in the inversion is the time damping, which leads to the so-called Laplace-Fourier domain FWI (Sirgue and Pratt, 2004; Shin and Ho Cha, 2009). The purpose of this outer loop is to first fit earlier arrivals, and then fit later arrivals. By first fitting early arrivals (with shorter travel-time) one reduces the risk of large phase differences between observed and modeled data which can cause cycle-skipping. Once the travel-time differences are solved for early arrivals, we can progressively increase ς. Introducing time damping requires the following transformation: $\Omega = \omega + i/\varsigma$, with ς being the time damping (Kamei et al., 2013). Thus, this transformation turns the real-valued angular frequency ω into a complex-valued angular frequency Ω. In order for the observed data be consistent with the damped modeled data, one must also scale it as $f_r(\mathbf{x}_r,t) = d_{obs}(\mathbf{x}_r,t)e^{-t/\varsigma}$ before the transformation to the frequency domain.

The low frequencies of the data are sensitive to the long wavelength (smooth) components of the earth model. However, if the data do not have such frequencies, data-domain wavefield tomography is unable to update such components. In contrast, focusing in extended images is mostly sensitive to the smooth components of the model. By implementing a joint workflow using image-domain wavefield tomography for updating the smooth components of the model and later using data-domain wavefield tomography for the high resolution features of the

model, we can obtain a more complete spectrum in the model. The first pass of using image-domain wavefield tomography has the ability to stabilize the cycle-skipping problems in data-domain wavefield tomography.

2.4 Synthetic data example

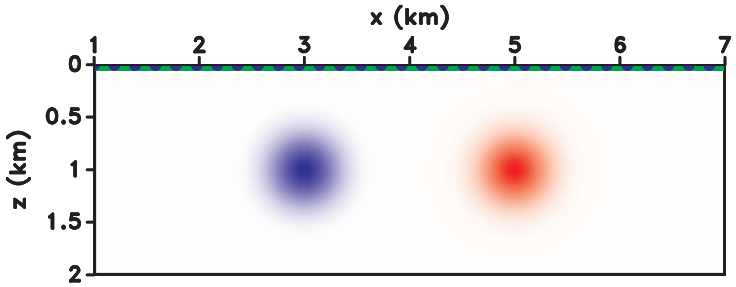
In this section, we use a simple synthetic example to demonstrate the image-domain tomography process. Figure 2.1(a) shows a background velocity model of 3km/s with two Gaussian anomalies of -0.5km/s and $+0.5$km/s from left to right, respectively. The four reflectors, shown in the image computed with the true velocity model (Figure 2.1(b)), are produced by contrasts in the density model. The starting model for tomography is set constant at 3km/s. The purpose of the tomographic step is to recover the effect of the Gaussian anomalies in the model. Figure 2.1(c) depicts the seismic image computed from the starting model.

We perform two inversions using CIP gathers: (1) with the conventional DSO penalty function (equation 2.8) and (2) with the illumination-based highlighting operator H (from equation 2.10). Figure 2.2(a) shows the image corresponding to the inverted model in Figure 2.2(b) using the conventional DSO penalty function. Figure 2.2(c) depicts the image corresponding to the inverted model in Figure 2.2(d) using the illumination-based maximization approach. Both inverted models introduce velocity anomalies with the expected sign. However, the model retrieved with the illumination-based framework better reconstructs the anomalies compared with the model from the DSO penalty functions. The RTM image corresponding to the illumination-based model (Figure 2.2(c)) is flatter than the corresponding DSO-based penalty model shown in Figure 2.2(a).

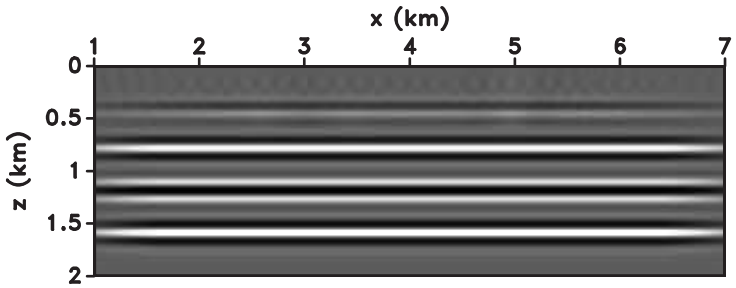
In order to understand the fundamental difference between the two penalty methods, one can take a look at an ideal CIP image obtained from the correct model shown in Figure 2.3(a) and compare it to a CIP image to a -10% velocity anomaly (Figure 2.3(b)). The conventional DSO-based penalty function depicted in Figure 2.3(c) is ideal for an unrealistic CIP image described by a delta function:

$$R(\mathbf{x}, \boldsymbol{\lambda}, \tau) = \delta(\boldsymbol{\lambda})\delta(\tau).$$

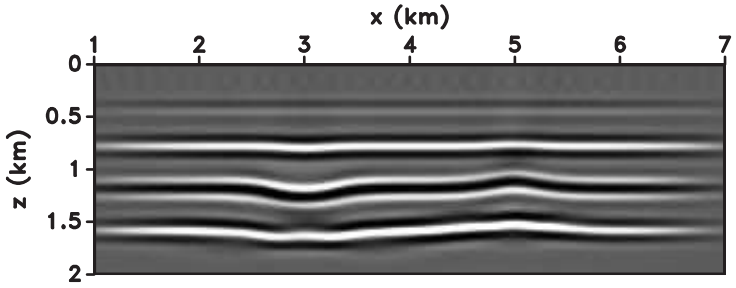
In contrast, Figure 2.3(d) is computed through a migration/demigration process that mimics the ideal focusing for the given acquisition parameters. One can observe the consistency between the illumination-based penalty function, Figure 2.3(d), and the ideal CIP image, Figure 2.3(a).



(a)

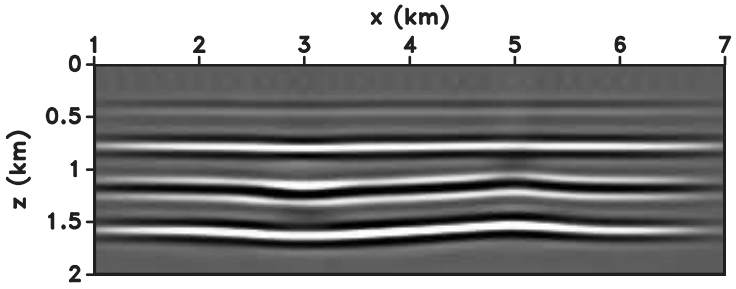


(b)

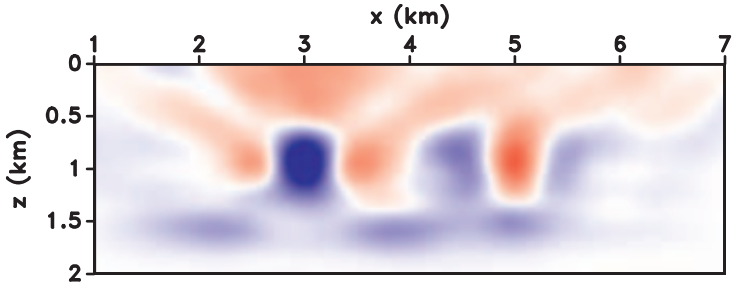


(c)

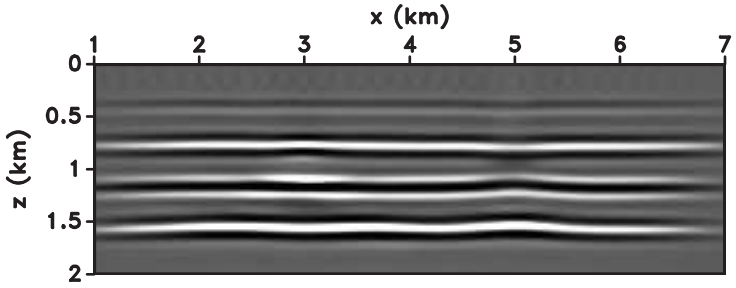
Figure 2.1: Synthetic examples: (a) true velocity model, (b) corresponding image, and (c) image after starting velocity model (constant background).



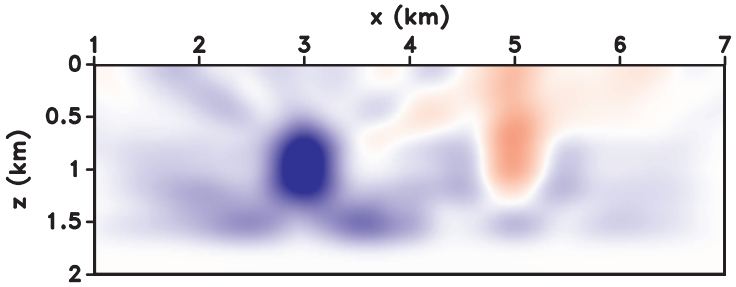
(a)



(b)



(c)



(d)

Figure 2.2: Synthetic example: (a) updated image and (b) model after DSO-based tomography, and (c) updated image and (d) model after illumination-based tomography. Note how the structure is better recovered with the illumination-based approach.

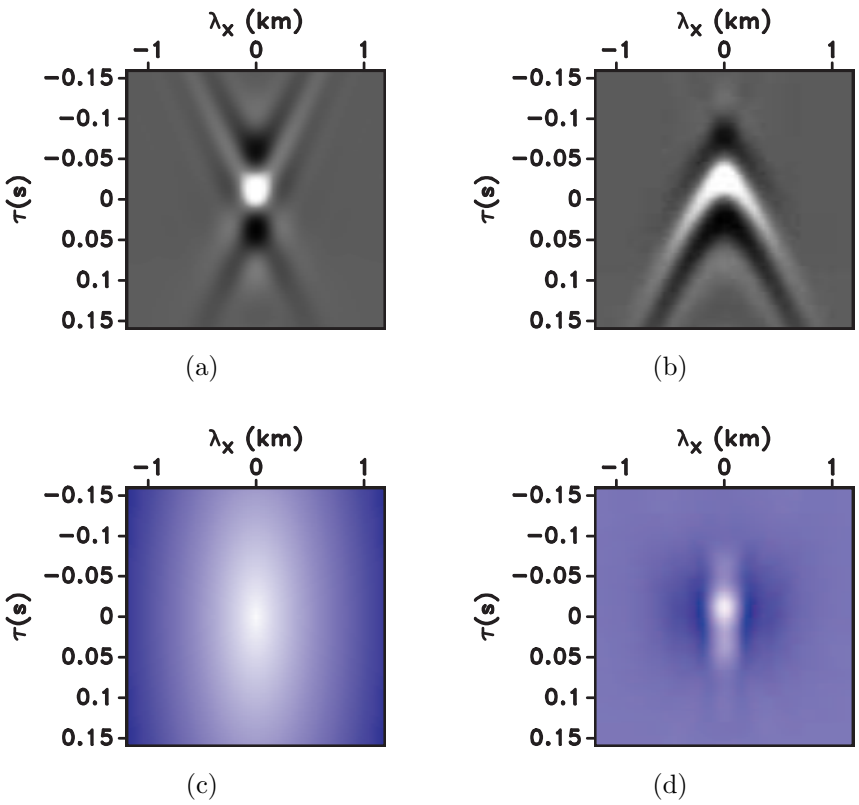


Figure 2.3: (a) a example of CIP gather from the correct model, (b) CIP gather with 10% velocity error, (c) conventional penalty function, and (d) illumination-based penalty. Note the consistency between the penalty functions in (d) and the correct gather in (a).

2.5 Field data example

In this section we apply the cascaded workflow of image-domain wavefield tomography followed by data-domain wavefield tomography. We use image-domain wavefield tomography to correct for most kinematics errors, and then data-domain wavefield tomography to refine the model and add details. We obtain 3 models: (a) the initial model, (b) the model obtained by image domain wavefield tomography, and (c) the data-domain wavefield tomography model.

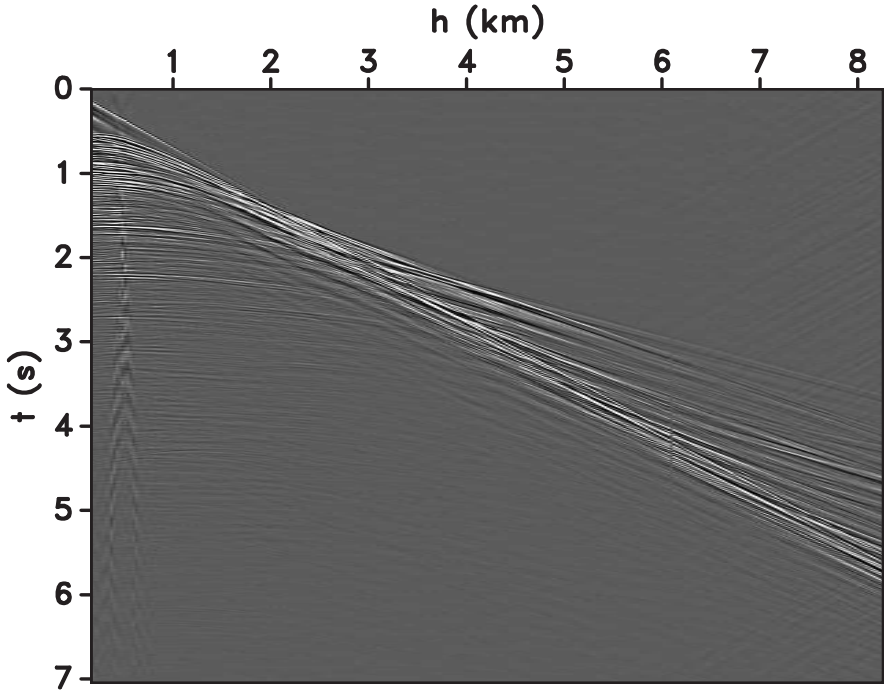
The dataset is a marine 2D line acquired with a variable depth cable. The towed streamer contains increasing depths as a function of offset, which enhances the frequency content of the data by producing a mixed-notch response. Hence, the increased cable depths improve the low-frequency content at intermediate and far offsets, which can be very helpful for data-domain wavefield tomography. The cable contains offsets ranging from 0.169 to 8.256 km. Figure 2.4(a) shows a shot gather from $x = 14$ km, and Figure 2.4(b) depicts the average amplitude spectrum for the same gather.

We build the initial model, Figure 2.7(a), by performing time-domain NMO analysis followed by smoothing, RMS (stacking) conversion to interval velocity (Dix, 1955), and time-to-depth conversion. Figure 2.5(a) shows the RTM image produced by the model in Figure 2.7(a), one can see poor focusing through the section and a relatively flat geology below $z = 1.5$ km. Figure 2.6(a) shows angle gathers extracted at sparse locations in the model. Note that we do not use the angle gathers for inversion; instead, we use the gathers as an independent quality control tool. The transformation from space-lag gathers $R(\mathbf{x}, \boldsymbol{\lambda})$ to angle domain $R(\mathbf{x}, \theta)$ follows the method of Sava and Fomel (2003). The angles vary from 0 to 45° for all the gathers shown in this article. The moveout in the gathers confirms that the velocity is too fast. Some of the events in the gathers, however, correspond to migrated surface-related multiples, and their moveout is not indicative of velocity error.

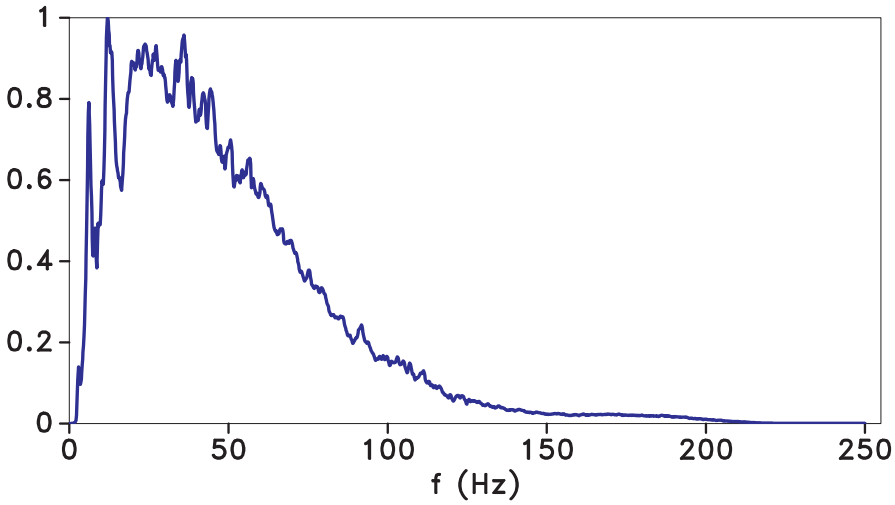
We generate the model in Figure 2.7(b) using the CIP tomography approach with illumination constraints. The CIP gathers contain a maximum space and time lag of $\boldsymbol{\lambda} = 1.5$km

and $\tau = 1$s, respectively. We fix the number of CIP observation points to 10000 per iteration, and perform 11 iterations. Figure 2.5(d) shows a subset of the selected CIP locations for the initial image. The automatic picker (Cullison and Sava, 2011) works by placing cip locations at areas of the image with high coherency, planarity, and amplitude strength. One can also restrict the locations is such a way that the picks are not clustered together. The idea of this tomographic step is to correct for the bulk of the kinematic errors in the model. Figure 2.7(b) shows that in the updated model, in general the model slows down, especially in the deep part of the section. Figure 2.5(b) shows the corresponding RTM image, where the focusing improves significantly throughout the section. This observation is confirmed in Figure 2.6(b), where the gathers are flatter along the section. In order to construct the illumination-based highlightor operators, we perform a demigration/migration step at each iteration. Assuming that the illumination does not vary significantly within iterations, one could update the operators every few iterations, which would help to diminish the overall cost of the tomography.

Finally, we update the image-domain wavefield tomographic model with the data-domain wavefield tomographic approach. Figure 2.7(c) depicts the updated model, which changes considerably in the interval $x = 6$ km to $x = 12$ km. Figures 2.5(c) and 2.6(c) are the corresponding RTM image and angle gathers, respectively. Note how further details are added into the velocity model, which changes the structure of the image and flattens the reflector at $z = 2.2$km. For data-domain wavefield tomography, we use 7 frequency blocks with 5 frequencies each. The center frequency for each block ranges from $f = 2.6$ Hz to $f = 8.9$ Hz. For the time damping constant, we use $\varsigma = 1.6s$. The first step in data-domain wavefield tomography involves estimating the source function $f_s(\Omega)$; later we compare the inverted source functions for each model. Figures 2.8(a)-2.8(b) show the estimated source signatures as a function of shot position before and after data-domain tomography, respectively. One can observe how the consistency of the source signature considerably improves after the model update.

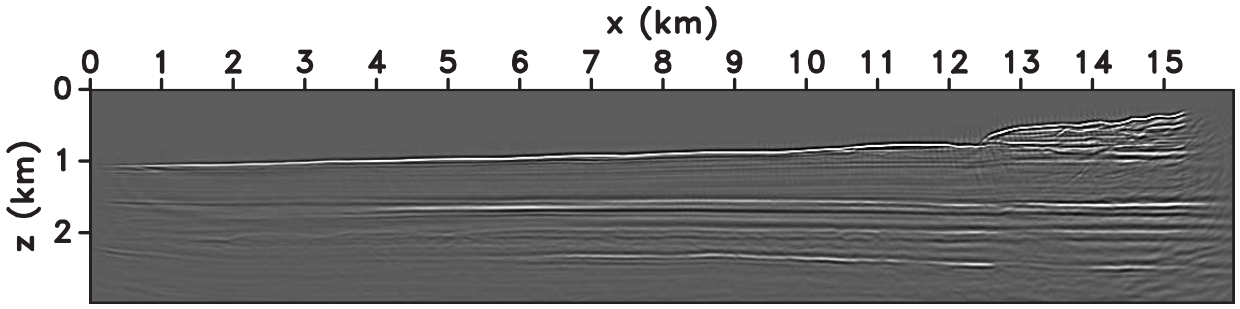


(a)

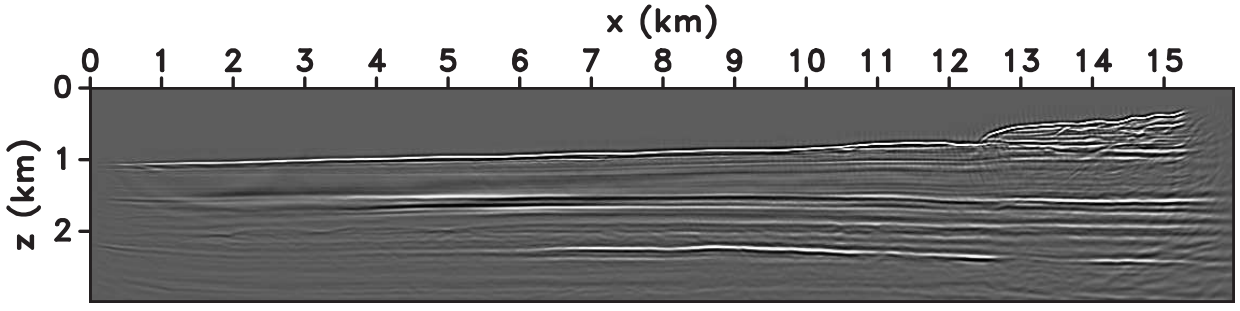


(b)

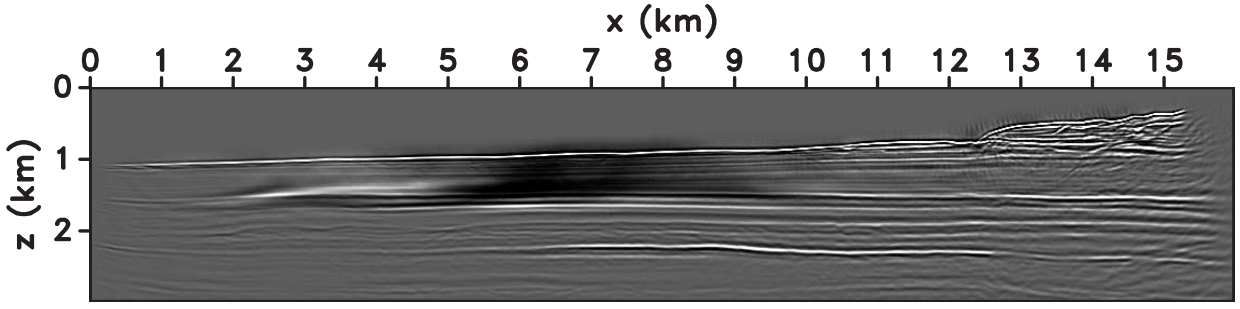
Figure 2.4: (a) Shot gather at $x = 14$ km and (b) its average amplitude spectrum. Note the low frequency response of the data with a good signal to noise ratio starting at around $f = 2.5$Hz.



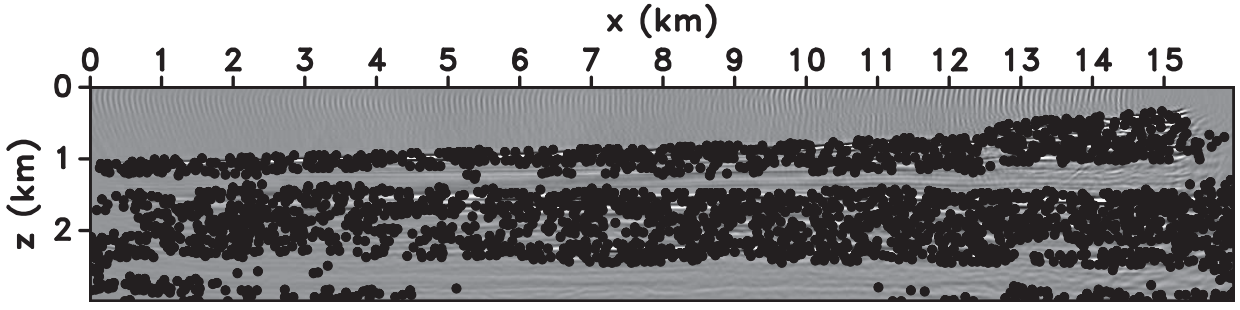
(a)



(b)

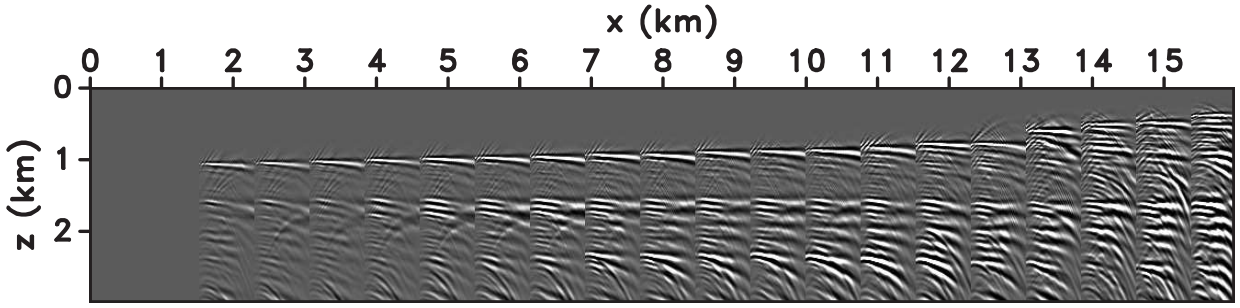


(c)

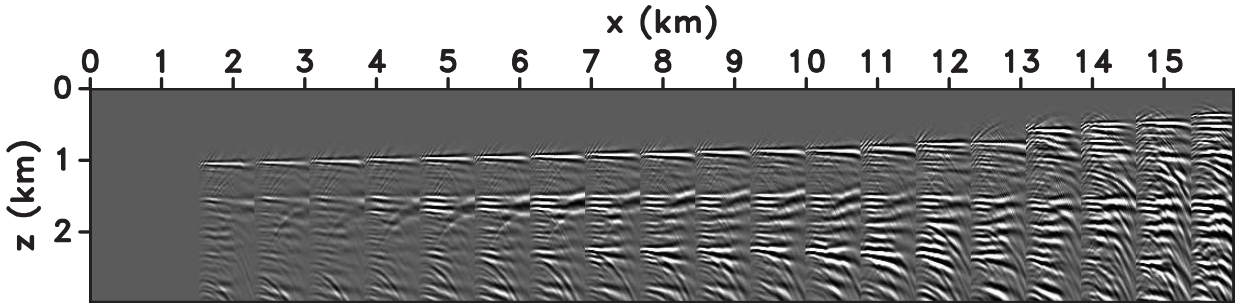


(d)

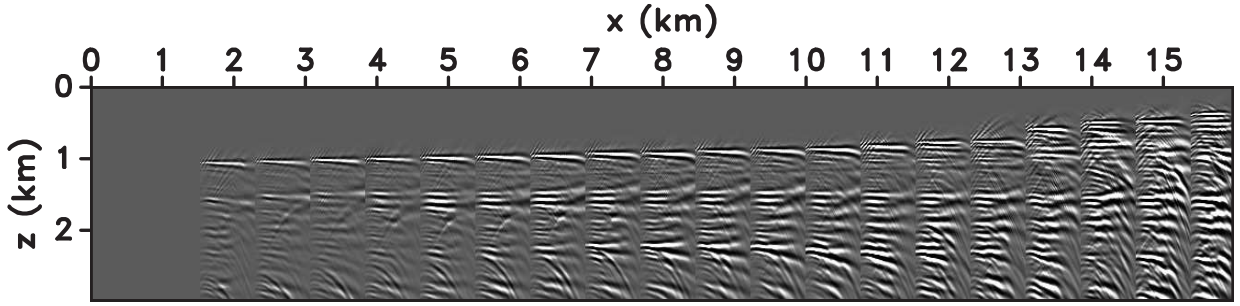
Figure 2.5: RTM images from (a) the initial model, (b) illumination-based tomographic model, and (c) data-domain inversion model. Note how the focusing (image strength) increases from (a) to (b). Also, one can see the shallow part of the image improved due to the data-domain tomography around $z = 1.1$km. Image (d) shows an example of the CIP locations selected during tomography.



(a)

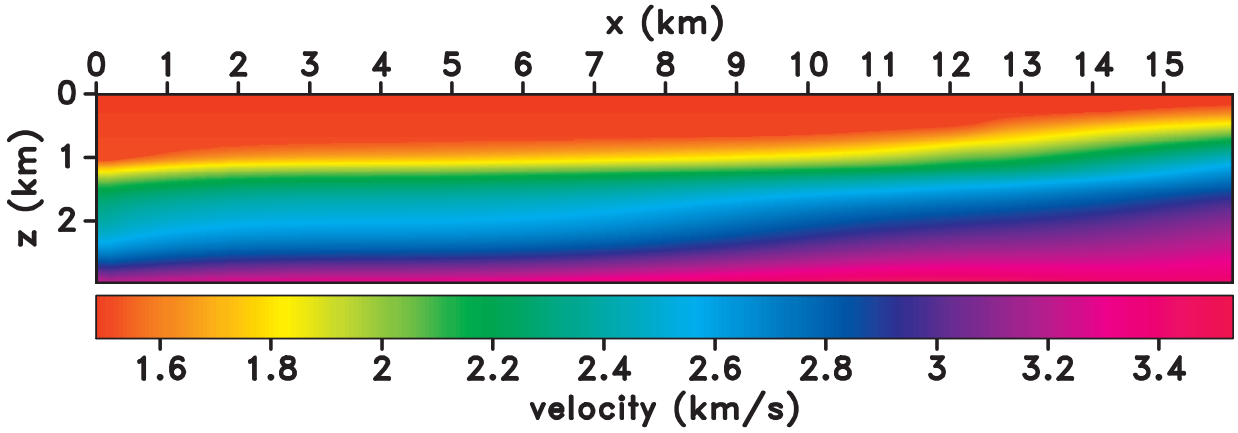


(b)

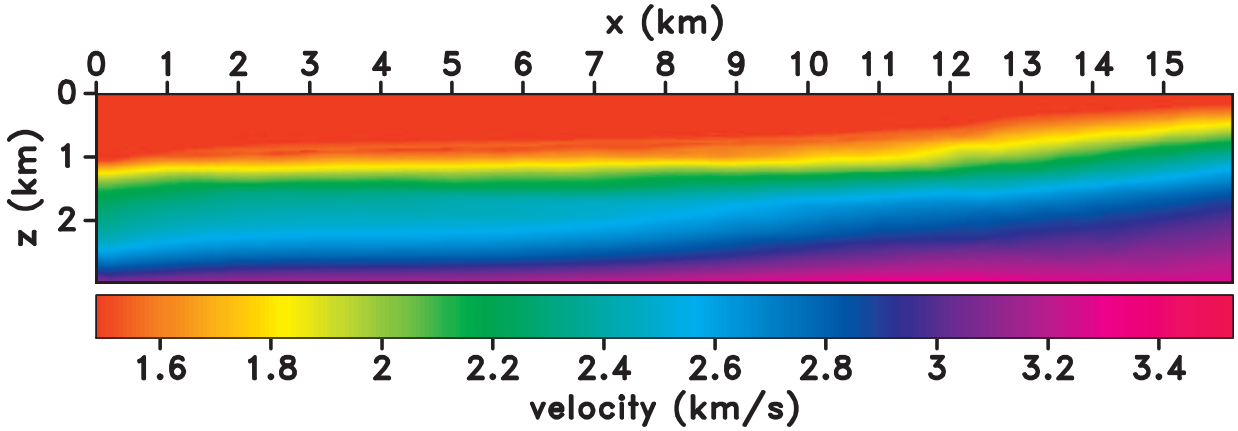


(c)

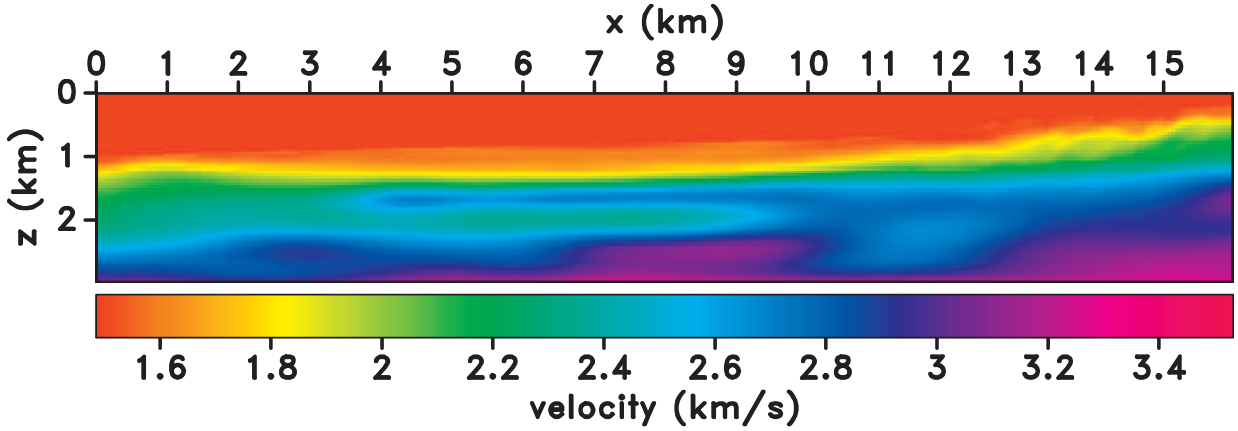
Figure 2.6: Angle gathers retrieved at sparse surface locations from (a) the initial model, (b) illumination-based tomographic model, and (c) data-domain inversion model. Observe how the flatness of the gather increases from (a) to (b). The improvement in (c) is especially evident at the shallow part $z < 1.2$km between $x = 10$km and $x = 15$km.



(a)

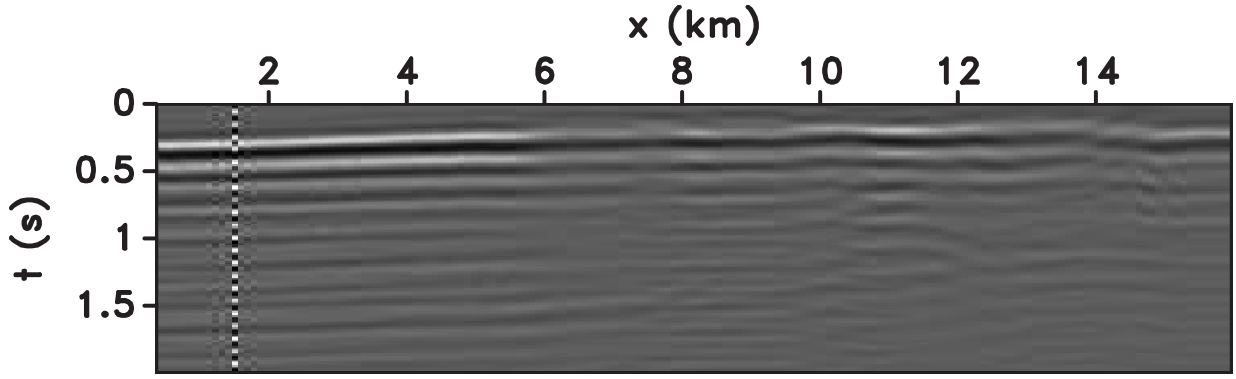


(b)

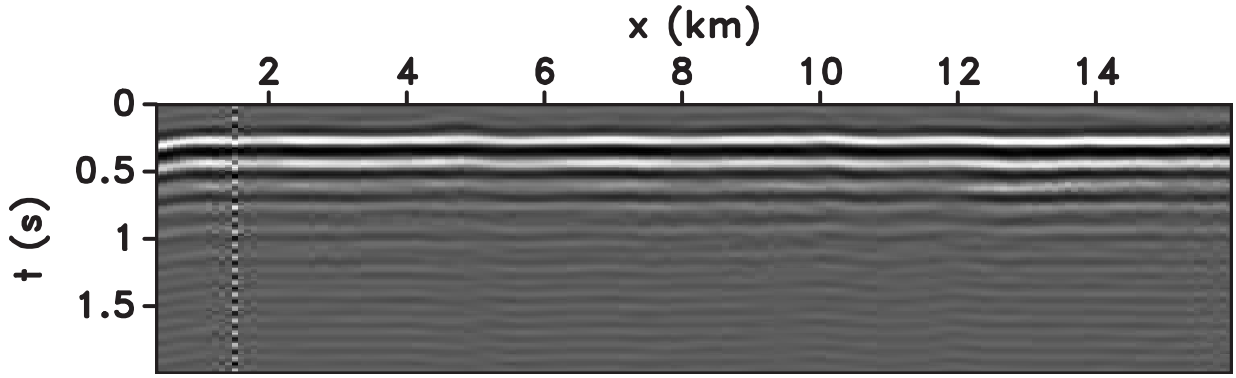


(c)

Figure 2.7: Velocity profiles from (a) the initial model, (b) illumination-based tomographic model, and (c) data-domain inversion model. Note how the velocity decreases from (a) to (b) and details are added in the model in (c).



(a)



(b)

Figure 2.8: Estimated source signatures as a function of shot coordinate: (a) before data-domain tomography and (b) after data-domain tomography. After data-domain tomography the source signatures the lateral consistency improves. For inversion, we use the average of the individual signatures.

2.6 Conclusions

The combination between image-domain and data-domain wavefield tomography seeks to exploit the features of each method. The image-domain wavefield tomography methods are sensitive to the smooth components of the model due to the definition of the inverse problem. Once we obtain a smooth model that improves focusing in the extended images, we can proceed to further refine the model using data-domain wavefield tomography. We demonstrate the cascaded workflow using a real 2D marine dataset. Our image-domain wavefield tomography is based on illumination constraints imposed in the inversion through the highlighting operator. This operator, allows to emphasize aspects of the gather one wants to improve. The criteria is based on how the gather should look like if the velocity model were correct. This knowledge is estimated by a demigration/migration process at each iteration. The illumination-based criterion is more realistic than the conventional DSO-based penalty functions as it incorporates information acquisition, data-bandwidth, and velocity information. Our tests with the 2D dataset show how the velocity model improves with the illumination-based tomography followed by the data-domain tomography. The second, adds details in the shallower part of the model; thus, increasing the consistency between modeled and observed early arrivals (not constrained by the image-domain tomography).

2.7 Acknowledgments

We thank sponsor companies of the Center for Wave Phenomena. The seismic data shown in this article is proprietary to and provided courtesy of CGG. We thank Bruce VerWest for the support with the dataset. The reproducible numeric examples in this paper use the Madagascar open-source software package (Fomel et al., 2013) freely available from http://www.ahay.org.

CHAPTER 3

SEISMIC TOMOGRAPHY USING LOCAL-CORRELATION FUNCTIONS[6]

Esteban Díaz and Paul Sava

Center for Wave Phenomena, Colorado School of Mines

Wavefield tomography in the data domain is usually formulated by using the data difference as the misfit criterion. The data difference, however, is susceptible to cycle skipping which leads to convergence into local minima. To overcome the strong non-linearity, a multi-scale approach to the inversion is often needed. The success of this approach relies on the low-frequency content of the data. Here, we define the data domain tomography misfit criterion using local correlations. Correlation-based inversions are less sensitive to local minima than difference-based inversions. Correlations, however, are often contaminated with cross-talk between events; in addition, the global correlations give just a general idea of the kinematic errors of the model because of the summation along the entire time axis. Alternatively, local correlations with Gaussian windows, advocated in this paper, are able to extract the local kinematic errors in the misfit between modeled and observed data. Local correlations are also less sensitive to cross-talk of seismic events than their global correlations counterpart because the summation is performed locally as a function of time. Less correlation cross-talk leads to cleaner adjoint sources and hence, cleaner gradients. We further improve the gradients using a penalty function that is consistent with the bandwidth of the seismic data, which is more realistic than linear penalty functions designed to annihilate infinite bandwidth data.

3.1 Introduction

Waveform tomography (WT) methods have the ability to recover velocity models that are consistent with the data bandwidth. Kernels of wavefield tomography contain higher

[6]IN PREPARATION FOR GEOPHYSICS

sensitivity to low wavenumbers compared to those of rayfield tomography. This is because the update is carried by wavepaths that are sensitive to the bandwidth of the data, as opposed to the update carried by rays that are defined in the context of high frequency asymptotic and that sample the model on an infinitesimal path. Typically, objective functions of WT use the data difference as a measure of the similarity between observed and modeled data (Tarantola, 1984; Lailly, 1983; Pratt, 1999). This criterion is highly non-linear and ill-posed (Virieux and Operto, 2009), mainly due to the fact that it compares oscillatory and band-limited functions; hence, an inversion can quickly fall into a local minima if the traveltime error is smaller than half a period (Virieux and Operto, 2009). This non-linearity requires a multi-scale approach where the model is updated from low wavenumbers to high wavenumbers (Bunks et al., 1995), and this approach often imposes serious low-frequency requirements on the data used for FWI.

Other waveform inversion methods sacrifice the resolution of the data difference for the robustness to errors beyond half a period. Luo and Schuster (1991) propose to use the time-lag (measured at the maximum of the correlation between observed and modeled data) as the misfit criteria. van Leeuwen and Mulder (2010) demonstrate that penalizing the correlation, instead of picking its maximum, is more robust to phase difference errors due to the wrong choice of source functions to generate the modeled data (for which the maximum of the correlation does not match the true travel time error). Luo and Sava (2012) suggest to use deconvolution (correlation with the inverse of the modeled data) as opposed to direct correlation and assert that the deconvolved misfit criterion contains better resolution than the conventional correlation approach. Warner and Guasch (2014) continue along the same line of deconvolution functionals through minimization of the non-zero lag components of the Wiener filter, which are computed from the observed data to match the modeled data. All these waveform tomography methods are optimized with gradient descent minimization methods. The gradients are computed efficiently subject to wave equation constraints implemented within the framework of the adjoint state method (Plessix, 2006).

Perrone et al. (2015) use local correlation functions between migrated seismic images to define the tomography problem. Here, we propose to use local temporal correlation functions to extract the misfit information in the data domain (at the receiver locations), instead of measuring such correlations in a global sense, as suggested by other approaches. To compute the local correlation functions, we use the method of Hale (2006a), which performs the local summation in an efficient manner using recursive Gaussian filters (Hale, 2006b). Local correlations minimize the cross-talk between seismic events, since the comparison between observed and simulated waveforms is performed locally in time. This is a key advantage of our method because it has the ability to produce cleaner adjoint sources, which translate into cleaner gradients (with less cross-talk than the global correlation functionals).

This paper is organized as follows: we start by reviewing the theory behind global and local-correlation functions, where we also define the corresponding adjoint pair for the local-correlation linear operator. Then, we introduce the theory for the wavefield tomography problem based on our new functional, and we show its corresponding implementation using the adjoint state method (Plessix, 2006). We then illustrate the method with a synthetic example and discuss different approaches for applying penalization functions to global and local correlations.

3.2 Correlation functionals.

Consider the problem of measuring the similarity between two signals $f(t)$ and $g(t)$ (Figures 3.1(a)-3.1(b)). One way is to measure the difference between f and g. This approach is highly sensitive if the two signals are close enough (in a kinematic sense) to each other. However, if the two signals contain a kinematic difference longer than half a period, this functional is not informative (Bunks et al., 1995). Instead of determining the similarity by measuring the difference, we can determine the similarity through correlation. In the most common case, one can use a global correlation function (van Leeuwen and Mulder, 2010).

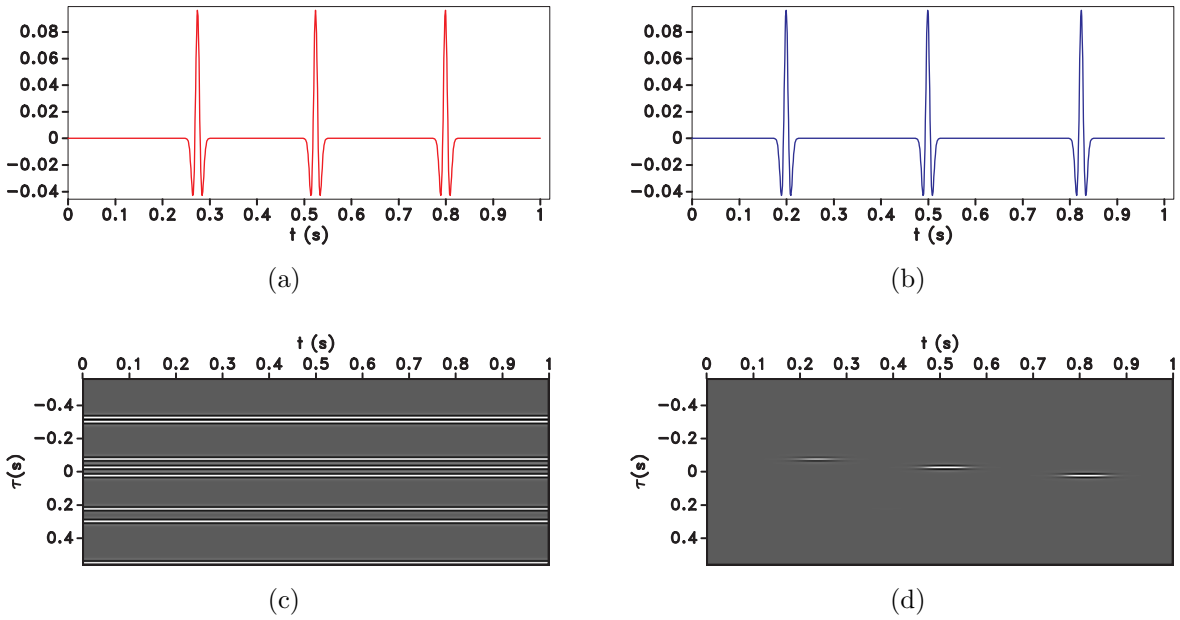


Figure 3.1: Signals (a) $f(t)$ and (b) $g(t)$ to exemplify the correlation functions, (c) is the global correlation $f \star g$ (invariant with time), and (d) $f \star_l g$ is the local correlation with a Gaussian window with $\sigma = 0.3s$.

Here, we follow the convention of Hale (2006a) and define centered correlations:

$$c(\tau) = (f \star g)(\tau) = \int_{-\infty}^{\infty} f(t' - \tau/2) g(t' + \tau/2) dt'. \tag{3.1}$$

The correlation function $c(\tau)$ gives us a general sense of the similarity between $f(t)$ and $g(t)$. Alternatively, we can estimate the similarity in a localized way (Hale, 2006a), where we can capture how the two signals are related to each other along the time axis:

$$c(t, \tau) = (f \star_l g)(t, \tau) = \int_{-\infty}^{\infty} w(t' - t) f(t' - \tau/2) g(t' + \tau/2) dt', \tag{3.2}$$

where $w(t' - t) \equiv e^{-\tau^2/4\sigma^2} e^{-(t'-t)^2/\sigma^2}$ is a Gaussian window centered at t', and σ is standard deviation of the Gaussian function. As $\sigma \to \infty$, the local correlations behave like the global correlations at each time t, i.e. they are invariant with respect to time t. This localized formulation is equivalent to convolving the shifted product $h_\tau(t') = f(t' - \tau/2) g(t' + \tau/2)$ with a Gaussian window: $c(t, \tau) = (G * h_\tau)(t)$. The convolution is efficiently implemented

using recursive Gaussian filters (for which the convolution cost is independent of the Gaussian half-width σ) (Hale, 2006b). One can think about equation 3.2 as a linear operator defined as $\mathbf{c} = \mathbf{C}\mathbf{g} = \mathbf{G}\mathbf{S}\mathbf{g}$, where $\mathbf{C} = \mathbf{G}\mathbf{S}$ is the local correlation matrix, $\mathbf{G}$ is a block diagonal matrix (with each block being a convolution matrix with the Gaussian function), and $\mathbf{S}$ contains shifted versions of f. The transposed (adjoint) operator $\mathbf{C}^\top$ implements a local convolution as $\mathbf{C}^\top = \mathbf{S}^\top\mathbf{G}$. By construction, the matrix $\mathbf{G}$ is SPD since all of the blocks are SPD: hence, $\mathbf{G}^\top = \mathbf{G}$. Mathematically, the adjoint operator is implemented as follows:

$$f^*(t) = \int_{\tau_{min}}^{\tau_{max}} \left(\int_{-\infty}^{\infty} w(t' - t + \tau/2) c(t' + \tau/2, \tau) dt' \right) g(t + \tau) d\tau, \tag{3.3}$$

where τ_{min} and τ_{max} are the minimum and maximum time-lags, respectively, and f^* represents the adjoint of f.

Figure 3.1(c) shows the global correlation between the signals in Figures 3.1(a) and 3.1(b). Note that the global correlation is time-invariant and contains cross-talk between events (around $\tau = -0.3$s and $\tau = 0.3$s), the cross-talk is captured because the correlation lags are big enough to facilitate interference between different events. In contrast, Figure 3.1(d) shows the local correlation between $f(t)$ and $g(t)$; no cross-talk is observed because the Gaussian window prevents interference between distant events. However, it is still possible to observe cross-talk between events using local correlations if the events lie within the correlation window of standard deviation σ.

3.3 tomography with correlation functions

Here, we show the application of correlation functionals for the inverse tomographic problem under the physics of the acoustic wave equation

$$m(\mathbf{x}) \frac{\partial^2 u}{\partial t^2} - \nabla^2 u = s, \tag{3.4}$$

where $u(\mathbf{x}, t)$ is the wavefield excited by the source function $s(\mathbf{x}, t)$, and $v(\mathbf{x})$ is the medium P-wave velocity. We can extract data from the wavefield by using a restriction operator at the receiver positions $\mathbf{x}_r$ as $d(\mathbf{x}_r, t) = u(\mathbf{x}, t)\delta(\mathbf{x} - \mathbf{x}_r)$. For the remainder of this paper, we

express the wave equation as the linear operator $\mathbf{L}(m)$ with its corresponding adjoint operator $\mathbf{L}^{\top}(m)$, where $m = 1/v^2(\mathbf{x})$ is the model parameter representing the slowness squared. We can design a correlation-based objective function that maximizes the correlation between observed and modeled data:

$$J(m) = \frac{||P(\tau)d^o \star_l d^m||_2}{||d^o \star_l d^m||_2} = \frac{||\mathbf{PCd^m}||_2}{||\mathbf{Cd^m}||_2}, \tag{3.5}$$

where $P(\tau)$ is a penalty operator that enhances energy outside $\tau = 0$ (which we want to minimize), m is the model parameter, and $d^o(e, \mathbf{x}_r, t)$ and $d^m(e, \mathbf{x}_r, t)$ are the observed and modeled data for each experiment e, respectively. The correlation matrix $\mathbf{C}$ is built using shifted versions of the observed data d^o along with the Gaussian convolution, as defined by the process in equation 3.2. The application of the penalty operator $P(\tau)$ is equivalent to a diagonal penalty matrix $\mathbf{P}$. Here, the L2 norm squared $||.||_2$ is defined over receiver coordinates $\mathbf{x}_r$, experiment index e, time t, and correlation lag τ. Later, we discuss strategies for choosing the penalty operator $P(\tau)$.

The objective function in equation 3.5 is similar to that of van Leeuwen and Mulder (2010) except that we use local instead of global correlation, and we normalize the denominator. The normalization balances the objective function by the energy of the correlation function (Warner and Guasch, 2014). This step eliminates the objective function bias for positive and negative errors.

To solve the optimization problem, we use gradient descent methods. The adjoint state method framework (Plessix, 2006) allows us to obtain the gradient by defining our problem under equality constraints given by the wave equation as

$$\begin{aligned} \underset{m}{\text{minimize}} \quad & J(m) \\ \text{subject to} \quad & L_e u = s. \end{aligned} \tag{3.6}$$

These constraints satisfy the wave equation for the modeled data d^m for every experiment e. We define the augmented functional as

$$H(u_s, a_s, m) = J(m) - \sum_e < a_s, Lu_s - s >, \tag{3.7}$$

from which we seek to obtain the appropriate Lagrange multipliers $a_s(e, \mathbf{x}, t)$ (adjoint wavefields) that satisfy the constrained equation. We use the notation $< ., . >$ for the dot product between two vectors.

The stationary points of the Lagrangian provide us with the equations to solve the problem. The first condition

$$\partial_{a_s} H = 0 \tag{3.8}$$

leads to the state variable problem

$$\mathbf{L}(m)u_s(e, \mathbf{x}, t) = s(e, \mathbf{x}, t) \tag{3.9}$$

for each experiment e. The second condition

$$\partial_{u_s} H = 0 \tag{3.10}$$

leads to the adjoint state variable system

$$\mathbf{L}^\top(m)a_s(e, \mathbf{x}, t) = g_s(e, \mathbf{x}, t), \tag{3.11}$$

where $g_s(e, \mathbf{x}, t)$ is the adjoint source for our specific choice of objective function. The adjoint sources are given by

$$g_s = -\frac{1}{||\mathbf{C}d^m||_2}\mathbf{C}^\top \left(\mathbf{P}^\top\mathbf{P} - J\mathbf{I}\right)\mathbf{C}d^m. \tag{3.12}$$

This expression tells us to twice penalize the correlation and subtract the value of the objective function from it before computing the adjoint correlation. This process effectively balances the influence between the numerator and denominator in the adjoint source. At early iterations, when J is larger, the norm of the correlation has greater influence on the adjoint source; as iterations continue, the penalty term becomes more prominent. The cor-

rect calculation of the adjoint source relies on the fact that the linear operators and their corresponding transpose (adjoint) operators are correctly implemented. We make sure this is the case by applying the dot product test (Claerbout, 2008) on each linear operator. Once we solve for u_s and a_s, we can compute the gradient of J with the last condition

$$\partial_m H = 0, \tag{3.13}$$

from which we obtain the gradient

$$\partial_m J = \sum_{t,e} \ddot{u}_s a_s, \tag{3.14}$$

where the double dot indicates the second derivative in time.

3.3.1 Penalty operators

The penalty operator $P(\tau)$ has the objective of enhancing the energy that does not comply with our assumption of maximum correlation near zero lag for a correct velocity model. A conventional penalty function is $P(\tau) = |\tau|$ (van Leeuwen and Mulder, 2010; Luo and Sava, 2012) based on the concept of differential semblance optimization (Symes, 2008), and is designed to annihilate energy linearly as a function of τ. This penalty function goes to zero only if the data, and thus the correlation, has infinite bandwidth. This linear weight becomes less sensitive as the data bandwidth decreases. For band-limited data, this penalty can produce high residuals even for a correct model. Alternatively, we can construct a penalty function that is consistent with the bandwidth of the data, similar to the procedure used by Yang et al. (2013) for image-domain wavefield tomography. Thus, we define the penalty using the auto-correlation of the observed data:

$$P(\tau) = \sum_{\mathbf{x}_r} \frac{1}{Env(d^o \star d^o(\tau, \mathbf{x}_r)) + \epsilon}. \tag{3.15}$$

The envelope function Env captures the energy of the auto-correlation, and ϵ is a stabilization factor for the division.

We illustrate the method with a simple model with constant velocity and with a free surface. We perturb the medium with a Gaussian anomaly in such a way that the free sur-

face reflection arrival time is not affected. Figures 3.2(a), 3.2(c), and 3.2(e) show the models for positive, zero, and negative Gaussian anomaly, respectively. Figures 3.2(b), 3.2(d), and 3.2(f) show the corresponding data, for positive, zero, and negative Gaussian anomalies. Figures 3.3(a), 3.3(c), and 3.3(e) show the global correlations for negative, correct, and positive perturbations, respectively. One can see the cross-talk in the correlations for lags longer than $\tau = \pm 0.1$s. Figures 3.3(b), 3.3(d), and 3.3(f) depict the local correlations for the same events. In this case, the cross-talk is greatly attenuated, and the correlation events are localized without interference along the time axis. The cross-talk in the global correlations is propagated into the corresponding gradients through the adjoint sources. Figures 3.4(a), 3.4(c), and 3.4(e) show the gradients from the global correlations. Strong cross-talk between events is visible as high wavenumber (reflection-like) events coincident at the free surface and elsewhere. However, the gradients from local correlations (Figures 3.4(b), 3.4(d), and 3.4(f)) are cleaner and present the correct gradient direction without the contamination from cross-talk, which is almost non-existent for the chosen window. In this simple setting, we expect two sensitivity paths: (1) from source to receiver, and (2) from source to the free surface and from the free surface to the receivers. However, we can observe several more events in the global correlation gradients. These extra events come from the cross-talk present in global correlations, which is then propagated into the adjoint wavefields.

3.4 Transmission tomography with Gaussian anomalies

In this example, we use a transmission setting with the model shown in Figure 3.5(a). The model contains a background velocity of 3km/s and two Gaussian anomalies of ± 0.6km/s. The dots on the left of the model represent the source locations, and the line on the right is the receiver line. We perform inversion with a starting model (Figure 3.5(b)) equal to the constant background velocity of 3km/s. The two Gaussian anomalies produce triplications in the wavefield as shown by the observed data in Figure 3.5(c). The data from the starting model is in Figure 3.5(d). Figures 3.5(e)-3.5(f) show the auto-correlation of observed data for a particular receiver at $z = 1.85$km, and the correlation of the initial data with observed

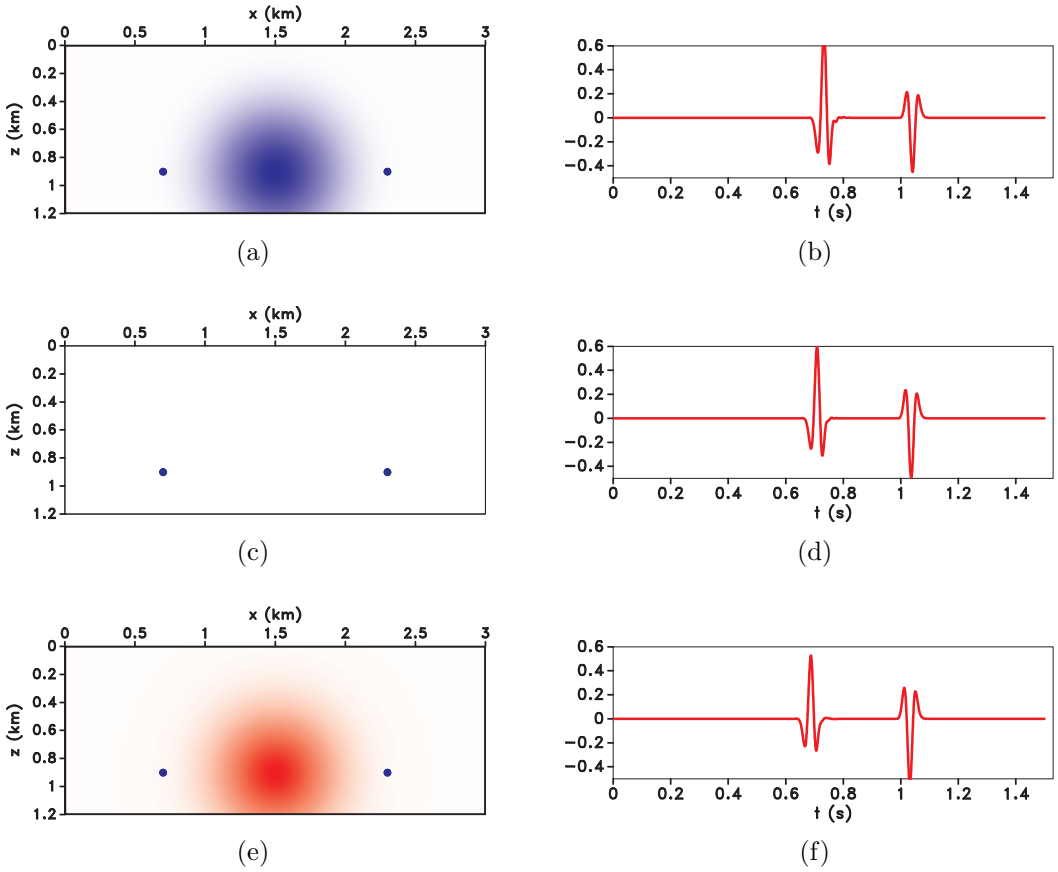


Figure 3.2: Sensitivity kernel experiment: the left column shows the models for (a) negative, (c) zero, and (e) positive Gaussian anomaly. The right column (b), (d), and (f) show the corresponding modeled data. For computing both the kernels we use the constant background model.

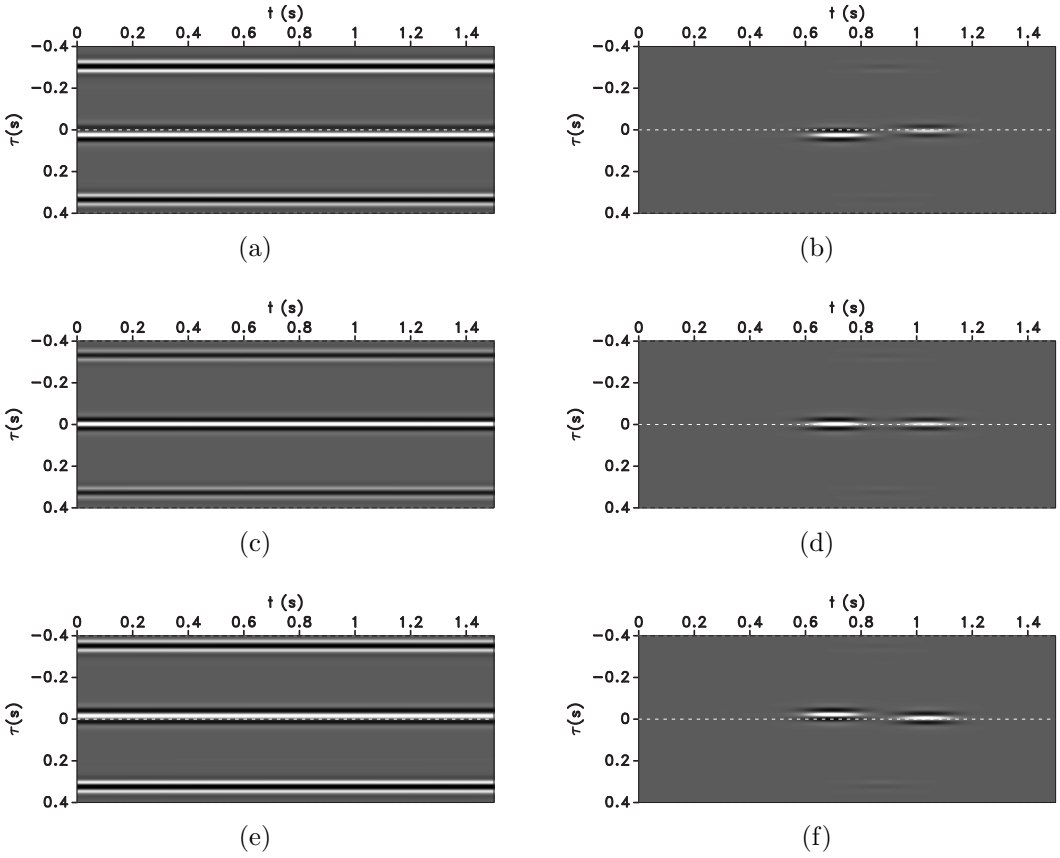


Figure 3.3: Sensitivity kernel experiment: the left and right columns show global and local correlations, respectively. From top to bottom one can see correlations for negative, zero, and positive Gaussian perturbations, respectively. Note how the local correlations clearly show the kinematic errors for the direct arrival event, whereas the global correlation shows the two events together. Also, the global correlation contains strong cross-talk between the two different events.

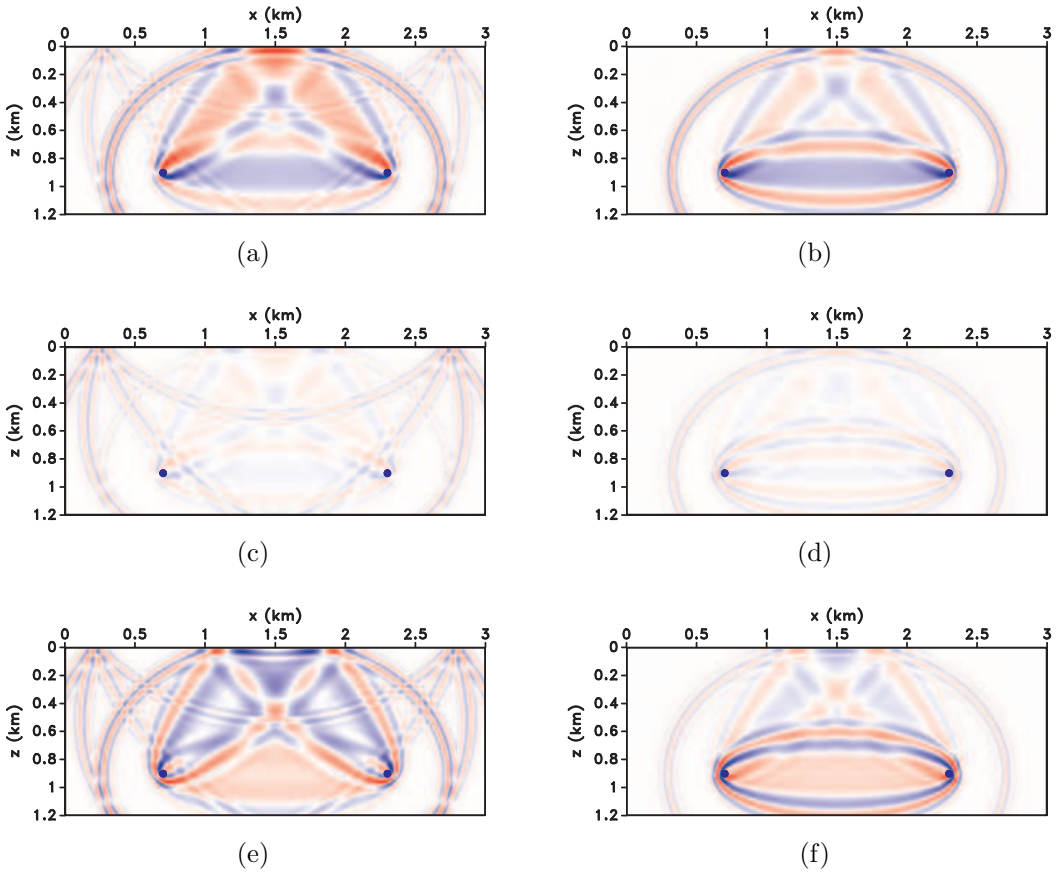


Figure 3.4: Sensitivity kernels for an experiment with free surface boundary condition and a Gaussian anomaly placed in between the source and the receiver (for which the surface reflection event is not sensitive). The left and right columns show gradients for global and local correlations, respectively. From top to bottom, the gradients correspond to low, correct, and high velocity.

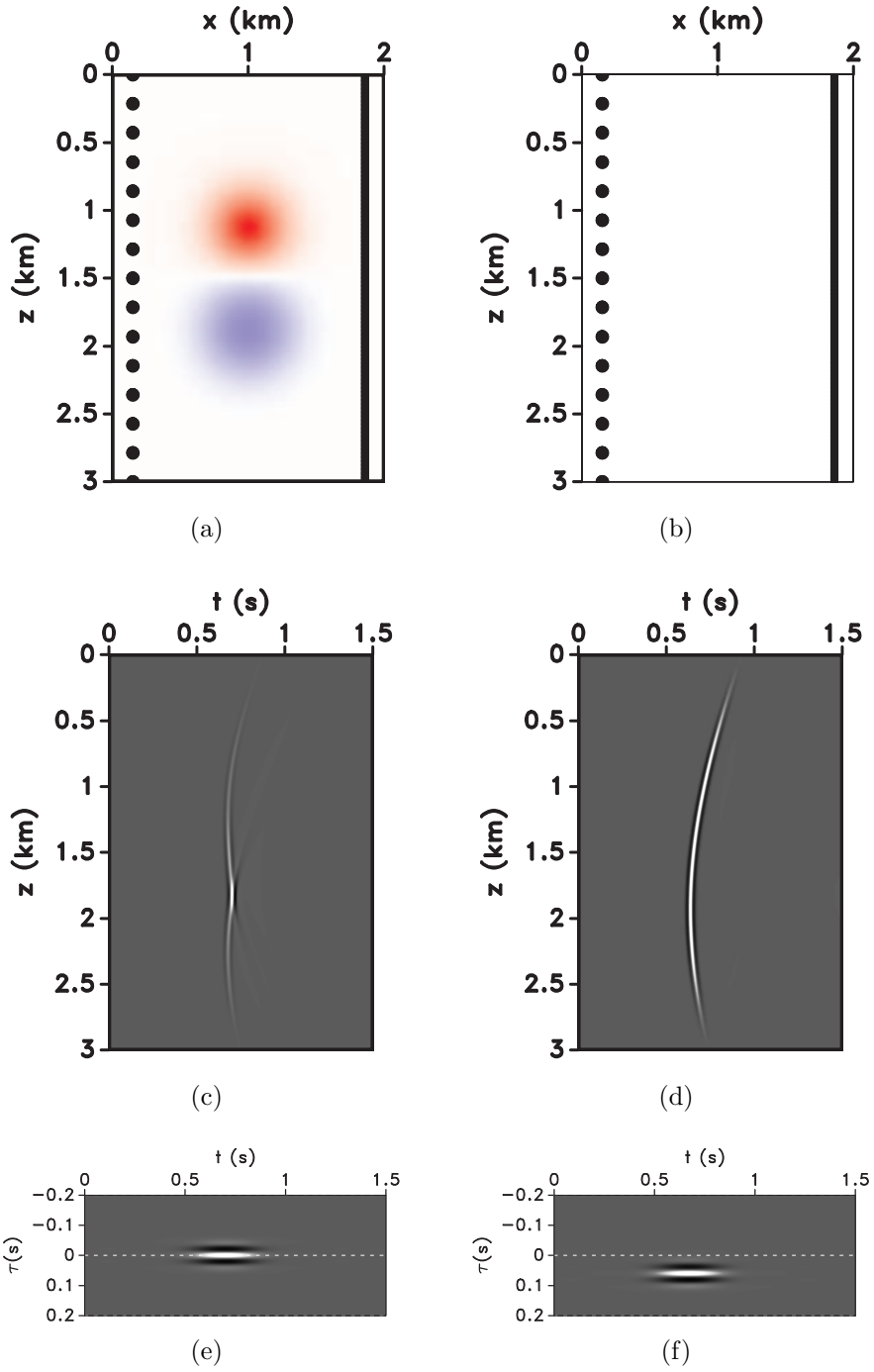


Figure 3.5: Transmission example with two Gaussian anomalies of opposite sign setup: (a) shows the true model, (b) the starting model (constant background velocity), (c) depicts the observed data from the highlighted shot point at $z = 1.9$km, (d) the modeled data from the background model, (e) the local auto-correlation of the observed data, and (f) the local correlation of initial data with observed data. The horizontal shows the location of the correlation functions.

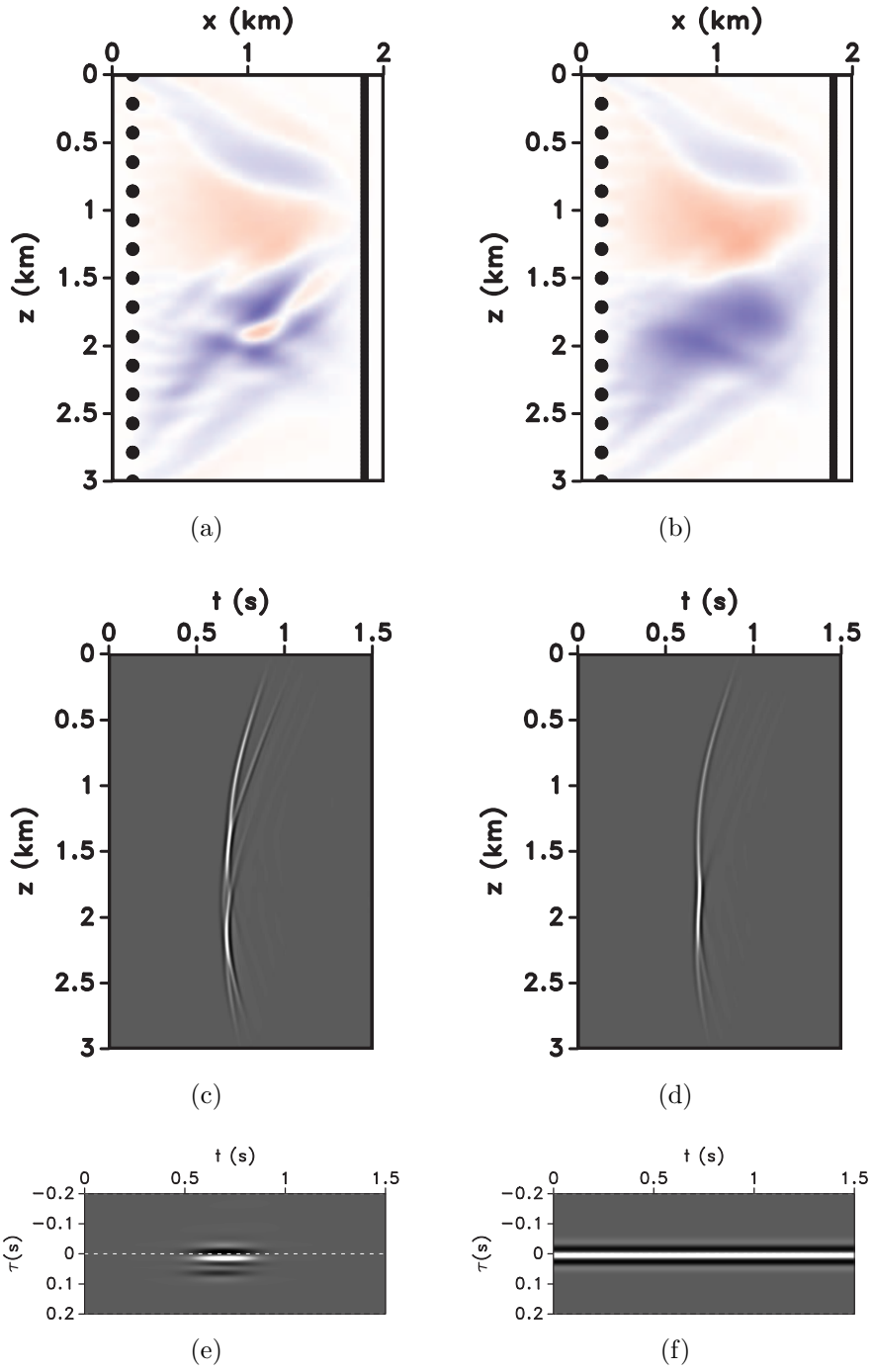


Figure 3.6: Inversion results: (a) shows the inverted model from local correlations, (b) the inverted model from global correlations, (c) depicts the modeled data from model in (a), (d) the modeled data from the model in (b), (e) the correlation at the final iteration for local correlation inversion, and (f) the correlation at the final iteration for the global correlation inversion. The horizontal line shows the location of the correlation functions.

data, respectively. Figure 3.6(a) is the retrieved model after using the local correlation objective function, whereas Figure 3.6(b) is the inverted model from the global correlation. A comparison of the two figures reveals that the model is better retrieved with the local correlation framework. Figures 3.5(c), 3.6(c), and 3.6(d) show the modeled data for the highlighted source point in the true model, local correlation model, and global correlation model, respectively. Even though the global correlation inversion recovers the correct sign of the anomaly, it is not able to create data similar to that of the true model. In contrast, the data from the local correlation model correctly recovers the triplication produced by the low velocity zone located in the lower half of the model. Figures 3.6(e)-3.6(f) show the correlation functions for the model retrieved with local and global correlations, respectively. The model retrieved with global correlations produces extra events when compared to the correlation from the local correlation inversion, which is simpler. The correlation in Figure 3.6(e) obtained from the model in Figure 3.6(a) matches the auto-correlation in Figure 3.5(e). Figure 3.7 shows the convergence curves for the local correlation inversion (solid line) and the global correlation inversion (dashed line). The local correlation inversion converges significantly faster than the global inversion.

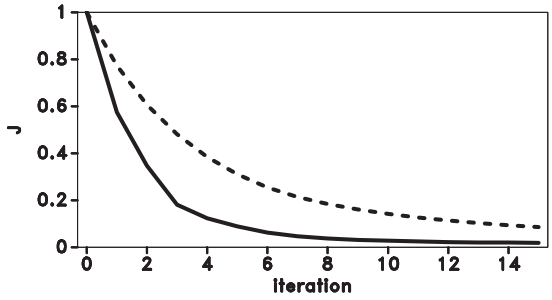


Figure 3.7: Convergence curves for the Gaussian anomalies model for local correlations (solid line) and global correlations (dashed line). Note how the local correlation convergence is faster than the global correlation inversion.

3.5 Transmission tomography with the Marmousi model

We also demonstrate our method with a portion of the Marmousi–2 model (Martin et al., 2002) in a transmission setup. We slightly smooth the original model to create the true model used in this example (Figure 3.10(a)). The acquisition geometry consists of 28 shots located at $x = 8.12$km equally spaced in the vertical direction. On the right hand side of the model, we place the receiver line at $x = 12.5$km. Figure 3.10(c) shows the data for the shot at $\mathbf{x}_s = (8.12, 1.95)$km.

We smooth the true model to create the starting model, Figure 3.10(b). Figures 3.10(c)-3.10(d) show the observed and modeled data from the true and starting models, respectively. The initial model is only able to generate a simple arrival, which differs from the observed data shown in Figure 3.10(c). Figures 3.10(e)-3.10(f) show the correlation functions for the data corresponding to the true model (which is an auto-correlation) and the starting model, respectively.

Before proceeding with the inversion, we perform different tests for the window parameter σ. Figures 3.8(a)-3.8(e) show the gradients for $\sigma = 0.1$s, $\sigma = 0.15$s, $\sigma = 0.25$s, $\sigma = 3$, respectively. As one increases σ, the correlation becomes effectively global. Hence, the more cross-talk is generated in the adjoint sources and is then transferred into the gradients. For lower σ, the gradient is simpler and smoother. For comparison, we include a conventional full waveform inversion (FWI) data difference gradient in Figure 3.8(f). For some σ values, the gradient is smoother than the FWI gradient. To better understand the differences between different σ, one can observe the corresponding adjoint sources in Figures 3.9(a)-3.9(e). As the σ parameter becomes smaller, the adjoint source is more influenced by the model data and contains less cross-talk. As σ increases, there are more possibilities for cross-talk between events, which generates coherent noise that is transfered into the gradients. The simplicity of the adjoint source for small σ is similar to that of the modeled data. The FWI adjoint source is dominated by the events present in the observed data. Based on our tests, we chose $\sigma = 0.15$s for the inversion since the retrieved gradients are smoother, and the adjoint source

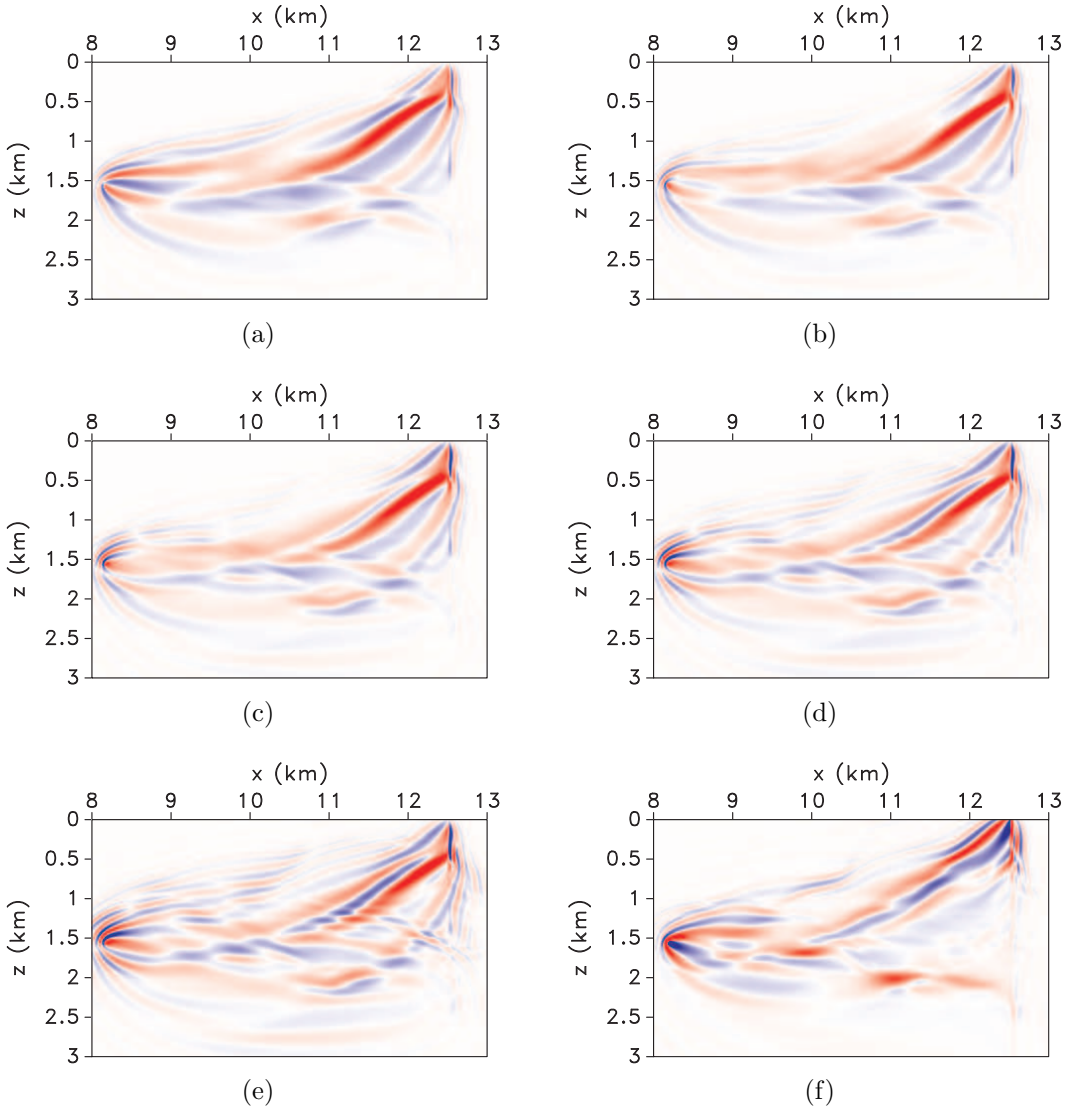


Figure 3.8: Marmousi model parametrization testing: (a)-(e) show local correlation gradients for $\sigma = 0.1$s, $\sigma = 0.15$s, $\sigma = 0.25$s, $\sigma = 3$s; (f) shows the conventional FWI gradient for comparison. Note how the complexity and cross-talk increases with σ.

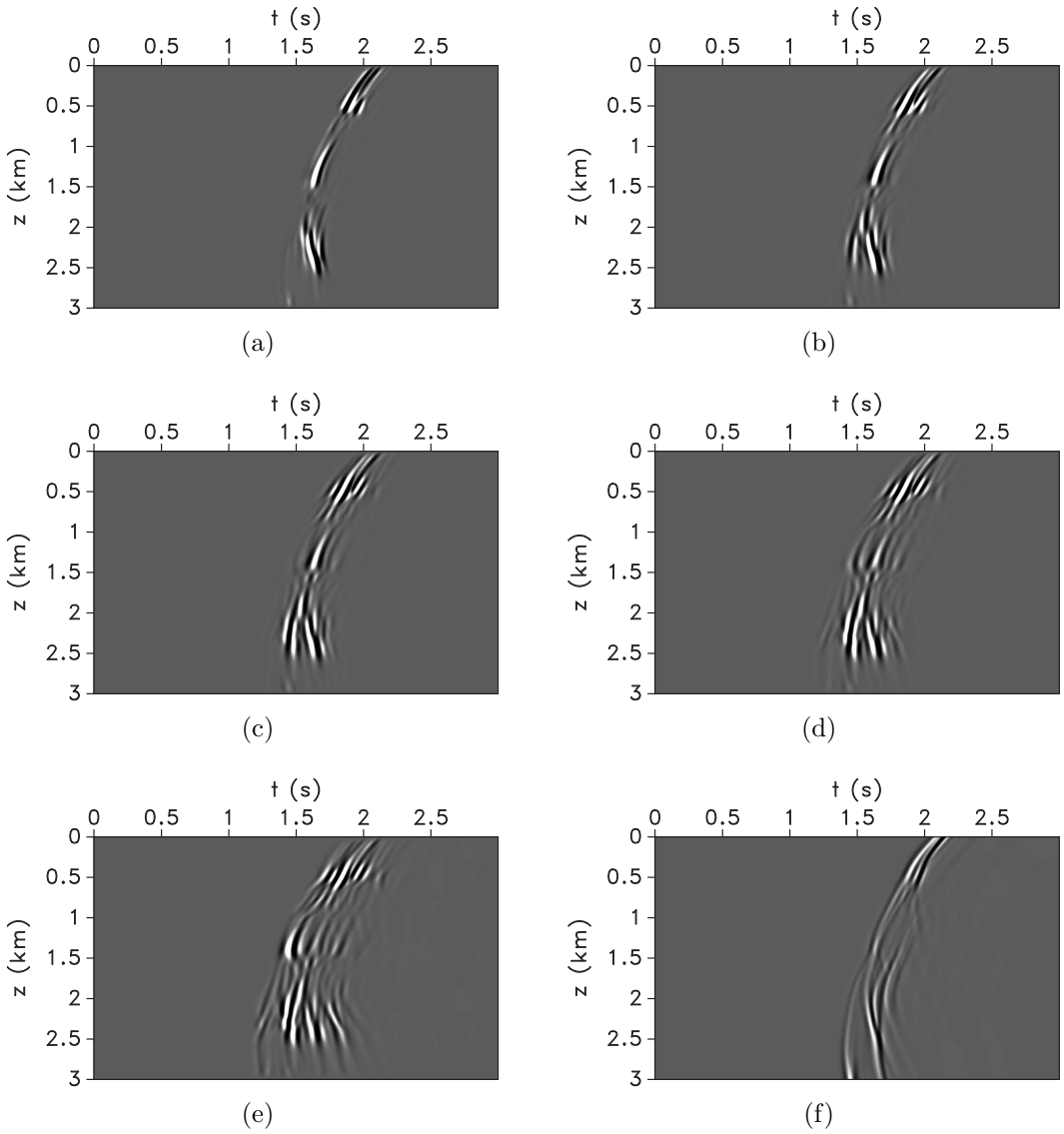


Figure 3.9: Marmousi model parametrization testing: (a)-(e) show local correlation adjoint sources for $\sigma = 0.1$s, $\sigma = 0.15$s, $\sigma = 0.25$s, $\sigma = 3$s; (f) shows the conventional adjoint source for FWI. Note how the complexity and cross-talk in the adjoint source increases with σ.

preserves the simplicity of the model data while avoiding cross-talk.

Figures 3.11(a)-3.11(b) show the inverted models for local and global correlations, respectively, and Figures 3.11(c)-3.11(d) show the corresponding gradients at the first iteration ; the gradient from the local framework contains smoother features, whereas the global correlation gradient contains higher wavenumber events which arise from cross-talk injected into the adjoint sources. Figures 3.11(e)-3.11(f) show the corresponding modeled data for local and global correlations, respectively. The inverted models add complexity into the seismic arrivals, which correctly matches the phase of the early arrivals from the observed data (Figure 3.10(c)). However, both models are unable to retrieve most of the later arrivals from the observed data. One reason for this behavior could be the strong gradient in the velocity model, which can produce poor illumination in the shallow section of the model. In the deeper part of the model, both inversions (in particular, the one from the local correlation framework) recover the high velocity layer near the reservoir. Underneath the reservoir, both inversions introduce a low velocity zone, which is also present in the true velocity model. Figures 3.11(g)-3.11(h) show the correlation functions in the inverted models, for local and global correlations, respectively. Both of the inverted models produce correlations closer to $\tau = 0$ compared to those of the starting model. Figure 3.12 shows the convergence curves for the local correlation (solid line) and for the global correlation inversion (dashed line). Again, the local correlation converges faster.

3.6 Conclusions

Local as opposed to global correlations provide cleaner measures of similarity between two functions. The main difference is that local correlation performs a local sum instead of summing over all the time samples. This is the key to eliminating the comparison between unrelated seismic events present in the observed and modeled data. This feature allows us to extract local kinematic errors that are more instructive than the misfit extracted using global correlations. Less cross-talk in the correlations translates into cleaner adjoint sources, which can produce more informative model updates than those of global correlation framework.

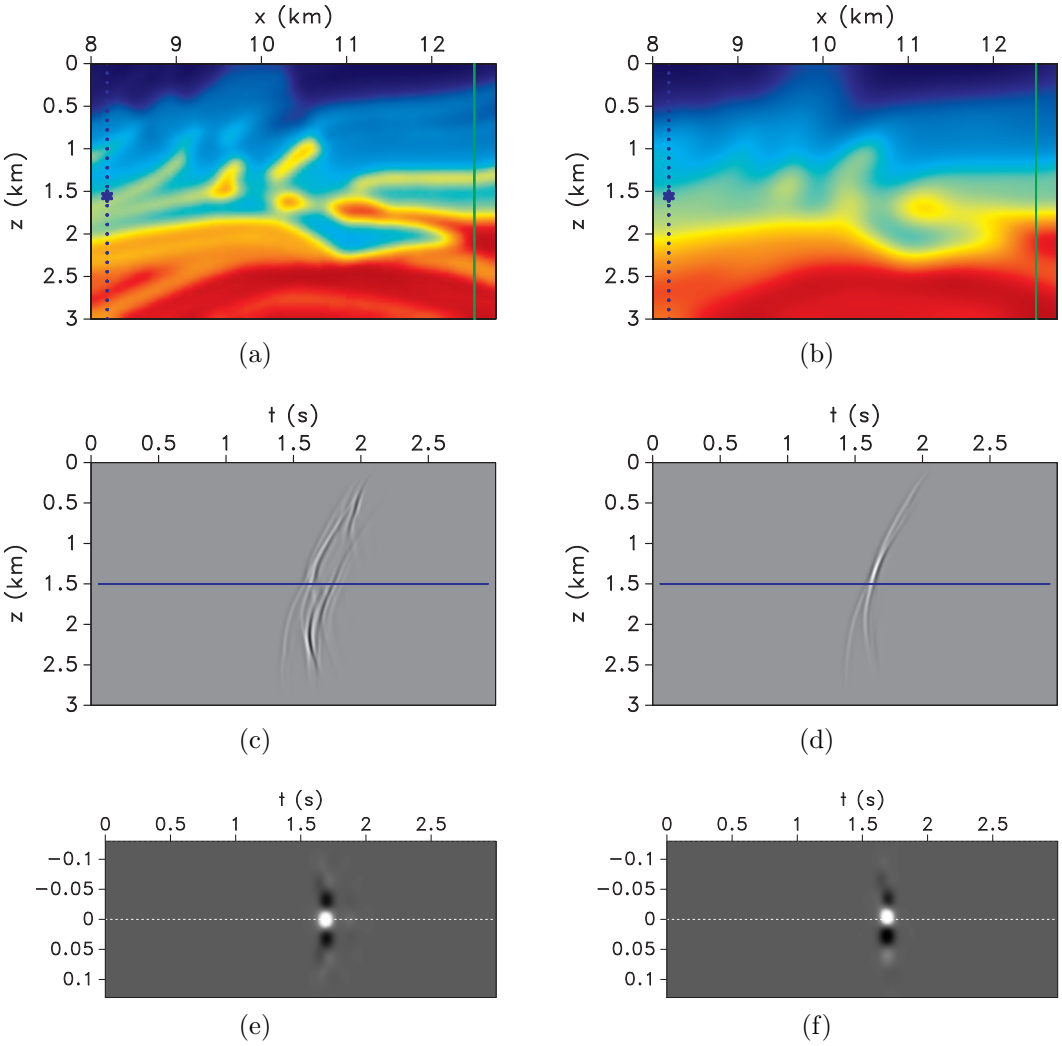


Figure 3.10: Marmousi cross-well experiment: (a) the true velocity model, (b) the starting velocity model, (c) the observed data, (d) the same shot gather in the initial model, (e) the local auto-correlation extracted at $z = 1.2$km, and (f) shows the local correlation from the initial data. Points in the left side of the model show the shot positions, the largest point shows the shot position for the depicted data panels. The line on the right shows the receiver array.

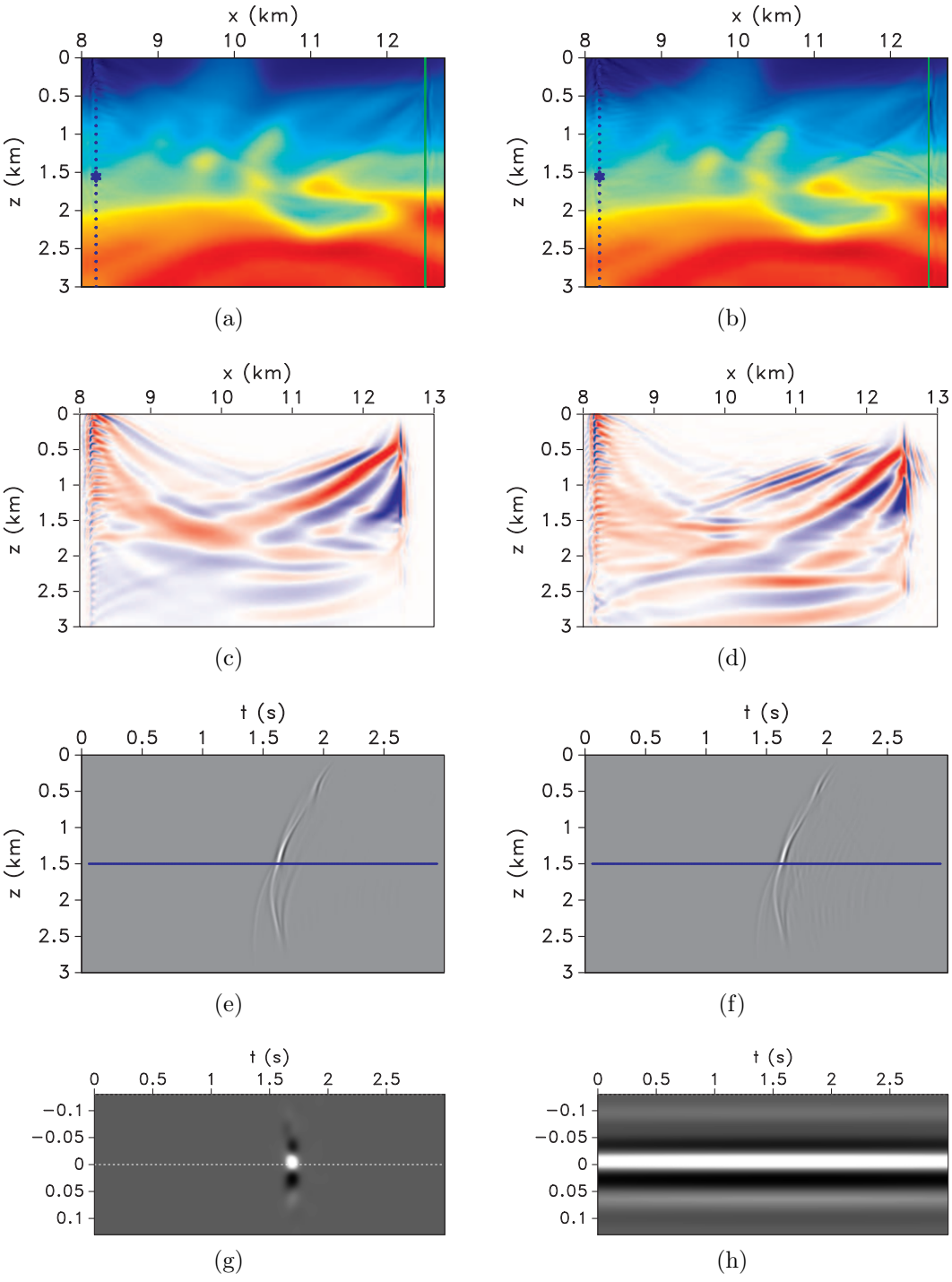


Figure 3.11: Inversion results from (a) local , and (b) global correlations. Panel (c) shows the initial gradient from the local correlation framework, whereas (d) depicts the initial gradient from the global correlation. The data in (e) and (f) are modeled from the local and global correlation inversion results, respectively. The bottom panels (g) and (h) show the correlation extracted at $z = 1.2$km for local and global correlations, respectively.

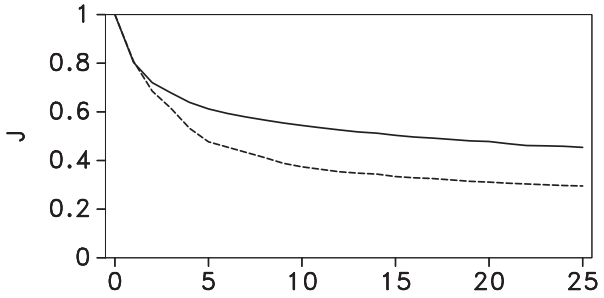


Figure 3.12: Convergence curves for the Marmousi model for local correlations (solid line) and global correlations (dashed line). Note how the local correlation convergence is faster than the global correlation inversion.

We also propose to use a bandlimited penalty function, which is constructed from the auto-correlation of the source function. This penalty function ensures faster convergence since it annihilates events in the correlation function that are consistent with the data bandwidth. Once the local correlation misfit is minimum, one can continue with high resolution functionals based on the data difference, i.e., waveform inversion. We show that local correlations together with bandlimited penalty functions produce cleaner gradients and converge faster than a comparable inversion using global correlations. Our adjoint operator for the local correlations method suggests several possible applications including objective functions based on local deconvolution, which extends further to local-matched filters applications.

3.7 Acknowledgments

We thank sponsor companies of the Consortium Project on Seismic Inverse Methods for Complex Structures. The reproducible numeric examples in this paper use the Madagascar software package (Fomel et al., 2013).

CHAPTER 4

IMAGING THE MODEL THROUGH WAVE EQUATIONS[7]

Esteban Díaz, Satyan Singh, Roel Snieder, and Paul Sava

Center for Wave Phenomena, Colorado School of Mines

Marchenko two-way wavefields are able to reproduce complex wave phenomena that includes primaries, internal multiples and surface-related multiples. The wavefield contains reflections representative of the true model, but it is obtained by propagating with the kinematics of the background velocity model. We design an inverse problem from which we derive a model that explains the scattering phenomena present in the wavefield. Given the two-way nature of the Marchenko wavefields we can use the homogeneous wave equation to obtain images that are indicative of subsurface model parameters. Unlike conventional imaging methods, where the image is a representation of discontinuities in the subsurface, our method inverts for the actual physical properties throughout the subsurface. Numeric tests show that we can invert for images describing the true model even when the background velocity model is smooth. Furthermore, this technique enable us to image properties of the true model that are not explicit inputs in the algorithm used to generate the wavefield.

4.1 Introduction

Seismic imaging methods aim to estimate the Earth's reflectivity, i.e., a map of the interfaces that scatter propagating wavefields. However, a seismic image does not only have structural information, but also dynamic properties that contain important information about the subsurface parameters (Etgen et al., 2009). Over the years, despite great advances in seismic exploration technology, the underlying imaging mechanism (the imaging condition) has remained similar to the one proposed by Claerbout (1971). The imaging condition simply

[7]IN PREPARATION FOR GEOPHYSICS

states that reflectivity exists where the incident and scattered wavefields coincide in space and time. The amplitude is then given by the deconvolution or correlation of the incoming and outgoing fields.

The wavefield reconstruction step is where many advances occurred in recent years. The source and receiver wavefields can be extrapolated with integral formulations as in Kirchhoff migration (Schneider, 1978), one-way equations as in wave equation migration (WEM) (Gazdag, 1978; Gazdag and Sguazzero, 1984), or two-way equations as in reverse time migration (RTM) (Baysal et al., 1983; Whitmore, 1983; McMechan, 1983). The complexity of the velocity model and the geologic structure dictates which method is best suited for accurate imaging. However, the three outlined families of methods share a basic assumption: the single-scattering (Born) approximation. Hence, the multiple-scattered waves present in the data are imaged at inaccurate positions. Therefore, the multiples are usually eliminated from the data as a pre-processing step (Verschuur et al., 1992; Weglein et al., 1997; Guitton, 2005; Lin and Herrmann, 2013) or from the migrated gathers as a post-imaging step (Sava and Guitton, 2005; Wang et al., 2010; Weibull and Arntsen, 2013b).

In order to consider imaging beyond the Born approximation, several authors (Guitton, 2002; Grion et al., 2007; Verschuur and Berkhout, 2011; Whitmore et al., 2010; Wong et al., 2015) propose ideas for adapting more conventional technologies to handle multiples (surface-related in most cases). The adaptation requires prior separation of primaries and surface-related multiples. The primaries can then be used for imaging as an areal source for the down-going source wavefield, which correlates at the reflector position with the back propagated multiples. These approaches require an independent migration for each additional order of the multiple one seeks to incorporate into the seismic image, this progressively increasing the imaging cost.

Rose (2002a,b) proposes an inverse scattering approach based on focusing and time reversal which retrieves a function that focuses at an arbitrary point in an unknown 1D medium. This approach is extended to 2D by Broggini and Snieder (2012). Their method yields the

correct Green's function and properly handles primaries and internal multiples given a background velocity model with the appropriate kinematics. The Marchenko modeling framework (Broggini and Snieder, 2012; Behura et al., 2014; Wapenaar et al., 2014b; Singh et al., 2015a) solves for focusing functions that are used to compute the total Green's function, which can be decomposed into up- and down-going components. This framework is also applicable to surface-related multiples (Singh et al., 2015a,b). The Marchenko modeling framework possesses interesting characteristics and differences with respect to the conventional propagators mentioned earlier: (1) it solves for the up- and down-going Green's functions in one step (as opposed to two independent propagation problems), (2) the up- and down-going Green's functions add-up to the total Green's function (whereas the conventional source and receiver wavefields do not), and (3) it naturally handles all scattering orders present in the data (as opposed to Born-modeling).

We advocate the use of the total Green's function instead of the individual up- or down-going components (or source and receiver wavefields) for obtaining an image of the model (instead of the reflectivity). When using the individual components of the wavefield, one usually applies an imaging principle similar to the one formulated by Claerbout (1971) for a reflectivity image. Instead, we use the total Green's functions together with relations given by an assumed wave equation to obtain images sensitive to other subsurface properties like density or impedance. This paper is organized as follows: first we review the requirements and fundamentals for wave propagation with the Marchenko methodology, then we show how to obtain the subsurface properties from the wavefield; finally, we show synthetic examples in which we invert for the velocity model and density images.

4.2 Marchenko imaging

In this section, we review the Marchenko framework outlined in Singh et al. (2015a) for the Green's function retrieval with pressure-normalized fields. We outline the modeling equations with the objective of highlighting the technical requirements for application of the framework to model parameter estimation. We refer the readers to Wapenaar et al. (2014b)

and Singh et al. (2015a) for specific details about the derivation for flux- and pressure-normalized Marchenko frameworks, respectively.

4.2.1 Inputs for Marchenko imaging

Let $R(\mathbf{x}_s, \mathbf{x}_r, t)$ be the pressure normalized response of a medium for sources $\mathbf{x}_s$ and receivers $\mathbf{x}_r$ located at $z = 0$. Note that R contains primaries, internal multiples, and surface-related multiples. Given this reflection response, the Marchenko framework retrieves the Green's function from a virtual source located at an arbitrary location $\mathbf{x}$ inside the medium (Singh et al., 2015a). The process relies on the computation of focusing solutions that, when injected at the surface, produce a concentration of energy (a focus) at time $t = 0$ at location $\mathbf{x}$. In order to retrieve such focusing solutions, one must inject the time-reversed inverse of the transmission response $T_d(\mathbf{x}, \mathbf{x}_r, t)$ from the point of interest to the surface. In practice, the direct arrival in the background medium approximates the inverse of the transmission response (Singh et al., 2015a).

4.2.2 Modeling in the Marchenko framework

The Marchenko framework involves convolutions and correlations between the reflection response R, the focusing functions, and the transmission response T_d through an iterative process. Hence, the reflection response must be deconvolved with the source signature to avoid multiple convolutions with the source function.

The total wavefield $G(\mathbf{x}, \mathbf{x}_s, \omega)$ is the sum of its one-way components:

$$G(\mathbf{x}, \mathbf{x}_s, \omega) = G^+(\mathbf{x}, \mathbf{x}_s, \omega) + G^-(\mathbf{x}, \mathbf{x}_s, \omega), \tag{4.1}$$

and the up-going (G^-) and down-going (G^+) components are obtained through convolution and correlations of the reflection response R with the focusing functions f_1^+ and f_1^- (Singh et al., 2015a):

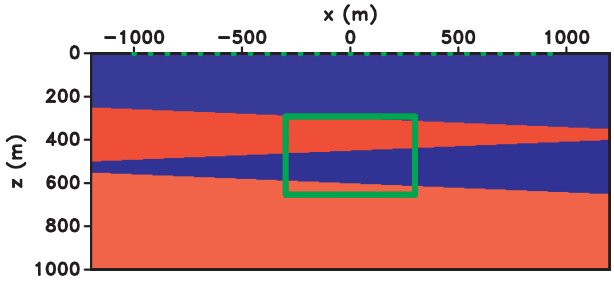
$$G^{-}(\mathbf{x},\mathbf{x}_r,\omega) = -f_1^{-}(\mathbf{x}_r,\mathbf{x},\omega) + \sum_{\mathbf{x}_s} f_1^{+}(\mathbf{x}_s,\mathbf{x},\omega)\, R(\mathbf{x}_r,\mathbf{x}_s,\omega) - r f_1^{-}(\mathbf{x}_s,\mathbf{x},\omega)\, R(\mathbf{x}_r,\mathbf{x}_s,\omega)\,, \tag{4.2}$$

$$G^{+}(\mathbf{x},\mathbf{x}_r,\omega) = +f_1^{+}(\mathbf{x}_r,\mathbf{x},\omega)^{*} - \sum_{\mathbf{x}_s} f_1^{-}(\mathbf{x}_s,\mathbf{x},\omega)^{*}\, R(\mathbf{x}_r,\mathbf{x}_s,\omega) - r f_1^{+}(\mathbf{x}_s,\mathbf{x},\omega)\, R(\mathbf{x}_r,\mathbf{x}_s,\omega)\,, \tag{4.3}$$

where r is the reflection coefficient at the free surface and * denotes complex conjugation. The individual components of the Green's function G satisfy the one-way reciprocity theorems shown by Wapenaar et al. (2014a), and the sum of its components satisfies the wave equation

$$\rho\nabla\cdot\left(\frac{1}{\rho}\nabla G\right) + \frac{\omega^2}{c^2}G = -j\omega\rho\delta(\mathbf{x}-\mathbf{x}_s). \tag{4.4}$$

The focusing functions are computed using the iterative framework described in Singh et al. (2015a).



(a)

Figure 4.1: Variable velocity model: the points on the surface denote source locations and the highlighted box denotes the area in which the Marchenko wavefield is retrieved.

4.3 Imaging model parameters

We use the total Green's function G to retrieve model parameters, like density and velocity, present in the wave equation and consider the modeling process in the previous section. The Marchenko modeling framework (Behura et al., 2014; Wapenaar et al., 2014b; Singh et al., 2015a) satisfies the acoustic wave equation; hence, we can use the retrieved wavefields as an input to the wave equation and solve for specific model parameters. A

similar idea have been successfully implemented in the medical community for finding elastic parameters of human tissues (Manduca et al., 2001). Here we show an inversion framework for acoustic parameters.

4.3.1 Constant velocity acoustics

Consider the constant density homogeneous wave equation

$$s^2 \frac{\partial^2 u}{\partial t^2} - \nabla^2 u = 0, \tag{4.5}$$

where $s^2(\mathbf{x})$ is the slowness squared and $u(\mathbf{x}_s, \mathbf{x}, t)$ the Marchenko wavefield generated from the source location $\mathbf{x}_s$. Note that the wavefield can be obtained inside the medium away from the source location $\mathbf{x}_s$ where the homogeneous equation is satisfied.

In order to solve for the unknown field $s^2(\mathbf{x})$, we set up a least-squares linear inverse problem based on the objective function

$$J(s^2) = \left|\left| s^2 \frac{\partial^2 u}{\partial t^2} - \nabla^2 u \right|\right|_2, \tag{4.6}$$

which has the formal solution:

$$s^2(\mathbf{x}) = \frac{\sum_{t,\mathbf{x}_s} \frac{\partial^2 u(\mathbf{x}_s,\mathbf{x},t)}{\partial t^2} \nabla^2 u(\mathbf{x}_s, \mathbf{x}, t)}{\sum_{t,\mathbf{x}_s} \left(\frac{\partial^2 u(\mathbf{x}_s,\mathbf{x},t)}{\partial t^2} \right)^2}. \tag{4.7}$$

This expression closely resembles the least squares solution for the deconvolution imaging condition (Claerbout, 1971; Guitton et al., 2007). Here, instead of correlating the source and receiver wavefields, we correlate the Laplacian of the total wavefield with its second time derivative.

The kinematics embedded in the Marchenko wavefield u are driven by the first arrival information $T_d(\mathbf{x}_s, \mathbf{x}, t)$, which is computed using a background velocity model. The inversion procedure outlined in equation 4.6 utilizes the geometric attributes of the wavefield; therefore, the wave equation is satisfied for the same background velocity model used for computing the transmission response $T_d(\mathbf{x}_s, \mathbf{x}, t)$.

4.3.2 Variable density acoustics

For a variable density field into the wave equation, the Marchenko wavefield u satisfies

$$\frac{s^2}{\rho}\frac{\partial^2 u}{\partial t^2} - \nabla \cdot \left(\frac{1}{\rho}\nabla u\right) = 0. \tag{4.8}$$

The relation of the wavefield $u(\mathbf{x}_s, \mathbf{x}, t)$ to the model parameters slowness squared s^2 and density $\rho(\mathbf{x})$ is not as trivial as in the previous case. This problem does not have a direct least squares solution. Considering that the kinematics are controlled by the model used as input to compute T_d, we can focus our attention on the density model ρ. It is important to realize that using only the homogeneous wave equation (for pressure), one cannot retrieve the magnitude of the density field because it scales all the terms in the equation. However, one should be able to retrieve the relative changes in the density field. Similarly to the constant density case, we set an inverse problem with the following objective function:

$$J\left(\frac{1}{\rho}\right) = \left|\left|\frac{s^2}{\rho}\frac{\partial^2 u}{\partial t^2} - \nabla \cdot \left(\frac{1}{\rho}\nabla u\right)\right|\right|_2. \tag{4.9}$$

Numerically, the problem is implemented as follows:

$$J(m) = ||\mathbf{D_t} m + \mathbf{D}\mathbf{G_u} m||_2 = ||L(m)||_2\,, \tag{4.10}$$

where $\mathbf{D_t} = diag\{s^2(\mathbf{x})\frac{\partial^2 u(\mathbf{x}_s,\mathbf{x},t)}{\partial t^2}\}$, $\mathbf{G_u}$ is a block diagonal matrix that contains the gradient of the pressure field u, and $\mathbf{D}$ is the matrix representation of the divergence operator.

The gradient with respect to the buoyancy $m = 1/\rho$ is given by

$$\nabla_m J = \mathbf{D_t}^\top L(m) + \mathbf{G_u}^\top \mathbf{D}^\top L(m), \tag{4.11}$$

where $^\top$ denotes adjoint. Once the gradient is defined, the model is updated iteratively using a non-linear solver.

4.3.3 Pressure and particle velocity relations

Another option for retrieving an image of the buoyancy is to use the relations between pressure fields and particle velocity. The Marchenko framework described in Singh et al. (2016) allows for the reconstruction of pressure u and the z component of the particle velocity

v_z fields. Hence, we can find the model that best matches the acoustodynamics relation (given by the first-order linear relations of Hooke's laws):

$$\frac{\partial V_z}{\partial t} = -\frac{1}{\rho}\frac{\partial u}{\partial z}. \tag{4.12}$$

From which we can setup a linear inverse problem with the following objective function:

$$J(m) = \left|\left|\frac{\partial V_z}{\partial t} + m\frac{\partial u}{\partial z}\right|\right|_2, \tag{4.13}$$

which has the following solution:

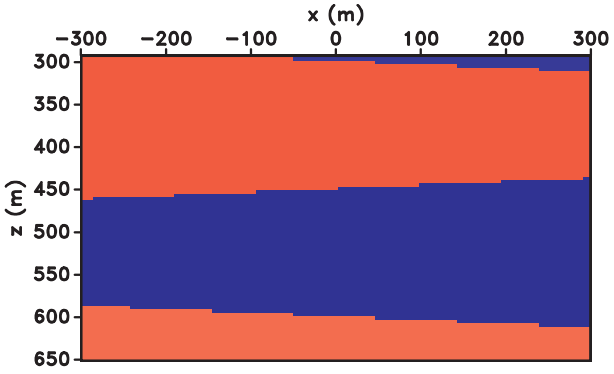
$$m = -\frac{\sum_{t,\mathbf{x}_s} \frac{\partial V_z}{\partial t}(\mathbf{x}, \mathbf{x}_s, t)\frac{\partial u}{\partial z}(\mathbf{x}, \mathbf{x}_s, t)}{\sum_{t,\mathbf{x}_s} \left(\frac{\partial u}{\partial z}(\mathbf{x}, \mathbf{x}_s, t)\right)^2}. \tag{4.14}$$

The approaches presented here are not unique and, one could try for instance, to re-parametrize the wave equation in terms of acoustic impedance or other suitable combination of acoustic parameters. However, the strategy should fall into one of the presented methods. With a wrong background velocity model, it would not be possible to decouple the effect of model contrasts from the original model. Hence, it could be convenient to use impedance as the parameter that explains the reflections in the retrieved Marchenko wavefield.

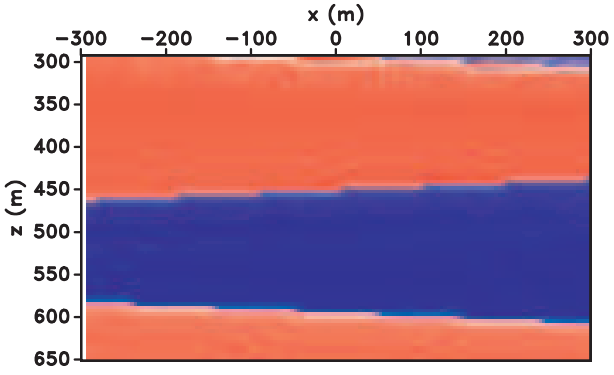
4.4 Examples

We test the imaging procedure described in the preceding section for two cases: with constant and variable density. For each case, we test the sensitivity of our technique to the background velocity model and compute acoustic data with a free surface boundary condition. In order to record both pressure and particle velocity data, we use the coupled acoustodynamics system of equations used in Thorbecke and Draganov (2011).

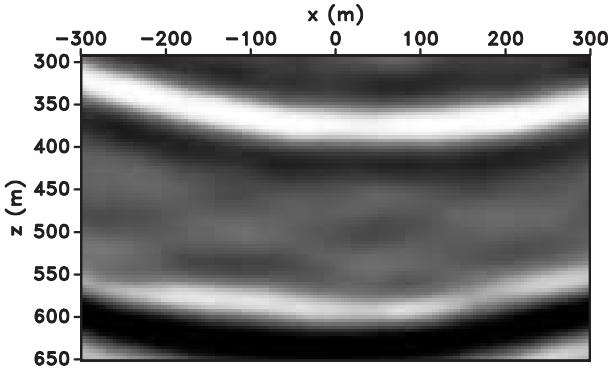
Figure 4.1 shows the velocity model and the highlighted box depicts the area where the Marchenko wavefields are retrieved. The dots in the surface indicate the source locations used for the imaging process (a subset of the sources present in the data). The model consists of four layers of alternating velocity.



(a)



(b)



(c)

Figure 4.2: Velocity retrieval: (a) is the (exact) background model used for computing the transmission response T_d, (b) is the retrieved velocity model, and (c) is a snapshot of the wavefield $G(\mathbf{x}, \mathbf{x}_s = (0,0), t = 0.6s)$. The inversion shows the sensitivity of the method to the information present in the input wavefield. The snapshot shows the downgoing primary field and the downgoing surface-related multiple.

To retrieve the Marchenko wavefields, we place 500 sources and receivers at $z = 0$m in the range $x = [-1000, +1000]$m (Figure 4.1). The input transmission data $T_d(\mathbf{x}, \mathbf{x}_s, t)$ is computed using an eikonal solver to obtain the (model-dependent) travel-times $t_d(\mathbf{x}, \mathbf{x}_s)$ followed by a convolution with a Ricker wavelet with peak frequency of 40Hz:

$$T_d(\mathbf{x}, \mathbf{x}_s, t) = \delta\left(t - t_d\left(\mathbf{x}, \mathbf{x}_s\right)\right) * w(t). \tag{4.15}$$

The Ricker wavelet is the source function in this example, but in a real case scenario one must obtain a suitable source function from the data themselves.

Figure 4.2(a) shows the background model inside the imaging box, which in this case is exact. The background model is used for computing the first arrivals (from the surface to the box locations). Figure 4.2(b) shows the retrieved velocity model (using equation 4.7), which converges to the true model in Figure 4.2(a). Figure 4.2(c) shows a wavefield snapshot at $t = 0.6$s for a source at $\mathbf{x}_s = (0, 0)$m. The retrieved Marchenko wavefield is the main input for the inversion. This result shows that the input wavefields do satisfy the homogeneous wave equation since the retrieved model converges to the input model used to create the data.

Figure 4.3(a) depicts a smooth version of the true velocity used as the background model. The Marchenko wavefields show propagating waves, with the corresponding reflections, even if the background velocity model is smooth. One could say that the retrieved wavefield are an expression of the true wavefield with the kinematics of the background model. Figure 4.3(b) is the inverted velocity model using the same procedure discussed above. Note that the inverted model still converges to the background model. This confirms that the kinematics of the retrieved wavefield are dominated by the background model, but the reflections present in the wavefield are due to the discontinuities of the original model. Figure 4.3(c) is a wavefield snapshot at the same time and location as in Figure 4.2(c) for the Marchenko wavefield propagated in the new background model which shows that the reflections from the original model propagate with the kinematics of the background velocity.

4.4.1 Variable density model

The model parameters for the variable density experiment are shown in Figures 4.4(a)-4.4(b). The density model contains two layers and its interface is decoupled from those in the velocity model. The model is designed in this fashion to test if the method is able to decouple the velocity from the density in the inversion. Since the kinematics of the wavefield converge to those of the background model, we focus our attention on the independent parameter, which is density in this case. Figure 4.5(a) shows the background model inside the box (the true model), whereas Figure 4.5(b) shows the inverted density model. The velocity interfaces leak into the inverted model, but the density layer and interface are well-imaged. Note that the density model is not an input for the computation of the wavefield shown in Figure 4.3(c) and yet the inverted image clearly delineates the density layer. In this example, we use the relation between pressure and particle velocity fields (equation 4.14) to compute the density image.

Figure 4.6(a) shows a smooth version of the true model used as background velocity field, and Figure 4.6(b) depicts the inverted density model. The density layer is clearly imaged despite the smooth background velocity model. The velocity interfaces again leak into the inverted model; interestingly, the background velocity model does not contain any interface (in a similar way as a conventional migrated image works) and yet the velocity interfaces from the original (true) model are imaged. One of the input wavefields for the inversion is shown in Figure 4.6(c). The wavefield propagates in the smooth background model (compare this with the snapshot in Figure 4.6(c)) and yet presents all the reflections from the true model.

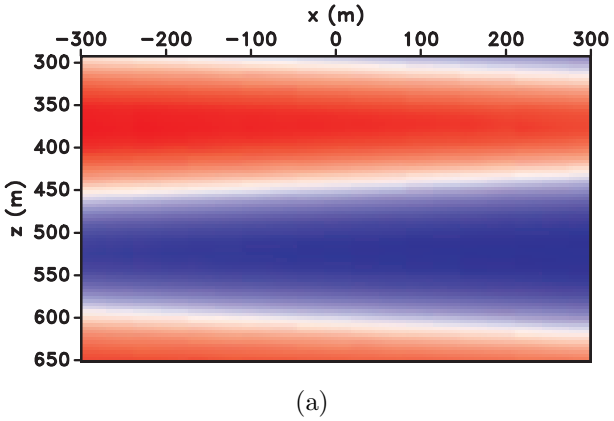
Figure 4.7 is a density inversion from the pressure field (equation 4.10). This type of inversion is more unconstrained (we have half the data). Also, by only using the pressure field, there is an ambiguity with the magnitude of the density model because many different density models can yield the same wavefield. The model is clearly imaged with this method, as the density layers are visible and there is only a hint of the velocity interfaces leaking into

the inverted model. This approach is a valid alternative if we only have the pressure data as input, instead of the pressure and particle velocity data needed for the previous variable density examples.

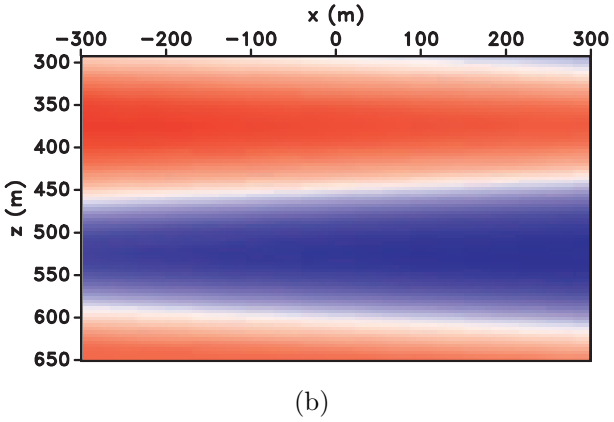
4.4.2 Gulf of Mexico inversion

We test the methodology with a 2D dataset from Gulf of Mexico (GoM). The data are acquired in the Mississippi Canyon and it has been used for multiple removal purposes (Dragoset, 1999), FWI (Guitton et al., 2012), image segmentation (Lomask and Biondi, 2005), among others. The dataset contains 683 gathers shot every 26.66m. The streamer cable has an offset range spanning from 100m to 4875m. For the processing of the data prior to Marchenko extrapolation we use the workflow from Jia et al. (2016) in which they (i) remove the surface-related multiples with SRME (Verschuur et al., 1992), (ii) remove the effect of the wavelet using the deconvolution method of (Guitton and Claerbout, 2015), and (iii) rearrange the data in split-spread configuration by means of reciprocity. Since the data contains primaries and internal multiples, the Marchenko wavefields are retrieved with the method of Broggini and Snieder (2012).

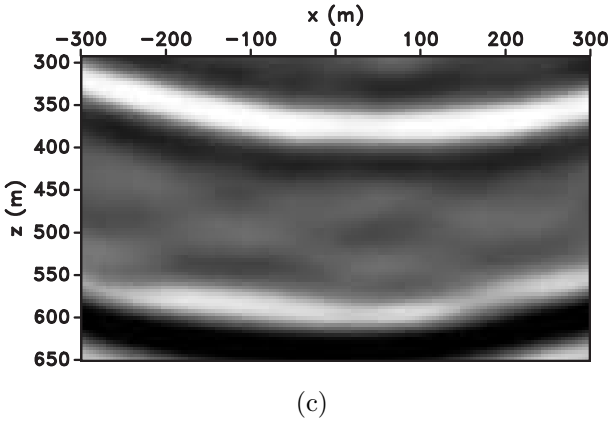
Since for this particular dataset we only have pressure data, we use the variable density acoustics framework. Figure 5.8 is the background velocity model which contains a tabular salt body. Figure 5.12(a) shows the corresponding Marchenko image inside the imaging target. The image contains strong reflectors in the sediments on the left side of the salt body. In the subsalt area, due to low illumination, some of the sediment reflectors vanish abruptly underneath the salt. Note that the boat is towed from left to right which means that is more difficult to record scattered energy from the subsalt area for those source locations placed above the salt body (since the salt body deflects a significant amount of energy). Figure 4.8(c) is the inverted model, in which we can see the imprint of the salt body and also some anomalously lower values underneath the salt. These anomalies are present in the lower illumination subsalt area. At around the same locations of the bright sediment reflector, high density anomalies are present. Deeper in the sediments area of the section



(a)

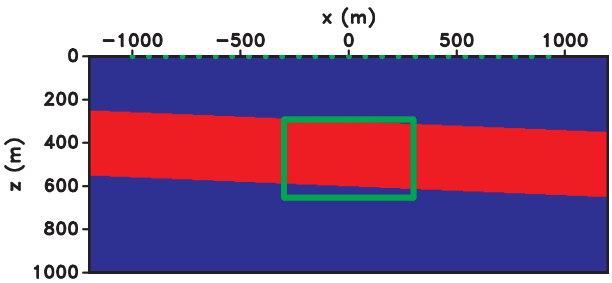


(b)

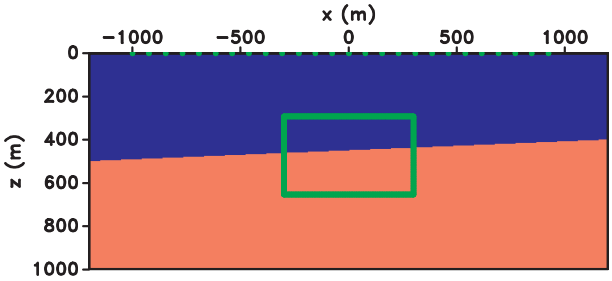


(c)

Figure 4.3: Velocity retrieval: (a) is a smooth background model used for computing the transmission response T_d, (b) the retrieved velocity model, and (c) a snapshot of the wavefield $G(\mathbf{x}, \mathbf{x}_s = (0,0), t = 0.6s)$. Note how the retrieved model converges to the background model

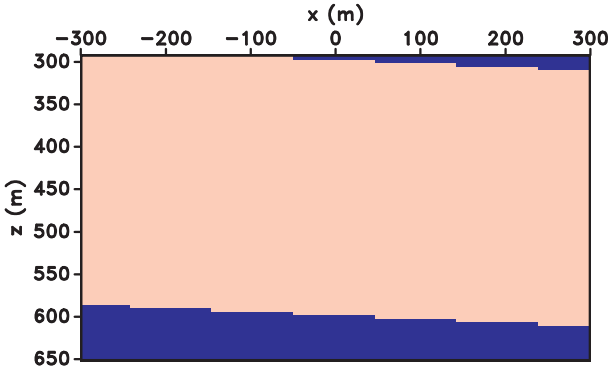


(a)

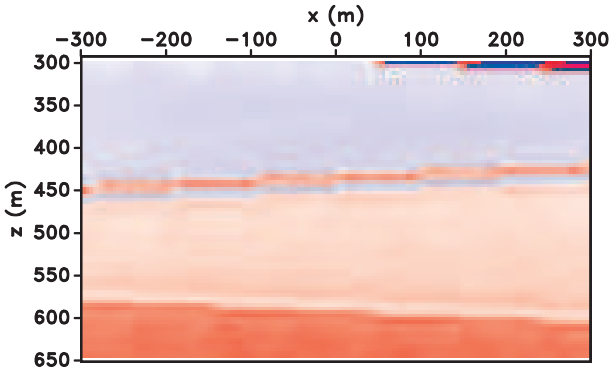


(b)

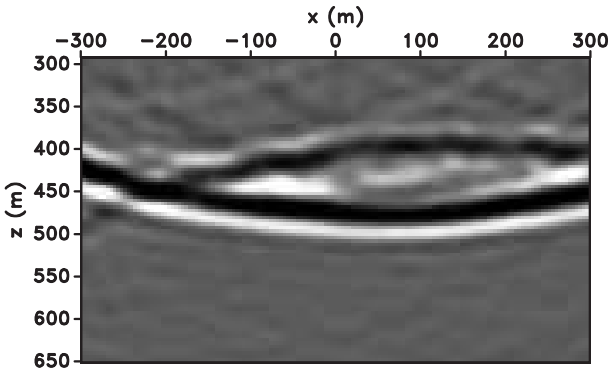
Figure 4.4: (a) Variable velocity and (b) density model. Note how the density and velocity model are uncorrelated inside the box. We construct the model this way to differentiate between velocity and density reflections during the inversion.



(a)

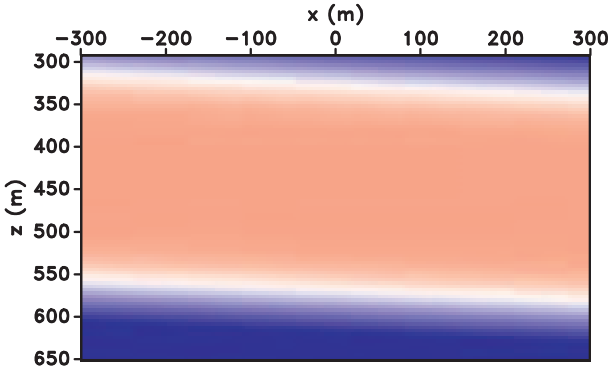


(b)

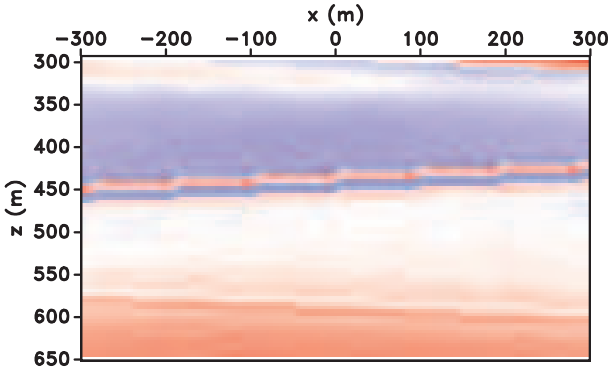


(c)

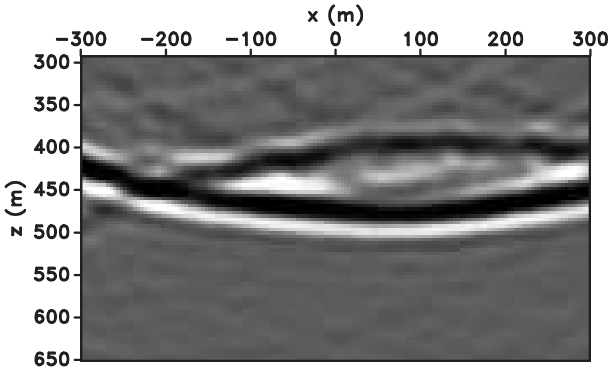
Figure 4.5: Density retrieval: (a) exact velocity model used as background, (b) inverted density model, and (c) a snapshot of the wavefield $G(\mathbf{x}, \mathbf{x}_s = (0,0), t = 0.6s)$. Note how the velocity layer leaks into the inverted model. However, the density layer and interface are clearly defined in the inverted model. Also, despite the density model not being one of the inputs for Marchenko, note how the reflection in very well defined in the wavefield snapshot (only existing above the density reflector).



(a)



(b)



(c)

Figure 4.6: Density retrieval: (a) smooth velocity model used as background, (b) inverted density model, and (c) a snapshot of the wavefield $G(\mathbf{x}, \mathbf{x}_s = (0,0), t = 0.6s)$. Note how the velocity layer leaks into the inverted model. However, the density layer and interface are clearly defined in the inverted model. The smoothness of the velocity model introduces small shifts in the wavefronts present in the snapshot.

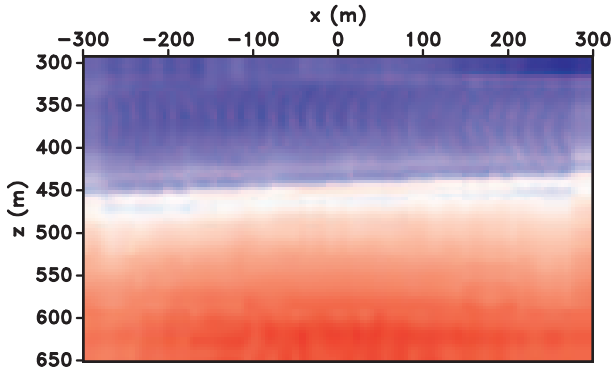


Figure 4.7: Density image retrieval using only the pressure wavefield in an iterative inversion process. Compare with Figure 4.6(b). Note how the leakage of the original velocity interfaces is attenuated (below 550m). The inversion is an image of the background density model.

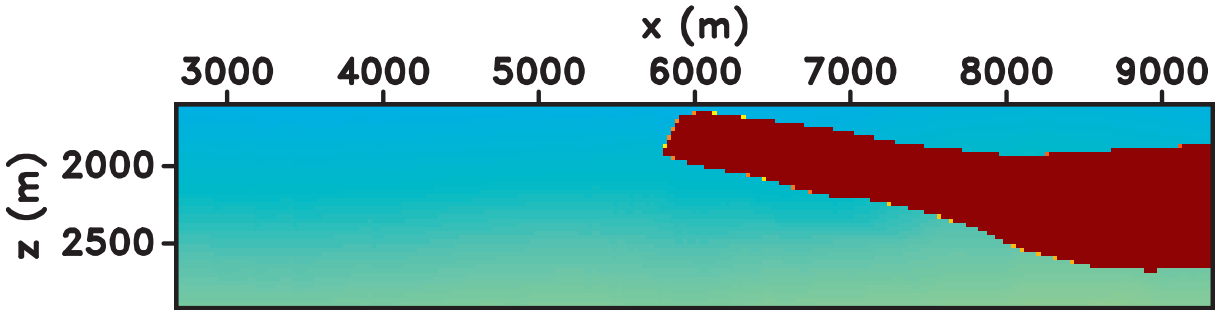
(below $z = 2200$m), the contrast is not very high (as shown by the seismic events amplitude in Figure 5.12(a)); hence, the inverted model (Figure 4.8(c)) is smooth in that area.

4.5 Conclusions

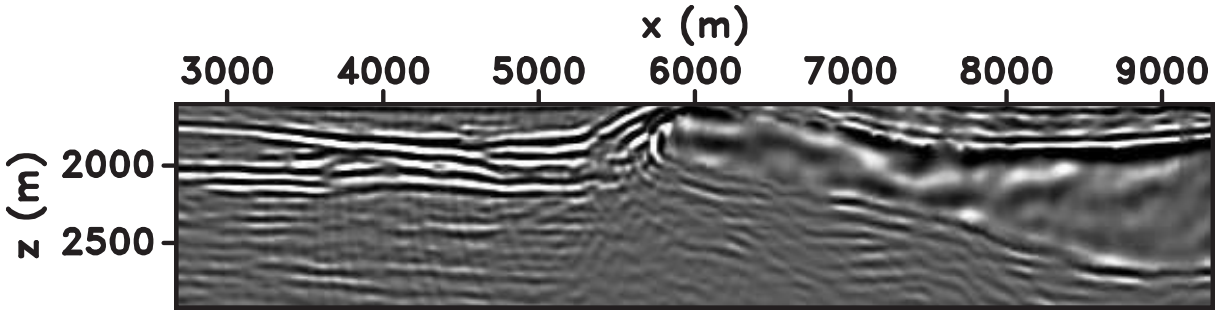
We present a framework to obtain images of acoustic model parameters that are different from reflectivity. Our method relies on two-way wavefields computed with Marchenko modeling. The wavefields are used with the wave equation together with the background velocity to find the model that explains, in a least-squares sense, the homogeneous wave equation. We show different options one could use depending on which type of data are available. One can exploit the relations between pressure and particle velocity and invert for the density model, or use an iterative approach if pressure is the only available data. Our tests show that it is possible to find an image representing the true model even when the background model is smooth. Our method is able to image the properties within the subsurface instead of the properties contrasts as it is done with conventional images.

4.6 Acknowledgments

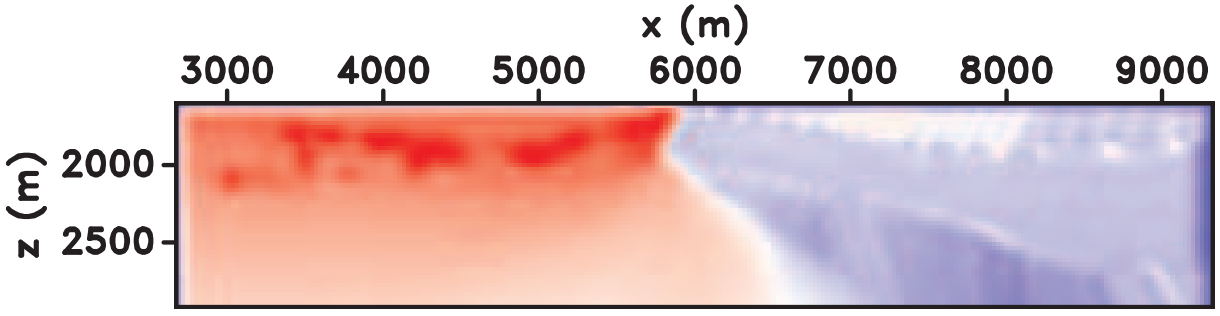
We thank the sponsors of the Center for Wave Phenomena, whose support made this research possible. The reproducible numeric examples in this paper use the Madagascar open-



(a)



(b)



(c)

Figure 4.8: Gulf of Mexico inversion: (a) background velocity model, (b) Marchenko image, and (c) inverted model. Note the strong contrast between the sediments and the salt body in the inverted image (i.e., at around $x = 5900$m).

source software package (Fomel et al., 2013) freely available from http://www.ahay.org. Esteban would like to thank Alex Jia (CWP) for providing his scripts to obtain the Marchenko wavefields for the GoM dataset.

CHAPTER 5

EXTENDED IMAGING, DECONVOLUTION, AND TWO-WAY WAVEFIELDS: A COMPARISON[8]

Esteban Díaz, Satyan Singh, and Paul Sava

Center for Wave Phenomena, Colorado School of Mines

Multiple-scattered waves contain information that is commonly disregarded during imaging and tomography. Marchenko wavefields are superior to time-reverse wavefields because of their ability to handle primaries together with internal and surface-related multiples. Using all types of waves for imaging can greatly improve the illumination and augment the sensitivity of the data to errors in the background velocity model. We compare extended images computed with reverse time and Marchenko wavefields, and investigate the potential of using multidimensional deconvolution for extended images in order to obtain higher image resolution. Our experiments show that the Marchenko wavefields are sensitive to errors in the background model in a way that is similar to the sensitivity of time-reverse wavefields. The main difference between these imaging strategies is the improved angle illumination with the Marchenko wavefields due to the correct use of multiples; this improvement can reduce the bias of tomography operators towards lower velocities when the data are contaminated with multiples.

5.1 Introduction

The quality of the seismic image depends on the choice of wave propagator, but more importantly on the quality of the background model. The better the velocity, the more focused the seismic image becomes. One can update the background velocity model in such a way that the image quality improves. By increasing the seismic image focusing, one

[8]IN PREPARATION FOR GEOPHYSICS

effectively improves the synchronization of the source and receiver wavefields at the image locations. Image focusing can be observed in zero-offset sections by measuring the reflector continuity and image strength, or by analyzing the focusing of point diffractors. However, if the events in the image consist mostly of continuous reflectors (layers inside the earth), the assessment of the background model accuracy becomes difficult. In order to retrieve the information about the kinematic error, one must extend the image away from zero offset, which facilitates interpretation of the image at different illumination directions. In the context of two-way operators, the extension is usually performed by extended correlation in the image domain (Rickett and Sava, 2002; Sava and Fomel, 2006; Sava and Vasconcelos, 2011b), in the angle domain (Sava and Fomel, 2003; Yoon et al., 2004; Jin et al., 2014; Yoon et al., 2011; Vyas et al., 2011), or in the surface offset domain (Giboli et al., 2012). Once a suitable extension is chosen, the velocity error information can be extracted (1) by measuring focusing in the extended gathers (Shen and Symes, 2008; Weibull and Arntsen, 2013a; Yang et al., 2013; Shan and Wang, 2013; Biondi and Almomin, 2014; Díaz and Sava, 2015; Yang and Sava, 2015), (2) by measuring moveout in the angle or surface offset gathers (Liu and Han, 2010; Shen et al., 2015; Fleury et al., 2014) or, (3) by measuring the consistency between individual shot record images (Perrone et al., 2015).

Another important factor that determines the quality of a seismic image is the amount of data illuminating a target in the subsurface. Great effort is invested in wider acquisition aperture to increase the illumination of a seismic target. Most of the current imaging (and modeling) technologies utilize the primaries in the data but disregard the multiples. By utilizing the multiples one can better illuminate the targets with smaller acquisition aperture (Verschuur and Berkhout, 2011). In order to incorporate the multiples into the seismic images, a conventional approach is to use the primaries as an areal source for the surface-related multiples; this allows one to image the first order surface-related multiples in the same position as the primaries (Guitton, 2002; Grion et al., 2007; Verschuur and Berkhout, 2011; Whitmore et al., 2010; Wong et al., 2015). However, these techniques require prior separation

of multiples and primaries. A recent alternative to incorporate the multiples into the imaging process is the Marchenko Green's function retrieval framework in which the multiples are used together with the primaries in a global prediction procedure to find focusing solutions that, when injected into the surface, produce wavefield focusing at an arbitrary location in the subsurface (Broggini and Snieder, 2012; Behura et al., 2014; Wapenaar et al., 2014b; Singh et al., 2015a). The focusing solutions are guaranteed to exist regardless of the background velocity model. Once the focusing functions are found the up- and down-going Green's functions are computed through convolutions and correlations between the focusing solutions and the surface data. The method was initially developed to handle internal multiples (Broggini and Snieder, 2012; Behura et al., 2014; Wapenaar et al., 2014b) and was later extended to incorporate surface-related multiples (Singh et al., 2015a) .

We show here the sensitivity of Marchenko wavefields to the background model which is similar to that of conventional one-way or RTM wavefields. We assess this sensitivity by using extended images and compare the moveout with conventional RTM extended images. We pose the extended imaging process as an inversion problem which can be thought of as an extended deconvolution imaging condition. Assessing the sensitivity to errors in the background is the first step towards the formulation of a tomographic problem. Furthermore, Marchenko wavefields can eliminate the need for the separation of primaries and multiples in tomography, and increase the illumination of the subsurface targets.

This paper is organized as follows: first, we compare conventional two-way RTM operators and Green's function retrieval with the Marchenko process; next, we review the extended imaging and deconvolution process; and finally, we show examples comparing the sensitivity of RTM and Marchenko modeling approaches to the background model.

5.2 Modeling differences

In this section, we review similarities and differences between Marchenko-derived and RTM propagators. While both propagators are two-way, Marchenko separates the total

wavefield u into up- and down-going components:

$$u(\mathbf{x}, \mathbf{x}_s, t) = u_s(\mathbf{x}, \mathbf{x}_s, t) + u_r(\mathbf{x}, \mathbf{x}_s, t), \tag{5.1}$$

where t is time, and $\mathbf{x}_s$ and $\mathbf{x}$ are source and space coordinates, respectively.

This separation is necessary because the method relies on one-way reciprocity theorems (Wapenaar et al., 2014a). Singh et al. (2015a) show how to recover the total wavefield u along with its directional components u_s and u_r using a free-surface condition which allows to correctly model primaries, internal multiples, and surface-related multiples.

The algorithm in Singh et al. (2015a) finds the focusing solutions $f_1^+(\mathbf{x}_s, \mathbf{x}, t)$ and $f_1^-(\mathbf{x}_s, \mathbf{x}, t)$, which produce a focus at an arbitrary location $\mathbf{x}$ in the subsurface by injection at the surface at coordinates $\mathbf{x}_s$. The focusing solutions always exist regardless of the background model. In order to obtain the focusing solutions, the method utilizes the recorded reflectivity $R(\mathbf{x}_s, \mathbf{x}_r, t)$ together with an estimate of the direct arrival $T_d(\mathbf{x}, \mathbf{x}_r, t)$, which goes from the subsurface point $\mathbf{x}$ to the surface locations $\mathbf{x}_r$. Embedded in the direct arrival information, T_d is the kinematic information from the background model. If the direct arrival information is accurate, the recovered wavefield satisfies the wave equation

$$\frac{1}{v^2}\partial_t^2 u - \rho\nabla\cdot\left(\frac{1}{\rho}\nabla u\right) = \rho\partial_t\delta(\mathbf{x} - \mathbf{x}_s, t) \tag{5.2}$$

for a volumetric injection source $\rho\partial_t\delta(\mathbf{x} - \mathbf{x}_s, t)$, density ρ, and velocity v from the true medium. If the direct arrival information T_d is obtained from an incorrect background velocity model, the retrieved wavefield is a representation of the true Green's function with the kinematics of the background model.

Once the focusing functions are computed through the iterative scheme described in Singh et al. (2015a), the individual wavefield components are obtained through the following equations (Singh et al., 2015a):

$$u_r(\mathbf{x},\mathbf{x}_r,t) = -f_1^-(\mathbf{x}_r,\mathbf{x},t) + \sum_{\mathbf{x}_s} f_1^+(\mathbf{x}_s,\mathbf{x},t) * R(\mathbf{x}_r,\mathbf{x}_s,t) - r f_1^-(\mathbf{x}_s,\mathbf{x},t) * R(\mathbf{x}_r,\mathbf{x}_s,t), \tag{5.3}$$

$$u_s(\mathbf{x},\mathbf{x}_r,t) = +f_1^+(\mathbf{x}_r,\mathbf{x},T-t) - \sum_{\mathbf{x}_s} f_1^-(\mathbf{x}_s,\mathbf{x},t-T) * R(\mathbf{x}_r,\mathbf{x}_s,t) - r f_1^+(\mathbf{x}_s,\mathbf{x},t) * R(\mathbf{x}_r,\mathbf{x}_s,t). \tag{5.4}$$

Note that the inputs for computing either wavefield are the same: the focusing solutions f_1^+ and f_1^-, the data at the surface R, and the reflection coefficient at the free surface $r = -1$. The focusing solutions f_1^- and f_1^+ can be thought of as propagators that redatum the data from the source position $\mathbf{x}_s$ to the image point $\mathbf{x}$.

In contrast, the RTM propagators obtain the source and receiver wavefields through two independent modeling problems:

$$L(v)u_s = f_s, \tag{5.5}$$

which is solved forward in time for the source function f_s, and

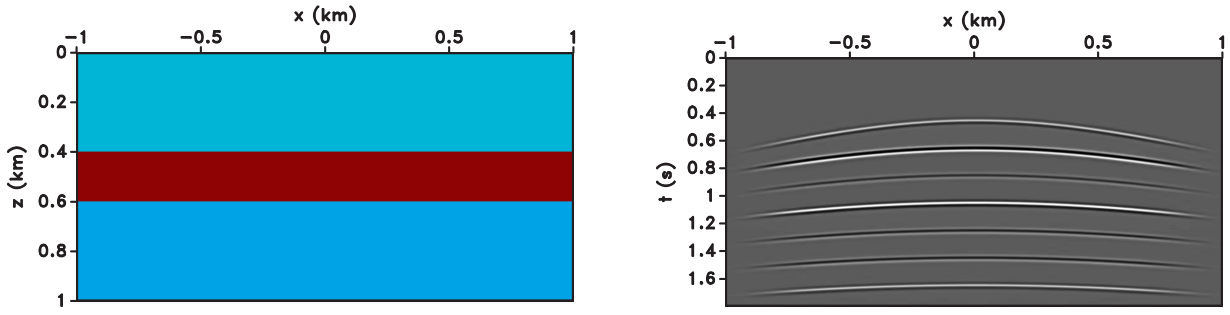
$$L^\top(v)u_r = f_r, \tag{5.6}$$

which is solved backward in time for the receiver data $f_r = \sum_{\mathbf{x}_r} R(\mathbf{x}_s,\mathbf{x}_r,t)\delta(\mathbf{x}-\mathbf{x}_r)$. Note that for the RTM propagators, the sum of the source and receiver wavefields does not satisfy equation 5.2 and does not correctly handle the multiples present in the data.

5.3 Sensitivity to the background model

In the Marchenko context, the wavefields u_r and u_s represent the up- and down-going components of the Green's function, respectively. Following the imaging principle (Claerbout, 1971) the image $I(\mathbf{x})$ is constructed when wavefields u_s and u_r coincide in space and time:

$$I(\mathbf{x}) = \sum_{\mathbf{x}_s,t} u_s(\mathbf{x}_s,\mathbf{x},t)u_r(\mathbf{x}_s,\mathbf{x},t). \tag{5.7}$$



(a) (b)

Figure 5.1: Free surface experiment with a constant velocity background: (a) variable density model and (b) the recorded data for a shot at $\mathbf{x}_s = (0, 0)$km. Note how the data contain primaries (first two arrivals), internal multiples (weak third arrival), and the corresponding surface-related multiples.

Although $I(\mathbf{x})$ is helpful for understanding the subsurface, it does not contain much information about the quality of the background velocity model. Instead, one can use the extended imaging condition (Rickett and Sava, 2002; Sava and Fomel, 2003; Sava and Vasconcelos, 2011b) to analyze the interaction between source and receiver wavefields at the vicinity of the image location $\mathbf{x}$ and $t = 0$:

$$I(\mathbf{x}, \boldsymbol{\lambda}, \tau) = \sum_{\mathbf{x}_s, t} u_s(e, \mathbf{x} - \boldsymbol{\lambda}, t - \tau) u_r(e, \mathbf{x} + \boldsymbol{\lambda}, t + \tau), \tag{5.8}$$

where $\boldsymbol{\lambda}$ and τ are the space- and time-lags of the correlation, respectively. The extended-imaging operator is linear and represents the down-going field u_s interacting with the up-going field u_r which in matrix-vector can be written as

$$I = \mathbf{U_s}^\top u_r, \tag{5.9}$$

where the matrix $\mathbf{U_s}^\top$ represents the extended-imaging operator applied to the receiver field u_r. Alternatively, one can think about the (extended) imaging process as an inversion problem such that $I(\mathbf{x}, \boldsymbol{\lambda}, \tau)$ satisfies the relation

$$\mathbf{U_s} I \approx u_r, \tag{5.10}$$

where the operator $\mathbf{U_s}$ implements

$$\widetilde{u_r}(\mathbf{x}_s, \mathbf{x}, t) = \sum_{\tau,\boldsymbol{\lambda}} I(\mathbf{x} - \boldsymbol{\lambda}, \boldsymbol{\lambda}, \tau) u_s(\mathbf{x}_s, \mathbf{x} - 2\boldsymbol{\lambda}, t + 2\tau). \tag{5.11}$$

The wavefield $\widetilde{u_r}$ denotes the action of $\mathbf{U_s}$ applied to I. The relation 5.10 implies that the extended reflectivity $I(\mathbf{x}, \boldsymbol{\lambda}, \tau)$ is the deconvolution between the down- and up-going fields at the surrounding locations of image point $\mathbf{x}$ and around time $t = 0$. In order to stabilize the inversion, one can also add a damping goal by trying to keep the extended reflectivity as small as possible:

$$\epsilon I \approx 0. \tag{5.12}$$

Both fitting goals can be casted as an inverse linear problem, where one minimizes

$$J = ||\mathbf{U_s} I - u_r|| + \epsilon \, ||I|| \,, \tag{5.13}$$

which can be solved through its equivalent normal equations system:

$$\left(\mathbf{U_s}^\top \mathbf{U_s} + \epsilon^2 \mathbf{I}\right) I = \mathbf{U_s}^\top u_r, \tag{5.14}$$

and can be solved with a linear conjugate gradient algorithm. Note that the matrix $\mathbf{U_s}^\top \mathbf{U_s}$ does not need to be formed; instead one needs to compute the action of the matrix to the input reflectivity $I(\mathbf{x}, \boldsymbol{\lambda}, \tau)$. The deconvolution process is similar to the one discussed by Valenciano et al. (2009) and is also referred to as multi-dimensional deconvolution (MDD) by van der Neut et al. (2011). Both, extended correlation and deconvolution images contain kinematic information about the quality of the background model. However, the latter is a better approximation of the reflectivity and is better focused.

5.4 Examples

In this section, we compare the sensitivity of RTM and Marchenko wavefields to the background velocity model. We use a simple 1D model with a constant background velocity and a variable density profile depicted in Figure 5.1(a). The pressure data are generated with a free-surface boundary condition shown in Figure 5.1(b). The main objective of this

exercise is to investigate the sensitivity of the Marchenko images to errors in the background model, which is the first step towards a formulation of a tomography problem.

We use the subset of the extended image with $\tau = 0$, i.e. space-lag gathers, at $x = 0$ for a slow, correct, and fast constant background velocity. Figures 5.2(a)-5.2(c) show the gathers computed using the conventional RTM propagator, where one can observe the strong response of the multiples present in the data. Among all of the reflectors present in the gather, only the first two correspond to the true interfaces; the rest correspond to misplaced surface-related and internal multiples events. Figures 5.2(d)-5.2(f) show the corresponding common angle gathers for slow, correct, and fast models, respectively. The transformation from space-lag $\boldsymbol{\lambda}$ to angle θ is described in Sava and Fomel (2006). For this particular model, the velocity is constant; hence, the reflection coefficient should remain constant as a function of the scattering angle. We recompute the same set of images with the deconvolution framework in equation 5.14: Figures 5.3(a)-5.3(c) show the deconvolved space-lag gathers with RTM wavefields, and Figures 5.3(d)-5.3(f) show the corresponding angle gathers with the RTM wavefields. In these figures, the amplitude across angles is more even. The deconvolved space-lag gathers contain more energy at larger lags; this is beneficial for improved tomographic updates because during tomography, the objective function measures the energy away from $\boldsymbol{\lambda} = 0$ (Shen and Symes, 2008):

$$J = ||\boldsymbol{\lambda} I(\mathbf{x}, \boldsymbol{\lambda})||_2 \,. \tag{5.15}$$

Hence, by using deconvolution, one can better highlight aspects of the gathers that need improvement through changes to the velocity.

We repeat the same experiment using the Marchenko wavefields. Figures 5.4(a)-5.4(c) show the space-lag gathers for low, correct, and fast background velocity, respectively. The main difference between these gathers and those in Figures 5.2(a)-5.2(c) is the correct handling of the surface-related and internal multiples. Despite the characteristics of the method, where the focusing solutions f_1^+ and f_1^- always exist, the gathers present moveout, indicating errors in the background model. However, both primaries and multiples correctly collapse

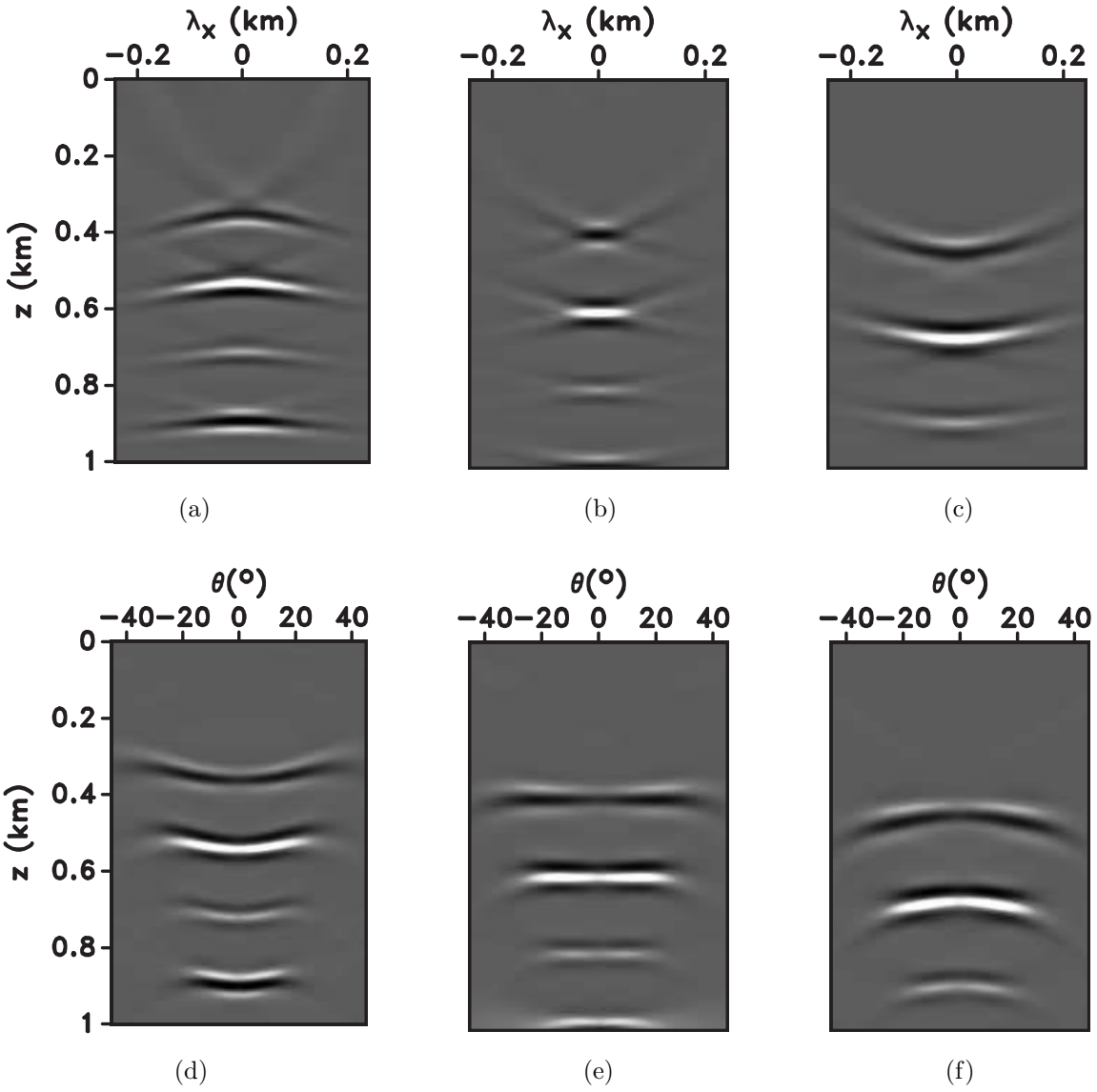


Figure 5.2: The top row shows space-lag gathers obtained from RTM wavefields for (a) low, (b) correct, and (c) fast velocity. The bottom row shows the corresponding angle gathers for (d) low, (e) correct, and (f) fast velocity.

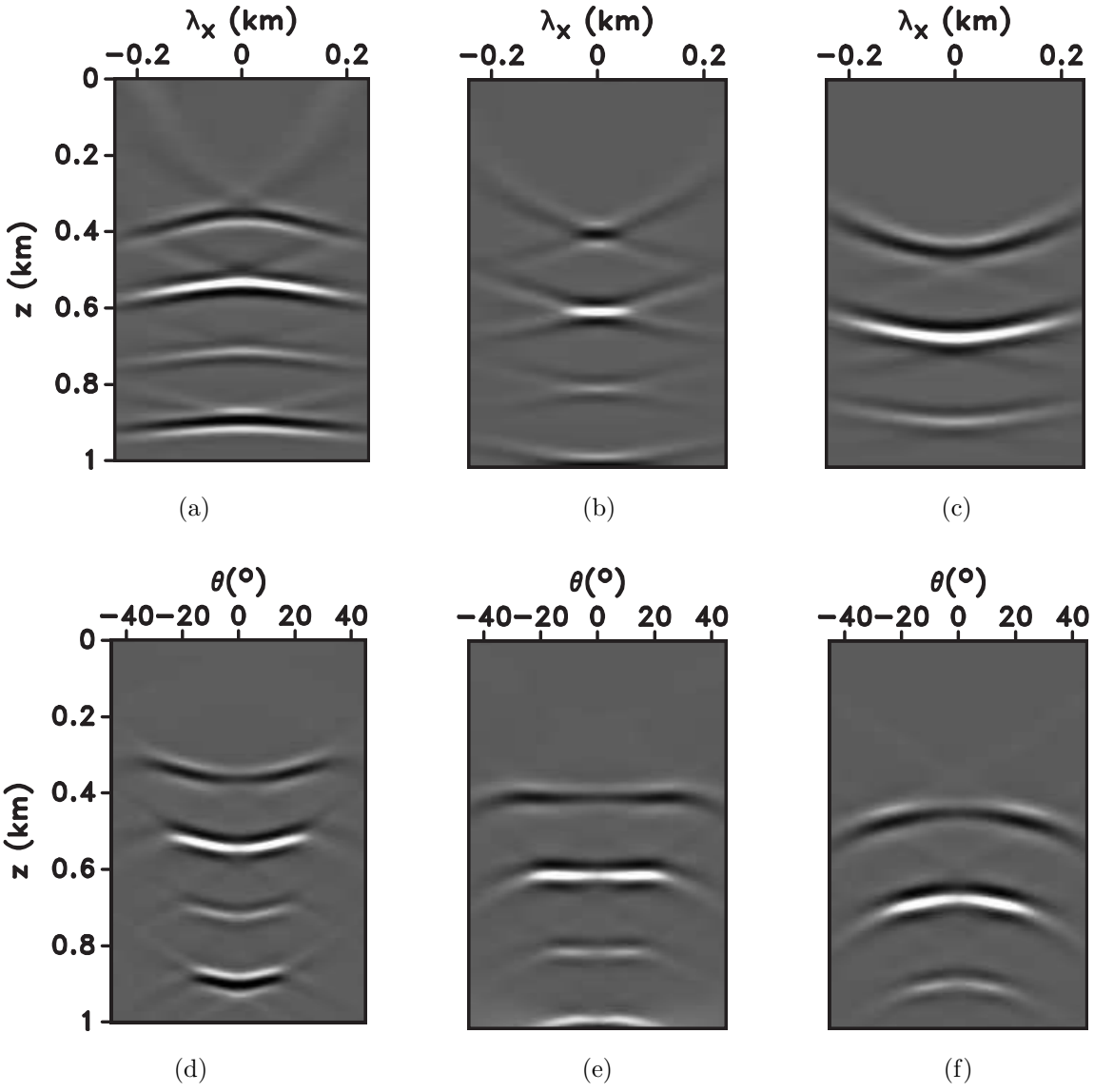


Figure 5.3: The top row shows deconvolved space-lag gathers obtained from RTM wavefields for (a) low, (b) correct, and (c) fast velocity. The bottom row shows the corresponding angle gathers for (d) low, (e) correct, and (f) fast velocity. Note how with the RTM propagator the multiples are misplaced in the gathers.

at the reflector position regardless of the errors in the background model. Also, the focusing for the correct velocity model, Figure 5.4(b), is improved compared to Figure 5.2(b). This enhancement in focusing is due to the improved illumination achieved by incorporating surface-related and internal multiples in the final image, which implies that one effectively increases the aperture with depth. In the conventional imaging, we see misplaced events due to multiples; similarly, with Marchenko wavefields, we can identify events above the true interfaces. Such events are caused by the interaction of multiples in the source wavefield with primaries in the receiver wavefield. However, these events are weak as shown in the Marchenko gathers (Singh and Snieder, 2016). The angle gathers from Figures 5.4(d)-5.4(f) contain a stronger response in the near angles, which shows the added illumination from internal and surface-related multiples. After deconvolving the Marchenko gathers (Figures 5.5(a)-5.5(c)) , the focus greatly improves, which can be better appreciated with the correct velocity case, Figure 5.5(b). We also see the same pattern as with Figures 5.3(a)-5.3(c) where the energy spreads more evenly over space-lags. The angle gathers (Figures 5.5(d)-5.5(f)) show a better amplitude distribution with angle. Also the angle range illuminated with the Marchenko fields remains somewhat constant with depth, whereas the angle gathers with conventional fields decrease the angle illumination with depth. This difference can have great impact on tomography where, with conventional propagators, the uncertainty increases rapidly in poorly illuminated areas.

5.4.1 Sigsbee model

In this example, we use a subset of the Sigsbee model (Paffenholz et al., 2002) from the sediment section, shown in Figure 5.8. The purpose for this example is to use a more challenging setting with more reflectors and lateral velocity variations. Using a free-surface boundary condition, we generate shot gathers at the surface with an Ormsby wavelet with a high-cut of 50Hz. We compute extended images from Marchenko wavefields for low, correct, and high velocity models that are obtained by scaling the correct model by 0.9, 1.0, and 1.1, respectively. For the three models, we compute conventional and deconvolved extended

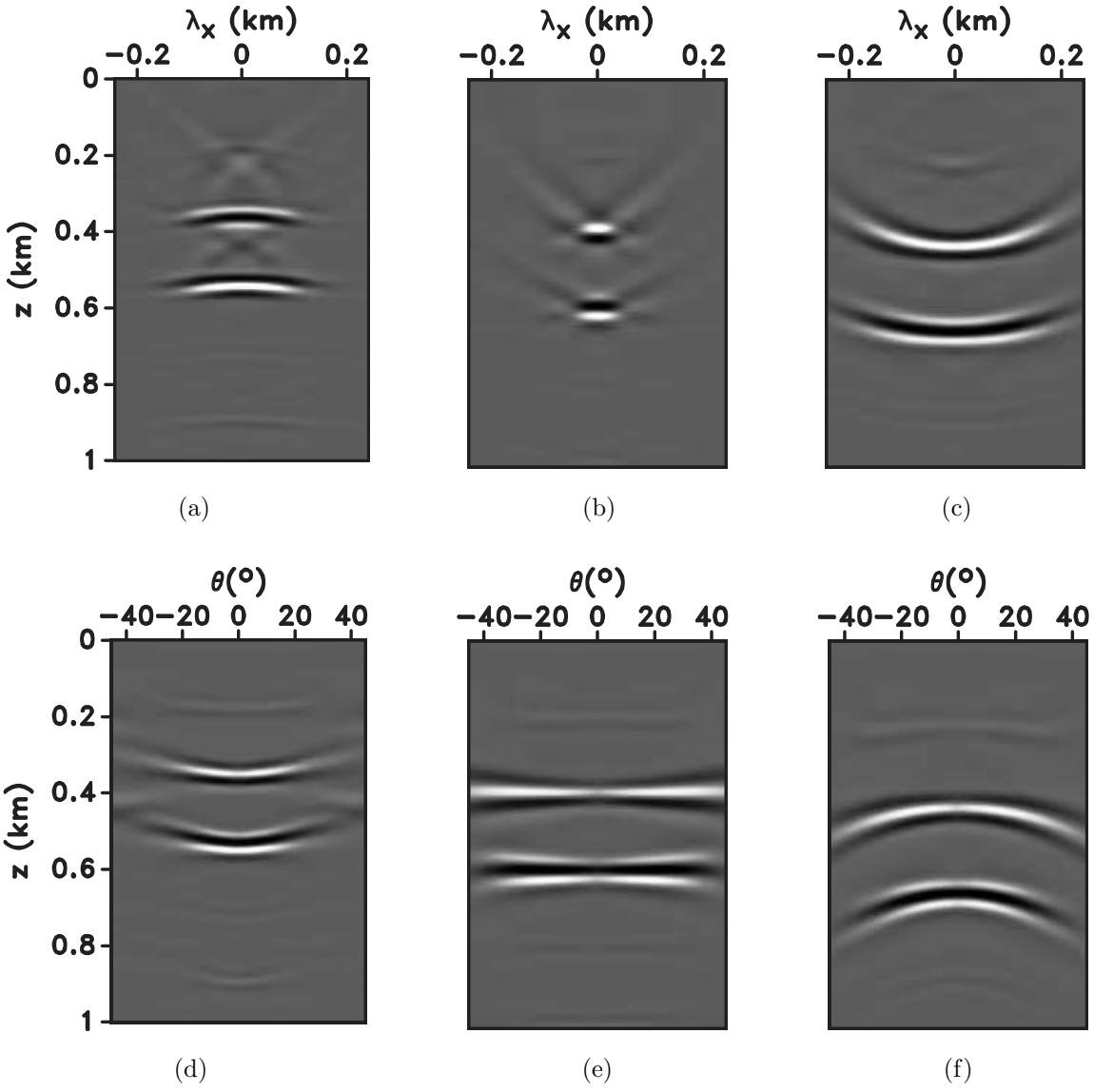


Figure 5.4: The top row shows space-lag gathers obtained from Marchenko wavefields for (a) low, (b) correct, and (c) fast velocity. The bottom row shows the corresponding angle gathers for (d) low, (e) correct, and (f) fast velocity. Note that the multiples have been imaged together with the primaries at the correct location.

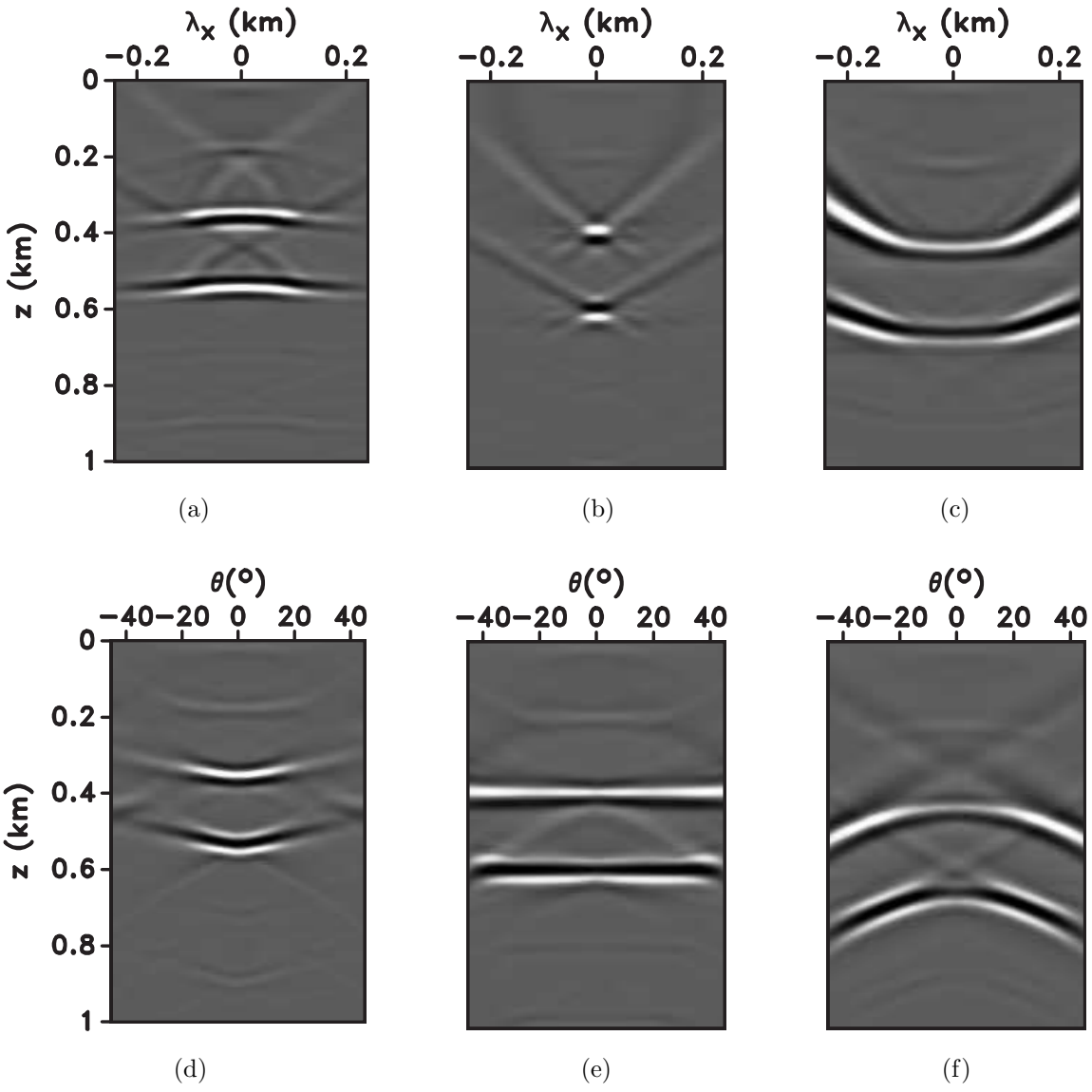


Figure 5.5: The top row shows deconvolved space-lag gathers obtained from Marchenko wavefields for (a) low, (b) correct, and (c) fast velocity. The bottom row shows the corresponding deconvolved angle gathers for (d) low, (e) correct, and (f) fast velocity. Note how the amplitude distribution is more homogeneous for the deconvolved angle gathers, also the focus for the correct velocity is much better than Figure 5.4(b).

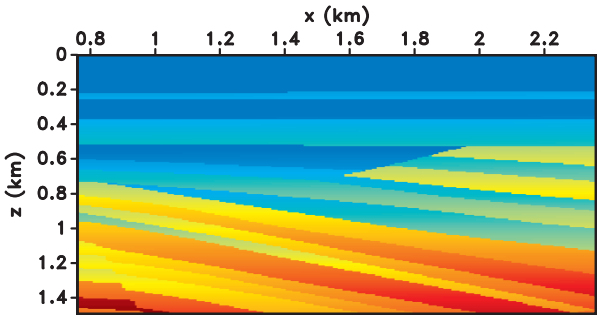


Figure 5.6: Subset from the sediment area of the Sigsbee model.

images; for each case, we show the corresponding angle gathers along the section. The angle range for each gather spans from $\theta = [-45°, 45°]$.

Figures 5.7(a), 5.7(c), and 5.7(e) depict the angle gathers from the Marchenko wavefields for low, correct, and fast velocity models, respectively. Note that the angle gathers do not contain visible artifacts from multiples in the data since there is absense of conflicting events with different moveout. Similar to the simple example from the preceding section, the angle gathers show moveout that is indicative of the errors in the background model, which is also the case for conventional RTM gathers. Figures 5.7(b), 5.7(d), and 5.7(f) show the corresponding angle gathers after the extended deconvolution process. The improved angle illumination can be seen when the amplitude is better balanced across angles, similar to the previous example. An area of such improvement is at around $z = 0.5$km, where the illumination increases at near and mid angles.

5.4.2 Gulf of Mexico example

This example features a Gulf of Mexico (GoM) dataset from the Mississippi Canyon. It has been used for multiple elimination benchmarking purposes (Dragoset, 1999) and for a full waveform inversion (Guitton et al., 2012). The subset used in this example contains 683 shots spaced every 26.67m. The streamer cable contains an offset range from 100m to 4875m. Jia et al. (2016) prepares the dataset for Marchenko Green's function retrieval with internal multiples (Broggini and Snieder, 2012). The steps in their workflow include surface-related

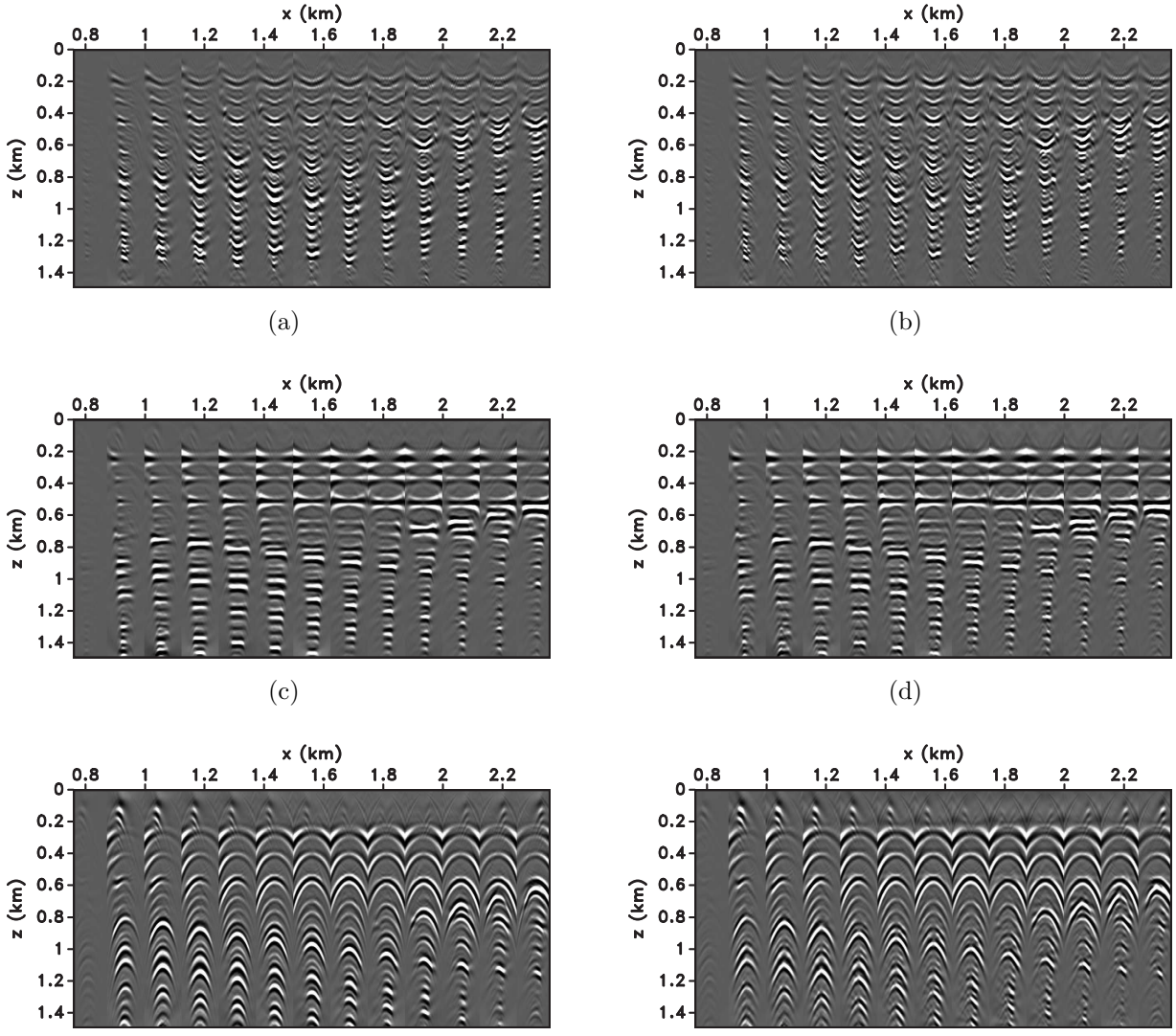


Figure 5.7: Retrieved Marchenko angle gathers obtained after mapping space-lag gathers to the angle domain. The left column shows the angle gathers transformed from extended images for (a) slow, (c) correct, and (e) fast model. The right column shows the angle gathers transformed from the deconvolved space-lag gathers for (b) slow, (d) correct, and (f) fast model. Note the improved response after the deconvolution process at near angles.

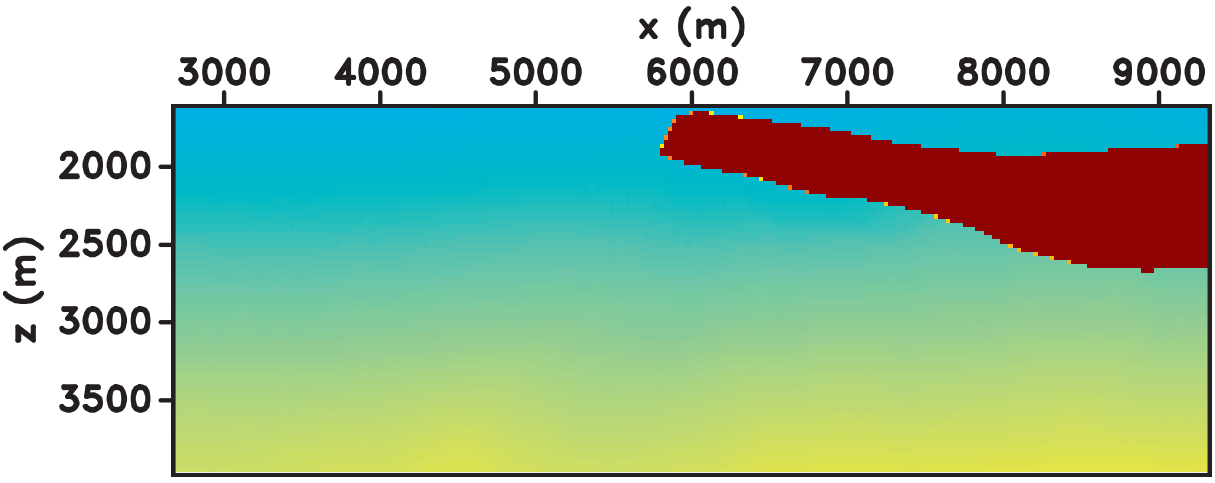
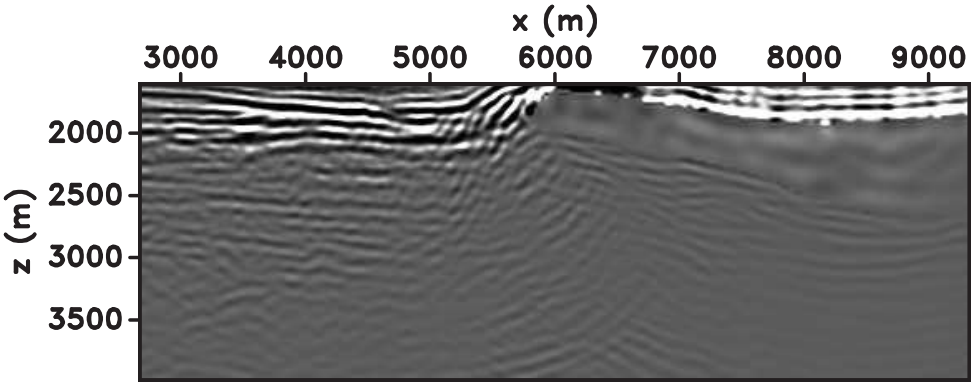


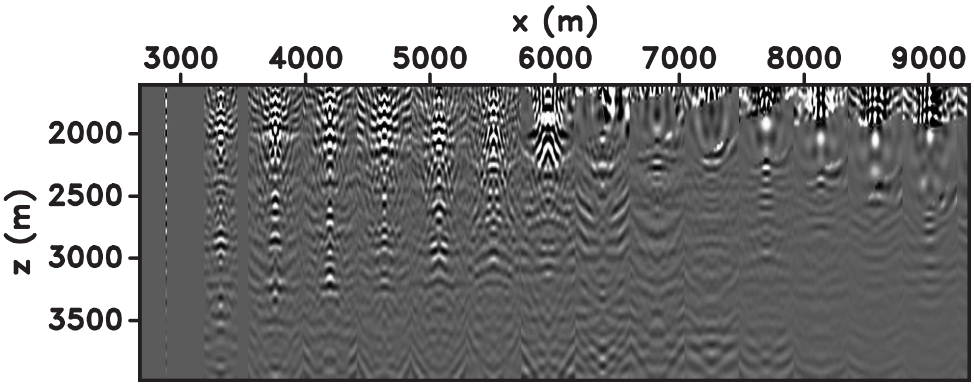
Figure 5.8: GoM velocity model obtained after traveltime tomography followed by salt flooding, interpretation of bottom of the salt, and traveltime tomography in the subsalt area.

multiple elimination (Verschuur et al., 1992), deconvolution (Guitton and Claerbout, 2015), and rearrangement of the data using reciprocity. We use their methodology to retrieve the Green's function in a target area of the line to the left of a salt body, featuring sediment and subsalt environments.

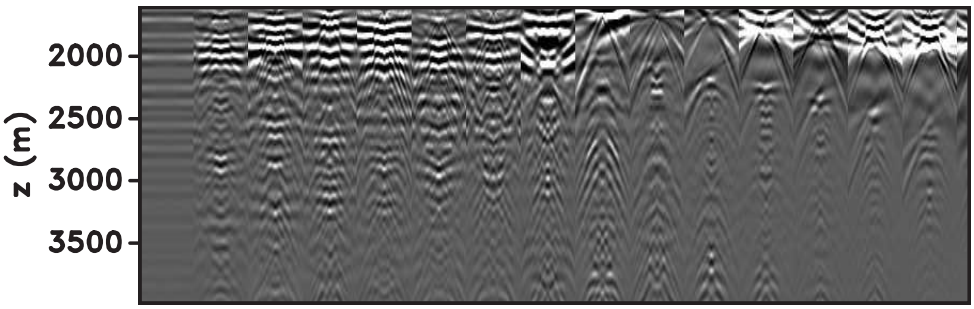
Figure 5.8 shows the velocity model used to retrieve the wavefields. Note that the tabular salt body together with the limited acquisition offset introduces illumination problems in the migrated images. Figure 5.9(a) depicts the RTM image in the target area, whereas Figures 5.9(b)-5.9(c) show the space-lag gathers and angle gathers, respectively. Note how the most coherent events in the target are sediments to left of the salt body. The space-lag gathers show relatively good focusing (i.e. most of the energy is concentrated around zero lag). Hence, most of the events in the angle gathers are flatter. However, the angle gathers show important illumination differences: in the sub-salt area the effective angle range is much narrower compared with the angle range of sediment events. After applying the extended deconvolution process, we obtain the RTM image shown in Figure 5.10(a). Overall, the images before (Figure 5.9(a)) and after deconvolution are very similar. An apparent difference is the energy balance in the image: in the sub-salt area, the correlation



(a)

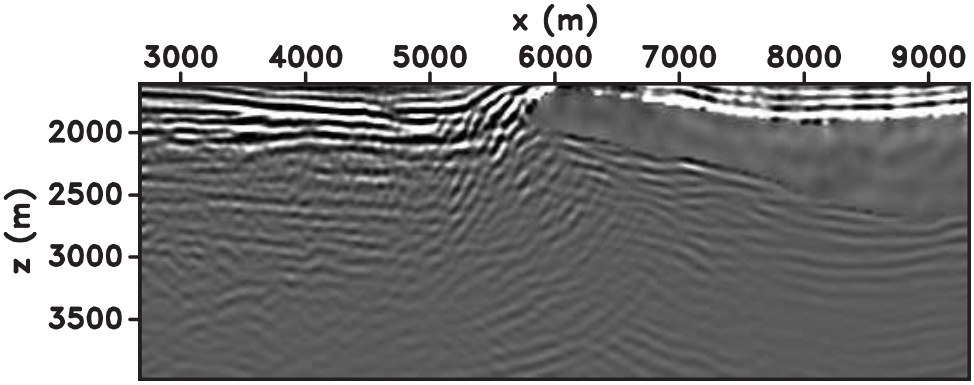


(b)

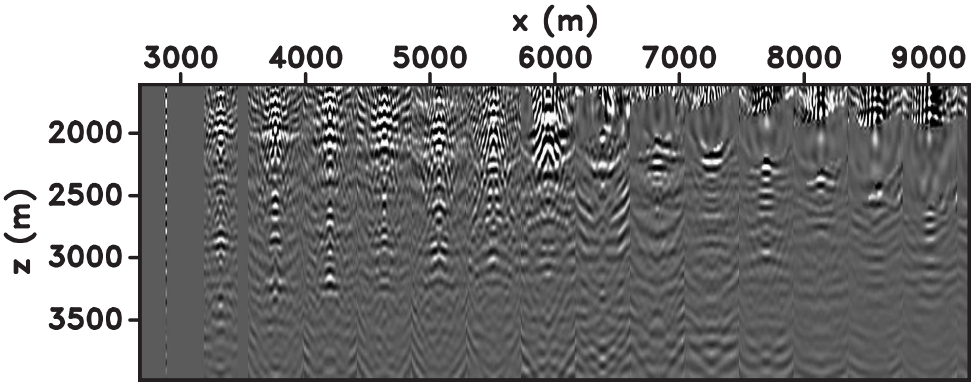


(c)

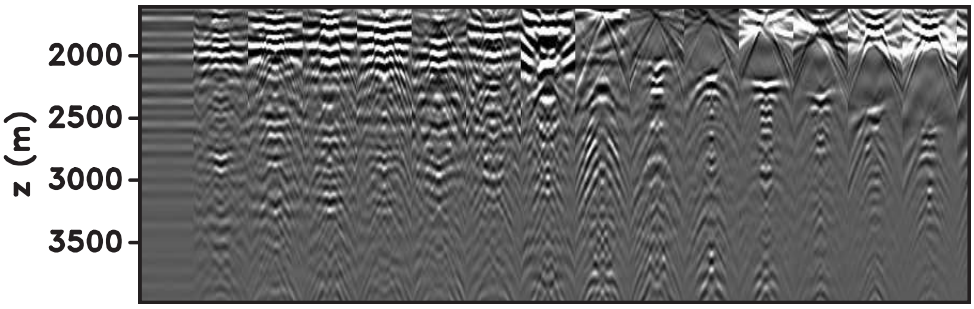
Figure 5.9: Gulf of Mexico dataset example: (a) shows the RTM image obtained within the target area, (b) shows space-lag gathers across the section, and (c) shows the corresponding angle gathers. Note the poor illumination underneath the salt body.



(a)



(b)



(c)

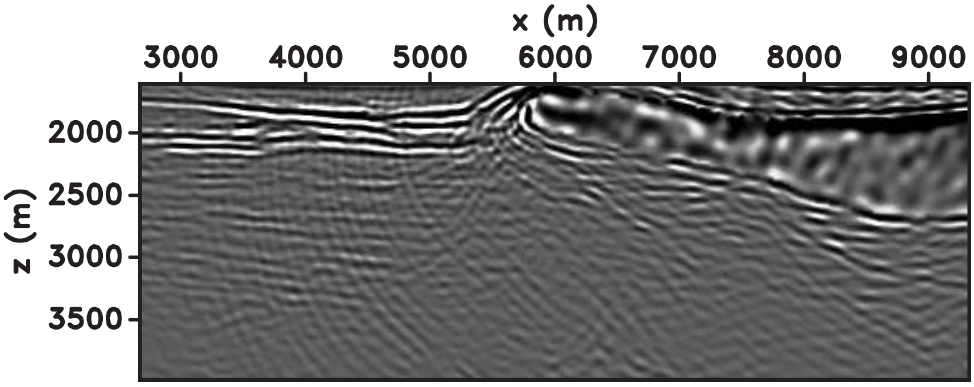
Figure 5.10: Gulf of Mexico dataset example: (a) depicts the RTM image obtained within the target area after the extended deconvolution process, (b) shows the deconvolved space-lag gathers across the section, and (c) shows the corresponding angle gathers (transformed from the deconvolved space-lag gathers. Note how the illumination underneath the salt body improves after deconvolution, now the amplitude is more even across the section.

image energy is very dim, while in the deconvolved RTM image, the sub-salt energy is boosted. However, the deconvolution also seems to highlight migration noise below the left edge of the salt body. The energy balancing effect can also be seen in the deconvolved space-lag and angle-gathers corresponding to Figures 5.10(b)-5.10(c).

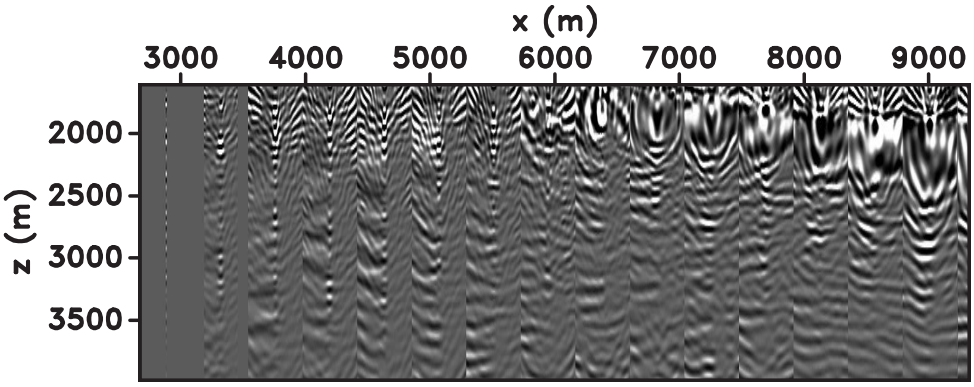
Figure 5.11(a) shows the Marchenko image in the same target area. Despite the fact we used the same input data as in Figure 5.9(a), one can see several differences in the coherency of the seismic events. One such important difference is below 2800m in the sediment area. In Figure 5.9(a) the reflectors end abruptly and migration swings are introduced in the image, whereas in Figure 5.11(a) the reflectors continue to the sub-salt region. After we perform the extended deconvolution, Figure 5.12(a), the signal to noise ratio of the section improves since much of the migration noise is attenuated. In this exercise, we show that both propagation methods behave in a similar way in complex environments. However in our tests, we can observe important differences, such as the improved coherency in the deeper events. Also, the deconvolution process balances the energy through the section and reduces migration artifacts, which effectively increases the signal to noise ratio in the image. The moveout and focusing properties are preserved in both methods, which shows consistency with the previous synthetic tests.

5.5 Conclusions

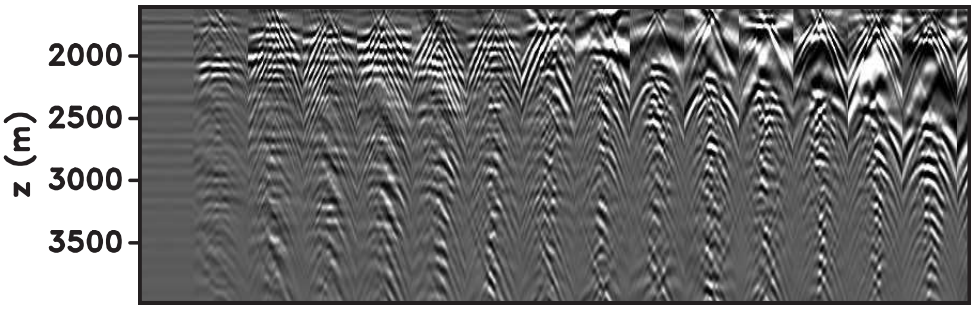
We compare image gathers obtained with Marchenko and RTM propagators. The Marchenko gathers handle primaries, internal multiples, and surface-related multiples. It also produces weak events above the correct interface depth due to the interaction of primaries with the surface-related multiples. In contrast, RTM gathers treat the multiples as primaries; hence, the events are imaged in the wrong location and can bias an inversion towards lower velocities. The Marchenko and RTM gathers show similar sensitivity to errors in the background model. However, the Marchenko gathers possess increased illumination due to increased effective aperture, as shown by the angle gathers in our experiments. We also show the benefits of posing the extended imaging as a deconvolution problem instead of correlation.



(a)

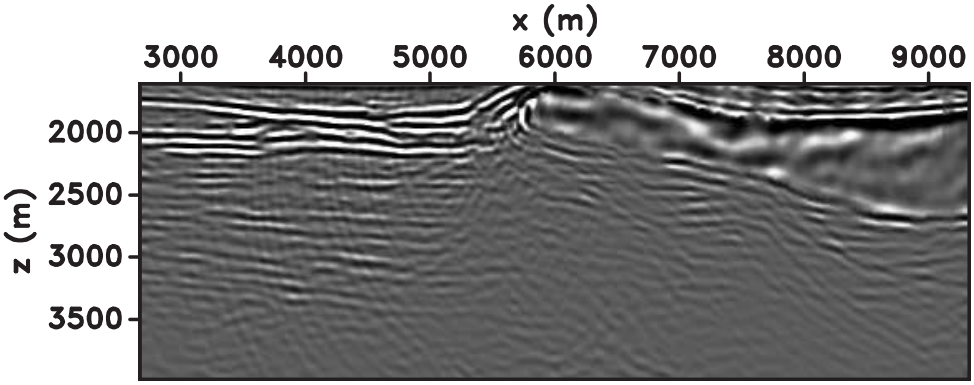


(b)

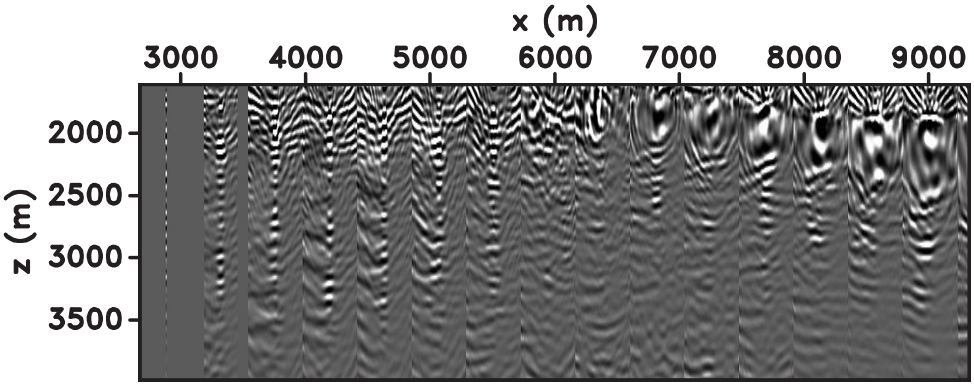


(c)

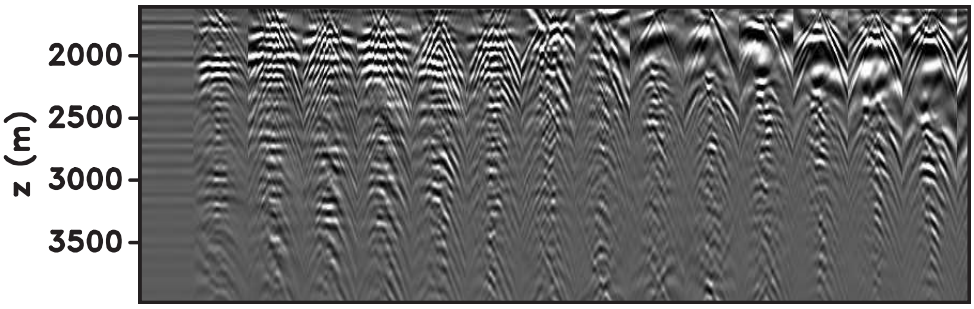
Figure 5.11: Gulf of Mexico dataset example: (a) shows the Marchenko target-oriented image, (b) shows space-lag gathers across the section, and (c) shows the corresponding angle gathers. Comparing with Figure 5.9(a) one can see how the sediments on the left side of the section are more continuous and coherent.



(a)



(b)



(c)

Figure 5.12: Gulf of Mexico dataset example: (a) shows the Marchenko target-oriented image after the extended deconvolution process, (b) shows the corresponding space-lag gathers, and (c) the angle gathers. Note how the continuity of the events improve after the deconvolution process, also the artifacts from migration (migration swings going through sediment reflections) are attenuated after the deconvolution process.

The gather focus improves, and the energy distribution is more uniform as a function of angle. The use of Marchenko wavefields together with extended deconvolution can benefit tomography. The dependency of the Marchenko wavefield with respect to the background velocity model is not as explicit as with the RTM wavefields. In Marchenko modeling, the kinematics are embedded in the direct arrival information used as input for the framework.

5.6 Acknowledgments

We thank the sponsors of the Center for Wave Phenomena, whose support made this research possible. The reproducible numeric examples in this paper use the Madagascar open-source software package (Fomel et al., 2013) freely available from http://www.ahay.org. Esteban would like to thank Alex Jia (CWP) for providing his scripts to obtain the Marchenko wavefields for the GoM dataset.

CHAPTER 6

CONCLUSIONS AND RECOMMENDATIONS

In this thesis, I present different methods for velocity model building with emphasis on extended images and applications to imaging. The wavefields are the main driver of information, by remapping the recorded data from the acquisition surface into the model domain. I start by demonstrating how to incorporate illumination constraints in tomography with common image point gathers which I later supplement with a data-domain inversion. I apply and demonstrate this process to a real dataset in which I show how each step improves the quality of the migrated image, and the synchronization between modeled and observed data. In the data-domain, I use local correlations which forms the misfit information as a function of time while avoiding the cross-talk between events that do not correspond to each other. Similarly to the image-domain, I use a band-limited penalty function that takes into account the bandwidth of the input data which is a more realistic penalty operator than the conventional DSO-based penalty.

Image-domain and data-domain methods have a similar objective: to obtain a velocity field that honors the kinematics of the model that is responsible for the recorded data. While the data-domain tomography aims to fit the data at the receiver locations, the image-domain methods aims to improve desirable features (like focusing) inside the model space. Conceptually, I find the image-domain objective to be more easier to understand since it deals directly the geophysical product we normally use for interpretation in exploration. I think that much of the geological knowledge is used in the interpretation stage could be applied and supplemented into the inversion process as constraints on the background model or the image itself.

Aside from the domain considerations for tomography, one important decision to make is the choice of wave extrapolator. Common extrapolators do not handle correctly the multiples

present in the recorded seismic data. As a result, the multiples contribute to the seismic image as coherent noise that interferes with the imaging process and produces a bias of the tomography operator towards low velocities. In the last two chapters, I use Marchenko-retrieved wavefields as an alternative to handle multiples present in the data. The total wavefield (the sum of its down- and up-going components) satisfies the wave equation. I use this feature for using the wavefield as an input to the wave equation from which I invert for the model parameters. The resulting model is the one that best explains, in a least-squares sense, the scattering present in the wavefields. I also demonstrate how the Marchenko wavefields are sensitive to errors in the background velocity model in a similar way to conventional RTM two-way wavefields. A key characteristic of the obtained gathers is that despite the errors in the background model, primaries and multiples map into the same event. This increases the illumination of a given reflector and reduces the bias of any tomographic operator towards low velocities, which contrasts with conventional extrapolators were the mishandling of multiples introduces a bias towards low velocities. I also apply a spatial deconvolution operator in the space-lag gathers which helps to mitigate noise in the image and it balances the amplitudes in areas with low illumination. The deconvolved gathers present better focusing while handling the multiples correctly. A natural future step is to use such gathers for tomographic updates. I predict that such updates using wavefields with multiples can influence more model parameters during the gradient computation, increasing the illuminated areas in the tomographic update. The cost of the extrapolator is a burden tomography; however, one could take advantage of the target oriented capabilities of the Marchenko process where the tomography could be applied in target areas in the subsurface with low illumination.

REFERENCES CITED

Al-Yahya, K., 1989, Velocity analysis by iterative profile migration: Geophysics, **54**, P. 718–729.

Alkhalifah, T., and Y. Choi, 2014, From tomography to full-waveform inversion with a single objective function: Geophysics, **79**, R55–R61.

Baysal, E., D. D. Kosloff, and J. W. C. Sherwood, 1983, Reverse time migration: Geophysics, **48**, 1514–1524.

Behura, J., K. Wapenaar, and R. Snieder, 2014, Autofocus imaging: Image reconstruction based on inverse scattering theory: Geophysics, **79**, A19–A26.

Biondi, B., and A. Almomin, 2014, Simultaneous inversion of full data bandwidth by tomographic full-waveform inversion: Geophysics, **79**, WA129–WA140.

Biondi, B., and W. Symes, 2004, Angle-domain common-image gathers for migration velocity analysis by wavefield-continuation imaging: Geophysics, **69**, 1283–1298.

Bleistein, N., 1987, On the imaging of reflectors in the earth: Geophysics, **52**, 931–942.

Broggini, F., and R. Snieder, 2012, Connection of scattering principles: a visual and mathematical tour: European Journal of Physics, **33**, 593.

Bunks, C., F. Saleck, S. Zaleski, and G. Chavent, 1995, Multiscale seismic waveform inversion: Geophysics, **60**, 1457–1473.

Claerbout, J. F., 1971, Toward a unified theory of reflector mapping: Geophysics, **36**, 467–481.

———, 2008, *in* Basic earth imaging: Stanford University, 11–22.

Cullison, T., and P. Sava, 2011, An image-guided method for automatically picking commonimagepoint gathers: 3908–3912.

Díaz, E., and P. Sava, 2015, Wavefield tomography using reverse time migration backscattering: Geophysics, **80**, R57–R69.

Dix, C., 1955, Seismic velocities from surface measurements: Geophysics, **20**, 68–86.

Dragoset, B., 1999, Introduction to the workshop papers: The Leading Edge, **18**, 85–85.

Engquist, B., and B. D. Froese, 2013, Application of the wasserstein metric to seismic signals: arXiv preprint arXiv:1311.4581.

Etgen, J., S. Gray, and Y. Zhang, 2009, An overview of depth imaging in exploration geophysics: Geophysics, **74**, WCA5–WCA17.

Fletcher, R. P., S. Archer, D. Nichols, and W. Mao, 2012, Inversion after depth imaging: 1–5.

Fleury, C., M. Giboli, and R. Baina, 2014, Moveout analysis for reflection tomography on wave-equation image gathers: SEG Technical Program Expanded Abstracts 2014, 4788–4792.

Fomel, S., P. Sava, I. Vlad, Y. Liu, and V. Bashkardin, 2013, Madagascar: open-source software project for multidimensional data analysis and reproducible computational experiments: Journal of Open Research Software, **1**, e8.

Gardner, G., L. Gardner, and A. Gregory, 1974, Formation velocity and densitythe diagnostic basics for stratigraphic traps: Geophysics, **39**, 770–780.

Gazdag, J., 1978, Wave equation migration with the phase-shift method: Geophysics, **43**, 1342–1351.

Gazdag, J., and P. Sguazzero, 1984, Migration of seismic data by phase shift plus interpolation: Geophysics, **49**, 124–131.

Giboli, M., R. Baina, L. Nicoletis, and B. Duquet, 2012, Reverse time migration surface offset gathers part 1: a new method to produce 'classical' common image gathers: SEG Technical Program Expanded Abstracts 2012, 1–5.

Grion, S., R. Exley, M. Manin, X. Miao, A. Pica, Y. Wang, P. Granger, and S. Ronen, 2007, Mirror imaging of obs data: first break, **25**.

Guitton, A., 2002, Shotprofile migration of multiple reflections: SEG Technical Program Expanded Abstracts 2002, 1296–1299.

———, 2005, Multiple attenuation in complex geology with a pattern-based approach: Geophysics, **70**, V97–V107.

Guitton, A., G. Ayeni, and E. Díaz, 2012, Constrained full-waveform inversion by model reparameterization: Geophysics, **77**, R117–R127.

Guitton, A., and J. Claerbout, 2015, Nonminimum phase deconvolution in the log domain: A sparse inversion approach: Geophysics, **80**, WD11–WD18.

Guitton, A., A. Valenciano, D. Bevc, and J. Claerbout, 2007, Smoothing imaging condition for shot-profile migration: Geophysics, **72**, S149–S154.

Hale, D., 2006a, An efficient method for computing local cross-correlations of multi-dimensional signals: CWP, **544**, 1–7.

———, 2006b, Recursive gaussian filters: CWP-546.

Jia, A., S. Singh, A. Guitton, and R. Snieder, 2016, Marchenko imaging for subsalt field data: CWP Annual Meeting 2016.

Jin, H., G. A. McMechan, and H. Guan, 2014, Comparison of methods for extracting adcigs from rtm: Geophysics, **79**, S89–S103.

Kamei, R., R. G. Pratt, and T. Tsuji, 2013, On acoustic waveform tomography of wide-angle obs data-strategies for pre-conditioning and inversion: Geophysical Journal International.

Lailly, P., 1983, The seismic inverse problem as a sequence of before stack migrations: Conference on inverse scattering: theory and application, 206–220.

Lin, T. T. Y., and F. J. Herrmann, 2013, Robust estimation of primaries by sparse inversion via one-norm minimization: Geophysics, **78**, R133–R150.

Liu, J., and W. Han, 2010, Automatic event picking and tomography on 3d rtm angle gathers: 4263–4268.

Lomask, J., and B. Biondi, 2005, Image segmentation with bounds: SEP report 120.

Luo, S., and P. Sava, 2012, A deconvolution–based objective function for wave–equation inversion: SEG Technical Program Expanded Abstracts 2011, 2788–2792.

Luo, Y., and G. Schuster, 1991, Wave-equation traveltime inversion: Geophysics, **56**, 645–653.

Manduca, A., T. Oliphant, M. Dresner, J. Mahowald, S. Kruse, E. Amromin, J. Felmlee, J. Greenleaf, and R. Ehman, 2001, Magnetic resonance elastography: Non-invasive mapping of tissue elasticity: Medical Image Analysis, **5**, 237 – 254.

Martin, G. S., K. J. Marfurt, and S. Larsen, 2002, Marmousi–2: an updated model for the investigation of avo in structurally complex areas: SEG Expanded Abstracts, 1979–1982.

McMechan, G. A., 1983, Migration by extrapolation of time-dependent boundary values: Geophysical Prospecting, **31**, 413–420.

Paffenholz, J., B. McLain, J. Zaske, and P. Keliher, 2002, Subsalt multiple attenuation and imaging: Observations from the sigsbee 2b synthetic dataset: 72nd Annual International Meeting, SEG, Expanded Abstracts, 2122–2125.

Perrone, F., P. Sava, and J. Panizzardi, 2015, Waveform tomography based on local image correlations: Geophysical Prospecting, **63**, 35–54.

Plessix, R.-E., 2006, A review of the adjoint-state method for computing the gradient of a functional with geophysical applications: Geophysical Journal International, **167**, 495–503.

Pratt, R., 1999, Seismic waveform inversion in the frequency domain, part 1: Theory and verification in a physical scale model: Geophysics, **64**, 888–901.

Rickett, J. E., and P. C. Sava, 2002, Offset and angle-domain common image-point gathers for shot-profile migration: Geophysics, **67**, 883–889.

Rose, J., 2002a, Time reversal, focusing and exact inverse scattering, *in* Imaging of Complex Media with Acoustic and Seismic Waves: Springer Berlin Heidelberg, volume **84** *of* Topics in Applied Physics, 97–106.

Rose, J. H., 2002b, 'single-sided' autofocusing of sound in layered materials: Inverse Problems, **18**, 1923.

Sava, P., and B. Biondi, 2004a, Wave-equation migration velocity analysis - I: Theory: Geophysical Prospecting, **52**, 593–606.

———, 2004b, Wave-equation migration velocity analysis - II: Subsalt imaging examples: Geophysical Prospecting, **52**, 607–623.

Sava, P., and S. Fomel, 2006, Time-shift imaging condition in seismic migration: Geophysics, **71**, S209–S217.

Sava, P., and A. Guitton, 2005, Multiple attenuation in the image space: Geophysics, **70**, V10–V20.

Sava, P., and I. Vasconcelos, 2011a, Extended imaging condition for wave-equation migration: Geophysical Prospecting, **59**, 35–55.

———, 2011b, Extended imaging conditions for wave-equation migration: Geophysical Prospecting, **59**, 35–55.

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

Schneider, W. A., 1978, Integral formulation for migration in two and three dimensions: Geophysics, **43**, 49–76.

Shan, G., and Y. Wang, 2013, Rtm based wave equation migration velocity analysis: 4726–4731.

Shen, P., and H. Calandra, 2005, One-way waveform inversion within the framework of adjoint state differential migration: 1709–1712.

Shen, P., and W. W. Symes, 2008, Automatic velocity analysis via shot profile migration: Geophysics, **73**, VE49–VE59.

Shen, Y., B. Biondi, and R. Clapp, 2015, Wave-equation based q tomography from angle-domain common image gathers: 4334–4338.

Shin, C., and W. Ha, 2008, A comparison between the behavior of objective functions for waveform inversion in the frequency and Laplace domains: Geophysics, **73**, VE119–VE133.

Shin, C., and Y. Ho Cha, 2009, Waveform inversion in the Laplace–Fourier domain: Geophysical Journal International, **177**, 1067–1079.

Singh, S., and R. Snieder, 2016, Strategies for imaging with Marchenko-retrieved Greens functions: Geophysics, submitted.

Singh, S., R. Snieder, J. Behura, J. van der Neut, K. Wapenaar, and E. Slob, 2015a, Marchenko imaging: Imaging with primaries, internal multiples, and free-surface multiples: Geophysics, **80**, S165–S174.

Singh, S., R. Snieder, J. van der Neut, J. Thorbecke, E. S. ob, and K. Wapenaar, 2015b, Incorporating free-surface multiples in marchenko imaging: SEG Technical Program Expanded Abstracts 2015, 4328–4333.

Singh, S., R. Snieder, J. van der Neut, J. Thorbecke, E. Slob, and K. Wapenaar, 2016, Accounting for free-surface multiples in marchenko imaging: SEG Annual meeting.

Sirgue, L., O. Barkved, J. Van Gestel, O. Askim, and J. Kommedal, 2009, 3d waveform inversion on valhall wide-azimuth obc: 71st EAGE Conference & Exhibition.

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

Symes, W. W., 2008, Migration velocity analysis and waveform inversion: Geophysical Prospecting, **56**, 765–790.

Tarantola, A., 1984, Inversion of seismic reflection data in the acoustic approximation: Geophysics, **49**, 1259–1266.

Thorbecke, J. W., and D. Draganov, 2011, Finite-difference modeling experiments for seismic interferometry: Geophysics, **76**, H1–H18.

Valenciano, A. A., B. L. Biondi, and R. G. Clapp, 2009, Imaging by target-oriented wave-equation inversion: Geophysics, **74**, WCA109–WCA120.

van der Neut, J., J. Thorbecke, K. Mehta, E. Slob, and K. Wapenaar, 2011, Controlled-source interferometric redatuming by crosscorrelation and multidimensional deconvolution in elastic media: Geophysics, **76**, SA63–SA76.

van Leeuwen, T., and W. A. Mulder, 2010, A correlation-based misfit criterion for wave-equation traveltime tomography: Geophysical Journal International, **182**, 1383–1394.

Verschuur, D. J., A. J. Berkhout, and C. P. A. Wapenaar, 1992, Adaptive surfacerelated multiple elimination: Geophysics, **57**, 1166–1177.

Verschuur, D. J. E., and A. J. G. Berkhout, 2011, Seismic migration of blended shot records with surface-related multiple scattering: Geophysics, **76**, A7–A13.

Virieux, J., and S. Operto, 2009, An overview of full-waveform inversion in exploration geophysics: Geophysics, **74**, WCC1–WCC26.

Vyas, M., D. Nichols, and E. Mobley, 2011, Efficient rtm angle gathers using source directions: 3104–3108.

Wang, B., C. Mason, M. Guo, J. Cai, S. Gajawada, and D. Epili, 2010, Interactive demultiple in the postmigration depth domain: 3436–3440.

Wapenaar, K., J. Thorbecke, J. van der Neut, F. Broggini, E. Slob, and R. Snieder, 2014a, Green's function retrieval from reflection data, in absence of a receiver at the virtual source position: The Journal of the Acoustical Society of America, **135**, 2847–2861.

———, 2014b, Marchenko imaging: Geophysics, **79**, WA39–WA57.

Warner, M., and L. Guasch, 2014, Adaptive waveform inversion: Theory: SEG Technical Program Expanded Abstracts 2014, 1089–1093.

Weglein, A. B., F. A. Gasparotto, P. M. Carvalho, and R. H. Stolt, 1997, An inversescattering series method for attenuating multiples in seismic reflection data: Geophysics, **62**, 1975–1989.

Weibull, W., and B. Arntsen, 2013a, Automatic velocity analysis with reverse-time migration: Geophysics, **78**, S179–S192.

———, 2013b, Reverse time demigration using the extended imaging condition: SEG Technical Program Expanded Abstracts 2013, 4015–4020.

Whitmore, N. D., 1983, Iterative depth migration by backward time propagation: SEG Technical Program Expanded Abstracts, **2**, 382–385.

Whitmore, N. D., A. A. Valenciano, W. Sollner, and S. Lu, 2010, Imaging of primaries and multiples using a dual-sensor towed streamer: SEG Technical Program Expanded Abstracts 2010, 3187–3192.

Wong, M., B. L. Biondi, and S. Ronen, 2015, Imaging with primaries and free-surface multiples by joint least-squares reverse time migration: Geophysics, **80**, S223–S235.

Woodward, M., 1992, Wave-equation tomography: Geophysics, **57**, 15–26.

Yang, T., and P. Sava, 2010, Moveout analysis of wave-equation extended images: Geophysics, **75**, S151–S161.

———, 2015, Image-domain wavefield tomography with extended common-image-point gathers: Geophysical Prospecting, **63**, 1086–1096.

Yang, T., J. Shragge, and P. Sava, 2013, Illumination compensation for image-domain wavefield tomography: Geophysics, **78**, U65–U76.

Yoon, K., M. Guo, J. Cai, and B. Wang, 2011, 3d rtm angle gathers from source wave propagation direction and dip of reflector: SEG Technical Program Expanded Abstracts 2011, 3136–3140.

Yoon, K., K. J. Marfurt, and W. Starr, 2004, Challenges in reversetime migration: SEG Technical Program Expanded Abstracts, 1057–1060.

Zhang, Y., and G. Shan, 2013, Wave-equation migration velocity analysis using partial stack-power maximization: SEG Technical Program Expanded Abstracts 2013, 4847–4852.