The Dissertation Committee for Yi Tao Certifies that this is the approved version of the following dissertation:

Time reversal and plane-wave decomposition in seismic interferometry, inversion and imaging

Committee:

Mrinal K. Sen, Supervisor

Robert H. Tatham

Clark R. Wilson

Kyle T. Spikes

Yosio Nakamura

Gail L. Christeson

Time reversal and plane-wave decomposition in seismic interferometry, inversion and imaging

by

Yi Tao, B.S.; M.S.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

December 2012

Dedication

To my parents

Acknowledgements

Time reversal and plane-wave decomposition in seismic interferometry, inversion and imaging

Yi Tao, PhD.

The University of Texas at Austin, 2012

Supervisor: Mrinal K. Sen

This thesis concerns the study of time reversal and plane-wave decomposition in various geophysical applications. Time reversal is a key step in seismic interferometry, reverse time migration and full waveform inversion. The plane-wave transform, also known as the tau-p transform or slant-stack, can separate waves based on their ray parameters or their emergence angles at the surface.

I propose a new approach to retrieve virtual full-wave seismic responses from crosscorrelating recorded seismic data in the plane-wave domain. Unlike a traditional approach where the correlogram is obtained from crosscorrelating recorded data, which contains the full range of ray parameters, this method directly chooses common ray parameters to cancel overlapping ray paths. Thus, it can sometime avoid spurious arrivals when the acquisition requirement of seismic interferometry is not strictly met. I demonstrate the method with synthetic examples and an ocean bottom seismometer data example. I show a multi-scale application of plane-wave based full waveform inversion (FWI) with the aid of frequency domain forward modeling. FWI uses the two-way wave-equation to produce high-resolution velocity models for seismic imaging. This technique is implemented by an adjoint-state approach, which

involves a time-reversal propagation of the residual wavefield at receivers, similar to seismic interferometry. With a plane-wave transformed gather, we can decompose the data by ray parameters and iteratively update the velocity model with selected ray parameters. This encoding approach can significantly reduce the number of shots and receivers required in gradient and Hessian calculations. Borrowing the idea of minimizing different data residual norms in FWI, I study the effect of different scaling methods to the receiver wavefield in the reverse time migration. I show that this type of scaling is able to significantly suppress outliers compared to conventional algorithms. I also show that scaling by its absolute norm generally produces better results than other approaches. I propose a robust stochastic time-lapse seismic inversion strategy with an application of monitoring Cranfield CO_2 injection site. This workflow involves two steps. The first step is the baseline inversion using a hybrid starting model that combines a fractal prior and the low-frequency prior from well log data. The second step is to use a double-difference inversion scheme to focus on the local areas where time-lapse changes have occurred. Synthetic data and field data show the effectiveness of this method.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

1.1 MOTIVATION

Seismic waves carry a wealth of information. By studying the seismic waves typically recorded at the Earth's surface, we are able to construct detailed images of the subsurface. This is the most cost-effective and reliable way to unravel the subsurface properties, to prospect oil and gas, and to predict natural hazards such as earthquakes and tsunamis. To derive more information from seismic waves, we need to use advanced geophysical tools. During the past decade, seismic interferometry, full waveform inversion (FWI) and reverse time migration (RTM) have become fairly popular in geophysical community. The study of these advanced techniques is fundamentally important in both academic and applied seismology.

One common aspect of these techniques is that they all use a time reversal process. In this process, the seismic waves from a source inside a medium are first recorded by an array of receivers located at the boundary of the medium, and then they are reversed in time and back-propagated into the medium. This back-propagation results in an energy focusing at or near the original source location. This was demonstrated by Fink in 1997 through physical experiments and later was extensively studied in different disciplines (e.g., Snieder et al. 2002 used this approach to study scattered coda waves). For practical implementations in interferometry, however, this back-propagation process is accomplished with a reciprocity principle and a crosscorrelation procedure.

Another common feature of these three methods is the fact that all of them are based on the full waveforms, i.e., they use both travel time and amplitude information. The travel time is related to long period (smooth) variation of the model parameters (typically velocity). The amplitude, on the other hand, is related to short period (rapid)

1

changes of the model parameters. Early stages of modern seismology relied mainly on travel time information (e.g., Oldham, 1906). Amplitude-related processing techniques began to become popular after 1980s when people realized that amplitude variation with offset (AVO) is important in exploration seismology (Ostrander, 1984). Compared to ray-based imaging approaches, full waveform methods honor the basic physics of seismic propagation which include multiple scattering, attenuation and anisotropic effects. It can also maintain the band-limited frequency effects instead of using infinite-frequency rays. It is therefore no surprise that full wave methods are now being studied extensively.

The main differences among the three methods are their distinct objectives and seemingly disconnected numerical implementations. Seismic interferometry is used to create virtual shot gathers (redatuming); FWI is used to derive elastic parameters (mostly velocity); and RTM is used to create an image section. In terms of implementation, the most significant difference is that both FWI and RTM require velocity information for wavefield extrapolation while interferometry does not require this information. Another difference is that FWI requires iterative processes to search for the optimal solution while interferometry and RTM are usually implemented as a one-step deterministic process.

The plane-wave transform (also known as the $\tau - p$ transform or slant stack) is used to generate a $\tau - p$ gather by applying a constant time shift to each offset trace and summing all offsets. This transform offers many advantages in multiple suppression (Liu et al., 2000), velocity analysis (Diebold and Stoffa, 1981), anisotropy analysis (Sen and Mukherjee 2003; Sil and Sen, 2009), direct inversion of plane-wave gather (Sen and Stoffa, 1991) and source-encoding for wave equation migration and inversion (Zhang et al., 2005, Vigh and Starr, 2008). In general, the $\tau - p$ domain is an excellent space for many geophysical applications.

For every inverse problem, non-uniqueness can be a challenge because even the data residual is small. Correspondingly, the inverted model can still be away from the "true" model. This situation can be improved if we consider building a good starting model and applying constraints in the inversion process.

The objective of this thesis is to investigate the applications of plane-waves and time-reversal methods to seismic redatuming, inversion and migration problems. Challenges in inverse problems such as non-uniqueness are also considered in this thesis. I will develop new methods to facilitate better applications of these algorithms. I will use theoretical analysis, synthetic numerical experiments and field data examples to demonstrate this.

1.2 FORWARD MODELING

Because of the complexity of real seismic problems, analytical solutions (e.g., Lamb, 1904; Trifunac, 1971) of wave propagations are rarely used in practice. Instead, forward modeling with numerical algorithms is used to generate synthetic data from given model(s). Forward modeling is the basis for redatuming, inversion and migration. There are many algorithms for forward modeling with different numerical approximations, accuracies and computational costs. These algorithms have several advantages and disadvantages; the choice of an algorithm depends on the problems we want to solve. Mathematically, all seismic forward modeling operators can be generally written as

$$\mathbf{d} = \mathbf{F}(\mathbf{m}), \tag{1.1}$$

where $\mathbf{m}$ are the model parameters - typically elastic parameters including P-wave velocity, S-wave velocity, density and so on; $\mathbf{d}$ is the data and $\mathbf{F}$ is the forward operator, which is typically nonlinear.

Perhaps the simplest forward modeling strategy for synthetic seismogram is the convolution model (e.g., Rowbotham et al., 2003; Hampson et al., 2005). This is used to generate flattened gathers that can match normal moveout (NMO) corrected common middle point (CMP) gathers or migrated gathers generated from acquired reflection seismic data. The convolution based forward modeling is used extensively in amplitude versus offset (AVO) analysis and in simultaneous inversion problems.

In contrast to a convolution model that characterizes the amplitude information of seismic data, ray-based methods (Červený, 2001) describe travel time. They track the movement of individual points of the propagating wavefront based on the asymptotic assumption of high-frequency solution of the elastodynamic equation. These methods are extensively used in ray-based migration and tomography applications.

Reflectivity modeling and wave equation modeling describe waveforms, i.e., both amplitude and travel time. Reflectivity modeling (e.g., Thomson, 1950; Kennett, 1983) is a semi-analytic approach that describes wave propagation in the frequency-wavenumber domain based on stratified earth assumption. Reflection/transmission coefficients are determined at the layer boundaries and the communication among layers is done with a propagator matrix approach. Full-wave signals including reflections, transmissions, converted waves and multiples, can be modeled with this approach.

Two-way wave equation methods are ideal for tackling many challenging exploration and production problems. These methods are derived from the elastodynamic wave equations for heterogeneous media. They do not have the limitations of the ray-based or 1D reflectivity methods. However, the grid size (or step length) of the propagation media should be small enough to satisfy the grid dispersion and stability conditions. Both FWI and RTM specifically require two-way wave equation methods as their forward modeling strategies. Seismic interferometry is not dependent on the forward

modeling methods. However, most of the numerical tests reported in literature are based on two-way wave equation methods.

There are several numerical methods for solving seismic wave equation – they differ mainly based on the schemes used for spatial or temporal discretization. Three most common numerical methods are the finite difference method (FDM), finite element method (FEM) and boundary element method (BEM). FDM is based on the truncated Taylor series of the partial derivatives. Standard-grid formulation (Alford et al., 1974) and staggered-grid formulation (Levander, 1988) are the two most common FD methods. FEM is based on a weak form (integral formula) of the wave equation. Finite volume method (Dormy and Tarantola, 1995), spectral element method (Komatitsch and Vilotte, 1998; De Basabe and Sen, 2007) and discontinuous Galerkin method (Rivière et al., 1999; De Basabe et al., 2008) can be considered as variants or special cases of FEM. BEM is also based on the weak form of the wave equation but the discretization is done only along the interfaces and thus it only can be applied to piecewise heterogeneous media (Schuster, 1985).

In this dissertation I will use FDM for two-way wave equation wavefield simulations because of its simplicity to implement and acceptable computational cost. A time-domain approach is used for seismic interferometry and RTM and a frequency-domain approach is used for FWI. The time-domain approach uses an explicit time marching algorithm and is therefore easier to implement. The frequency-domain approach requires a solution of a linear system and thus it is harder to implement but it is faster for 2-D problems when a large number of sources are needed for wavefield simulations.

Mathematically, the time-domain approach is to perform time marching on the following spatial discretized wave equation (isotropic case with no attenuation)

5

$$\mathbf{M\ddot{u}}(t) + \mathbf{Ku}(t) = \mathbf{f}(t) \quad , \tag{1.2}$$

where the matrices $\mathbf{M}$ and $\mathbf{K}$ denote the mass matrix and the stiffness matrix, respectively; the vector $\mathbf{f}$ is the source term. We write $\mathbf{M}$ and $\mathbf{K}$ as matrices to compare the frequency-domain solution, but in practice these two matrices are rarely formed and local finite difference stencils are used instead.

The frequency-domain modeling is based on the following space-discrete displacement formulation

$$-\omega^2 \mathbf{Mu}(\omega) + \mathbf{Ku}(\omega) = \mathbf{f}(\omega) \quad . \tag{1.3}$$

Equation 1.3 is the Fourier transformed version of equation 1.2. In order to solve this equation, $\mathbf{M}$ and $\mathbf{K}$ should be allocated explicitly and a linear system should be solved.

1.3 SEISMIC INTERFEROMETRY

Seismic interferometry refers to the method to retrieve a new Green's function between two different receivers from recorded wavefields. Similar to optical interferometry, which takes advantage of the interference pattern of the light beams to reveal the properties of the medium, seismic interferometry uses the interference pattern of seismic waves to study the elastic properties of the earth. The mathematical tool for this method in the time domain is crosscorrelation (Schuster, 2001; Schuster et al., 2004), crossconvolution (Curtis et al., 2006) or deconvolution (Vasconcelos et al., 2008a, 2008b). The most commonly used approach is crosscorrelation. All of these approaches can be expressed in the following general form

$$\mathbf{d} = \mathbf{R}(\mathbf{d}_0) \quad , \tag{1.4}$$

where $\mathbf{d}_0$ and $\mathbf{d}$ represent original data in the acquired geometry and interferometry data in the redatumed domain, respectively; $\mathbf{R}$ is an operator in data space.

Seismic interferometry can redatum the data from one type of acquisition geometry to another type of acquisition geometry. As suggested in equation 1.4, interferometry operates in the data domain. Thus, the redatuming operators used in interferometry do not require a velocity model, which is very different from the redatuming methods with equivalent source terms (Mulder, 2005).

Application of this redatuming technique offers several advantages. First, ambient noise recordings could be turned into useful signals by generating new shot gathers. Second, source and/or receiver arrays can be redatumed to be closer to the target without knowing the velocity model and thus, complex overburden can be avoided using this technique. Third, it enables us to view the earth with different viewing apertures and angles. Finally, multiple scattered arrivals can be transformed into primary reflections to provide super-illumination, super-resolution and elimination of statics.

The pioneering work of seismic interferometry can be attributed to Claerbout (1968). In that paper, Claerbout demonstrated that for a normal incidence plane wave in a stratified 1-D acoustic medium, by cross-correlating different transmitted signals from the free surface, one can get a new signal that appears to have propagated from one receiver to another receiver via subsurface. This wave, unlike the conventional reflected wave that has only causal part, contains a non-causal part and an infinite pulse. This work was extended to the elastic case with arbitrary incidence angle by Fraiser (1970). Wapenaar (2003, 2006) developed a theory for seismic interferometry based on seismic reciprocity. This is probably the most rigorous theoretical proof of seismic

interferometry. However, proof with wave modes separated reflectivity method is not well known to the geophysics community. I show this proof in Appendix A.

1.4 FULL WAVEFORM INVERSION

In the context of numerical mathematics, any non-linear inverse problem can be generally represented as (Menke, 1984)

$$\mathbf{m} = \mathbf{F}^{-1}(\mathbf{d}) \quad . \tag{1.5}$$

As shown in equation 1.5, the purpose of inversion is to derive subsurface model parameters from seismic data. It generally involves a process to minimize a misfit function (also known as error function or cost function) between observed and forward modeled data. A myriad of geophysical inversion approaches have been developed over the last several decades. They can be grouped into different categories mainly by their forward modeling strategies.

For FWI, the forward modeling usually refers to two-way wave equation waveform simulation methods. The purpose of FWI is to produce high-resolution elastic models for seismic imaging. Usually, the full elastic parameters of the subsurface for seismic imaging are reduced to a single parameter—acoustic or P-wave velocity. There are two main types of optimization approaches. The first one is a local optimization method based on the gradient of the misfit function. It usually requires a linearization of the non-linear problem based on the Born approximation. The second one is a global optimization method based on Monte Carlo simulations. It randomly calculates many models and decides the optimal model parameters and uncertainties with those calculations. So far, because of the computational cost and non-uniqueness, FWI has been implemented with an adjoint-state (Tarantola, 1984, Virieux and Operto, 2009) based local optimization method to iteratively update the model parameters.

8

1.5 REVERSE TIME MIGRATION

Migration aims to recover the true structure of the subsurface from recorded seismic data by moving dipping reflectors to their true subsurface positions. Mathematically, a general description of the migration operators can be written as follows (Claerbout, 1992)

$$\mathbf{m}_{mig} = \mathbf{F}^T(\mathbf{d}) \qquad , \qquad (1.6)$$

where $\mathbf{m}_{mig}$ is the migration image; $\mathbf{F}^T$ is the migration operator, which is the adjoint of the forward modeling operator. The term $\mathbf{m}_{mig}$ is different from the model parameter $\mathbf{m}$ because it has the same data size as NMO corrected data in the migrated gather, and the stacked image is created by stacking of the migrated data.

Migration algorithms can be classified based on extrapolation schemes. RTM uses two-way wave equation as its wavefield extrapolation scheme. Migration algorithms can also be classified based on the processing sequence (post-stack and pre-stack) and operation domain (time and depth). RTM is a depth migration approach, typically applied to the pre-stack data. Similar to all other wave equation based migration algorithms, it requires forward propagation of the source wavefield, back propagation of the receiver wavefield and an application of an imaging condition.

In addition to the general common and distinct features between interferometry and FWI, RTM can be specifically viewed as an interferometry process that requires subsurface velocity information and backward extrapolation. It can also be viewed as the gradient of the minimizing functionals (Tarantola, 1984, 1988; Mora, 1989).

1.6 ORGANIZATION

To give a complete description of the novel developments and applications I have made, this thesis is composed of the following chapters:

Chapter 2 describes an application of plane-wave transformation in the crosscorrelation-type seismic interferometry. I show synthetic examples for transmission to reflection retrieval and walkway vertical seismic profile (VSP) imaging of salt flank. I also show application of plane-wave for super-virtual seismic interferometry with synthetic data and real ocean bottom seismometer (OBS) data.

Chapter 3 describes the frequency-domain full waveform inversion with plane-wave encoded gradient and diagonal Hessian. This encoding can save computational cost because typically a ray parameter has wider illumination coverage than a shot gather.

Chapter 4 provides detailed analysis of the scaling of the receiver wavefield to the migrated data. This approach borrows the idea from full waveform inversion and it can be used to suppress the non-Gaussian noises.

Chapter 5 describes a robust stochastic time-lapse seismic inversion strategy with an application of monitoring a CO_2 injection site. This workflow involves a baseline inversion using a hybrid starting model that combines a fractal prior and the low-frequency prior from well log data. A second step of this workflow is to use a double-difference inversion scheme.

Chapter 6 contains the overall conclusions from my work. Possible future work is also described in this chapter.

Chapter 2: Plane-wave based seismic interferometry

2.1 ABSTRACT

In this chapter, I present a new approach to retrieve virtual seismic responses from crosscorrelating acquired seismic data in the plane-wave domain. Using this method, slant stacking is first performed over shot or receiver locations of observed seismic data to produce plane-wave transformed gathers. Crosscorrelation is then performed by selecting traces with the same ray parameters from different shot or receiver locations of the plane-wave gathers. Unlike traditional crosscorrelation-type time-space domain interferometry where full range of ray parameters are used for each survey location, this method directly selects common ray parameters to cancel overlapping ray paths. This approach can be used to select certain ranges of ray parameters and therefore can retrieve energy only from certain directions. It can also avoid spurious arrivals in super-virtual interferometry, where unwanted arrivals such as reflections violate the requirement of conventional interferometry. Computation time can be saved with this approach since plane-wave transform usually results in a reduction of the original data volume. I demonstrate this method with synthetic and field OBS data examples.

2.2 INTRODUCTION

In recent years, seismic interferometry, also known as Green's function retrieval, has grown into a thriving research area with many novel applications. Since it can redatum the seismic data from a given acquisition geometry to another source or receiver related geometry, and this redatuming does not require a velocity model, this technique has been applied extensively to controlled-source exploration data (e.g., Schuster, 2001; Wapenaar, 2006; Bakulin and Calvert 2006; Curtis, 2009). Another major type of data where seismic interferometry has been widely applied is passive seismic data with natural

sources (e.g., Rickett and Claerbout, 1999; Sabra et al., 2005; Draganov et al., 2007). For a comprehensive review of the theory and applications of seismic interferometry, the readers are referred to Schuster (2009), and Wapenaar et al. (2010a and 2010b).

In classical interferometry, a virtual trace is obtained by crosscorrelating traces from two different receiver or source locations. Under high-frequency approximation, this process can be viewed as cancelation of common part of a ray path from a physical source to the two different receivers. This cancelation results in a raypath connecting a virtual source to a receiver (Schuster et al., 2004). A stacking operator is then applied to the correlogram to take advantage of contributions of all the sources. This entire process is done in the time-space ($t - \mathbf{x}$) domain or in its equivalent frequency domain. However, in this domain, sources and receivers may be irregularly positioned and if a large number of traces need to be crosscorrelated, it can be computationally time consuming, especially when the crosscorrelation is done in time domain.

Transforming seismic data from time-space domain to the plane-wave ($\tau - \mathbf{p}$) domain involves a mapping from original source and receiver distributions to ray parameter ($\mathbf{p}$) distributions (e.g., Stoffa et al., 2006). This results in a regularized coordinate system even for irregular input data. This mapping is generally applicable to reflection seismic data, such as surface seismic profile (SSP) and vertical seismic profile (VSP), both of which are typically considered in seismic interferometry (Schuster and Zhou, 2006). This mapping is also applicable to refraction geometry used in a recent work of super-virtual interferometry (Mallison et al., 2011). In the plane-wave domain, each ray parameter corresponds to a certain angle of incidence. This transformation offers several advantages and numerous methods have been investigated for seismic wave filtering (Tatham, 1989), multiple attenuation (Liu et al., 2000), seismic forward

modeling (Vigh and Starr, 2008), migration (Stoffa et al., 2006) and inversion (Sen et al., 2003).

In this chapter, I propose retrieving seismic responses based on crosscorrelation in the plane-wave domain. Classical interferometry automatically finds a ray from a stationary source, passing through two different receivers and cancels the common ray path. Plane-wave domain interferometry decomposes the data into different ray parameters and cancels the common ray paths by crosscorrelating the same ray parameters. I will demonstrate that this method can obtain results similar to those obtained in time-space domain, but it offers several advantages such as a regular coordinate system, computational efficiency and the flexibility to choose ray parameters for seismic data redatuming.

2.3 THEORY

Plane-wave transform, also known as $\tau - \mathbf{p}$ transform or slant stack, has been investigated by many authors (e.g. Stoffa, 1989, Foster and Mosher 1992). The key idea of this transformation is to obtain "ray-parameter" gathers from common shot or common receiver gathers. It can be implemented either in time domain or in frequency domain. In the time domain, it involves summing of amplitudes along lines of constant slope called ray-parameters. In the frequency domain, this transform involves a phase shift determined by ray parameters and offsets and a summation over offsets. For acquisition geometries with surface sources and surface receivers or for transmission problems, the frequency domain version of plane-wave transform can be written as follows

$$\tilde{d}(\mathbf{p} \,|\, \mathbf{x}, \omega) = \int d(\mathbf{h} \,|\, \mathbf{x}, \omega) \exp(i\omega \mathbf{p} \mathbf{h}) d\mathbf{h}, \qquad (2.1)$$

where $d(\mathbf{h} \,|\, \mathbf{x}, \omega)$ is the recorded data in the frequency domain, $\mathbf{h}$ is offset and $\mathbf{h} = \mathbf{x_r} - \mathbf{x_s}$, $\mathbf{x}$ is the receiver position for a shot gather and it is the source position for a

receiver gather, $\tilde{d}(\mathbf{p} \mid \mathbf{x}, \omega)$ is the transformed plane-wave gather for a certain offset. The inverse plane-wave transform can be defined as

$$d(\mathbf{h} \mid \mathbf{x}, \omega) = \omega^2 \int \tilde{d}(\mathbf{p} \mid \mathbf{x}, \omega) \exp(-i\omega \mathbf{px}) d\mathbf{p}, \qquad (2.2)$$

For the interferometry problem considered here, plane-wave transform over either absolute receiver position $\mathbf{x_r}$ or absolute source position $\mathbf{x_s}$ more helpful. The forward and inverse transform of a shot gather are given by

$$\tilde{d}(\mathbf{p_r} \mid \mathbf{x_s}, \omega) = \int d(\mathbf{x_r} \mid \mathbf{x_s}, \omega) \exp(i\omega \mathbf{p_r x_r}) d\mathbf{x_r}, \qquad (2.3)$$

$$d(\mathbf{x_r} \mid \mathbf{x_s}, \omega) = \omega^2 \int \tilde{d}(\mathbf{p_r} \mid \mathbf{x_s}, \omega) \exp(-i\omega \mathbf{p_r x_r}) d\mathbf{p_r}. \qquad (2.4)$$

where $d(\mathbf{x_r} \mid \mathbf{x_s}, \omega)$ is a shot record with source coordinate $\mathbf{x_s}$ and receiver coordinate $\mathbf{x_r}$. Similarly, for a receiver gather, we can use the following formulas

$$\tilde{d}(\mathbf{p_s} \mid \mathbf{x_r}, \omega) = \int d(\mathbf{x_s} \mid \mathbf{x_r}, \omega) \exp(i\omega \mathbf{p_s x_s}) d\mathbf{x_s}. \qquad (2.5)$$

$$d(\mathbf{x_s} \mid \mathbf{x_r}, \omega) = \omega^2 \int \tilde{d}(\mathbf{p_s} \mid \mathbf{x_r}, \omega) \exp(-i\omega \mathbf{p_s x_s}) d\mathbf{p_s}. \qquad (2.6)$$

A general representation of the Green's function retrieval process between two positions is based on the Rayleigh reciprocity theorem (Wapenaar 2004, Schuster 2009). In this dissertation, I use a far-field approximation and ignore the absolute amplitude. The retrieval of seismic response using interferometry relation in the frequency domain can be written as (e.g., Wapenaar et al. 2010a)

$$d(\mathbf{x_B} \mid \mathbf{x_A}, \omega) + d^*(\mathbf{x_B} \mid \mathbf{x_A}, \omega) \approx \int d^*(\mathbf{x_A} \mid \mathbf{x}, \omega) d(\mathbf{x_B} \mid \mathbf{x}, \omega) d\mathbf{x}, \qquad (2.7)$$

where $d^*(\mathbf{x_A} \mid \mathbf{x_B}, \omega)$ represents the complex conjugate seismic wavefield with $\mathbf{x_B}$ and $\mathbf{x_A}$ being the source and receiver coordinate, respectively.

Using equation 2.3 or equation 2.5, decomposition of seismic wavefield $d(\mathbf{x_A} \,|\, \mathbf{x_B}, \omega)$ into frequency-ray parameter ($\omega - \mathbf{p}$) domain can be accomplished by summing the data over all the $\mathbf{x}$ positions, i.e.,

$$\tilde{d}(\mathbf{p} \,|\, \mathbf{x_A}, \omega) = \int d(\mathbf{x_A} \,|\, \mathbf{x}, \omega) \exp(i\omega\mathbf{px})d\mathbf{x}, \tag{2.8}$$

where the ray parameter $\mathbf{p}$ depends on the survey position $\mathbf{x_A}$. Similar to equation 2.4 or 2.6, the inverse transform is defined as

$$d(\mathbf{x_A} \,|\, \mathbf{x}, \omega) = \omega^2 \int \tilde{d}(\mathbf{p} \,|\, \mathbf{x_A}, \omega) \exp(-i\omega\mathbf{px})d\mathbf{p}. \tag{2.9}$$

We can also define a plane-wave transform by summing over all offsets, but here we need the absolute survey positions to derive interferometry relations in the plane-wave domain. After this transform, the complex conjugate wavefield $d(\mathbf{x_A} \,|\, \mathbf{x}, \omega)$ in the frequency domain is given by

$$d^*(\mathbf{x_A} \,|\, \mathbf{x}, \omega) = \omega^2 \int \tilde{d}^*(\mathbf{p} \,|\, \mathbf{x}_A, \omega) \exp(i\omega\mathbf{px})d\mathbf{p}. \tag{2.10}$$

Inserting equation 2.10 into equation 2.7, we obtain

$$\begin{aligned} & d(\mathbf{x_B} \,|\, \mathbf{x_A}, \omega) + d^*(\mathbf{x_B} \,|\, \mathbf{x_A}, \omega) \\ & \approx \omega^2 \iint \tilde{d}^*(\mathbf{p} \,|\, \mathbf{x_A}, \omega) d(\mathbf{x_B} \,|\, \mathbf{x}, \omega) \exp(i\omega\mathbf{px})d\mathbf{x}d\mathbf{p}. \end{aligned} \tag{2.11}$$

Notice that $\int d(\mathbf{x_B} \,|\, \mathbf{x}, \omega) \exp(i\omega\mathbf{px})d\mathbf{x}$ is a forward plane-wave transformation. Thus, we have

$$\begin{aligned} & d(\mathbf{x_B} \,|\, \mathbf{x_A}, \omega) + d^*(\mathbf{x_B} \,|\, \mathbf{x_A}, \omega) \\ & = \omega^2 \int \tilde{d}^*(\mathbf{p} \,|\, \mathbf{x_A}, \omega) \tilde{d}(\mathbf{p} \,|\, \mathbf{x_B}, \omega)d\mathbf{p}. \end{aligned} \tag{2.12}$$

When the absolute amplitude is not taken into account, the ω^2 filter ($\partial^2 / \partial t^2$ in the time domain) in the right-hand side can be considered as processing filtering effect and can be ignored (Claerbout, 1985). However, this filter is necessary to match the

frequency spectrum of the original data. Thus, equation 2.12 in the $\tau - \mathbf{p}$ domain is equivalent to

$$
\begin{aligned}
&d(\mathbf{x}_\mathrm{B} \mid \mathbf{x}_\mathrm{A}, t) + d(\mathbf{x}_\mathrm{B} \mid \mathbf{x}_\mathrm{A}, -t) \\
&\approx \frac{\partial^2}{\partial t^2} * \int \widetilde{d}(\mathbf{p}_\mathrm{A} \mid \mathbf{x}, t) \otimes \widetilde{d}(\mathbf{p}_\mathrm{B} \mid \mathbf{x}, t) d\mathbf{p},
\end{aligned}
\tag{2.13}
$$

where $\otimes$ denotes crosscorrelation.

Equation 2.13 suggests that analogous to $t - \mathbf{x}$ domain, seismic interferometry can be performed in the $\tau - \mathbf{p}$ domain by crosscorrelating two $\mathbf{p}$ traces which have the same $\mathbf{p}$ value but are from different positions. Following a procedure of summing over all $\mathbf{p}$ traces, this will produce a virtual trace between the two points.

Suppose that the input seismic data is a common shot gather. The steps to perform seismic interferometry in the tau-p domain can be described as follows:

1. Input t-x domain shot gathers and sort them into common receiver gathers $d(t, \mathbf{r}, \mathbf{s})$.

2. Transform the time-space domain data $d(t, \mathbf{r}, \mathbf{s})$ into plane wave data $d(\tau, \mathbf{r}, \mathbf{p})$ using equation 2.8.

3. Select master trace A and slave trace B based on the fixed receiver positions A and B. Then decide the range of $\mathbf{p}$ values to perform crosscorrelation and sum over all $\mathbf{p}$ values. And then apply an ω^2 filter of the data if necessary. This results in a virtual seismic data that looks like A is the virtual source and B is the virtual receiver.

4. Loop step 3 over all the receivers and output redatumed seismic data for further processing, e.g. migration.

Plane wave interferometry can be understood by considering a simple example (Figure 2.1). Interferometric redatuming in the $t - \mathbf{x}$ domain can be explained by cancelation of overlapping stationary ray paths recorded at different receivers. The Green's function between receivers B and A is retrieved by crosscorrelating a ray from

16

source S with cancelation of the common path. In the plane-wave domain, it may be more physically acceptable using asymptotic ray explanations. There are several raypaths from source S to receivers B and A with many ray parameters; stationary condition in interferometry suggests that only those that have the same source ray parameter can lead to cancelation of overlapping ray paths (e.g., Schuster 2009). Thus, interferometry in the plane-wave domain is conducted by selecting the same source ray parameters from receivers B and A for crosscorrelation. Simply put, interferometry in the $t-\mathbf{x}$ domain implies finding common path automatically from a given source-receiver geometry; in the plane-wave domain this common path corresponds to the same ray parameter.
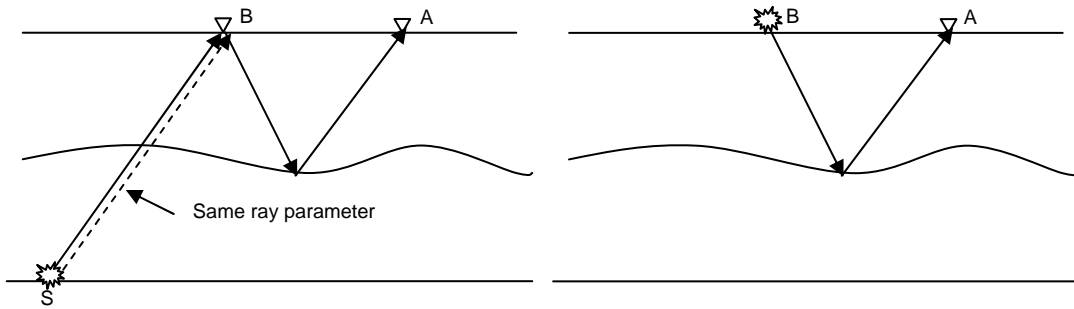
Figure 2.1: Crosscorrelating traces recorded at receiver locations B and A produces a new trace with virtual source at B, receiver at A. The canceled overlapping raypaths in the plane-wave domain corresponds to the same ray parameter.

2.4 EXAMPLES

2.4.1 Transmission to reflection retrieval

My first example is a transmission to reflection retrieval example. The 2D acoustic subsurface model is shown in Figure 2.2 (modified from Draganov et al., 2006). This model consists of an irregular sea floor and a dipping layer. Six hundred transient sources are located at $z = 2.0\,\mathrm{km}$ with an equal distance of 10m; 600 receivers are laterally positioned at the surface. After redatuming, we will get virtual surface

17

acquisitioned marine data with both sources and receivers on the surface for this geometry.

Original shot gathers are obtained numerically by 2-D finite difference modeling with absorbing boundary conditions. They are then sorted into common receiver gathers. The $\tau - \mathbf{p}$ transformed data using equation 2.2 for the first receiver gather (located at (0,0)) is shown in Figure 2.3 (right). The ray parameters range from -0.6 s/km to 0.6 s/km. Full wave information from $t - \mathbf{x}$ gather (Figure 2.3 left) is contained in this transformed gather. Interferometry is then performed over the $\tau - \mathbf{p}$ gather using step 3 and step 4.
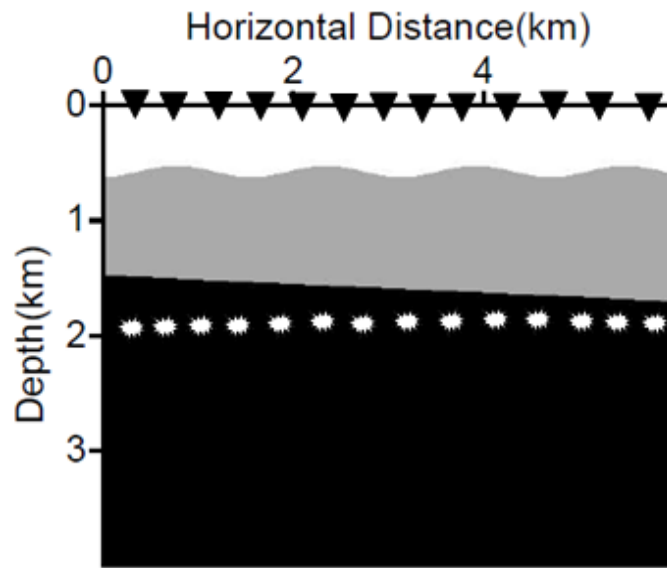
Figure 2.2: An acoustic velocity structure for transmission to reflection retrieval for marine data. This model has a water layer with velocity=1.5 km/s, an irregular sea floor followed by a layer with velocity=2.0 km/s, velocity of the dipping layer is 3.0 km/s.

An example crosscorrelogram of this model is shown in Figure 2.4. For $t - \mathbf{x}$ domain interferometry, this is obtained by first selecting a master trace with receiver

18

located at (3km, 0) for the first subsurface source and a slave trace with source and receiver at the same location, and then correlating these two traces. This procedure is repeated for all 600 sources to produce 600 correlation traces. For plane-wave interferometry, this is obtained simply by selecting the same ray parameter from $\tau - \mathbf{p}$ transformed common receiver gather, autocorrelating it and then repeating for another ray parameter. While all 600 sources for $t - \mathbf{x}$ domain interferometry needed to correlate to produce the correlation panel (Figure 2.4 left), for $\tau - \mathbf{p}$ domain interferometry (Figure 2.4 right), I can observe that most of the energy is produced within a certain range of ray parameters (about -0.3 s/km to 0.3 s/km for this case). This range corresponds to the stationary-phase region (Draganov et al., 2006). Also, note that the range of ray parameters in the correlation panel is smaller than the range of ray parameters in the receiver gather (about -0.5 s/km to 0.5 s/km).
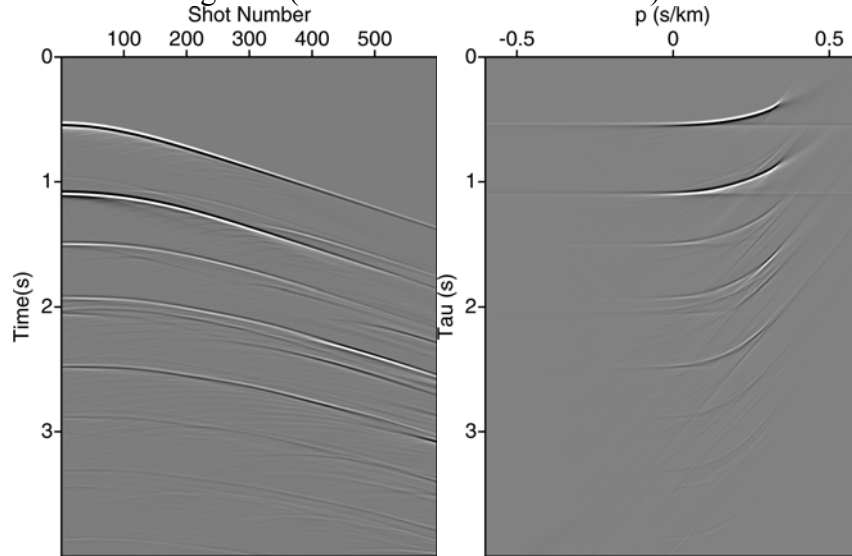
Figure 2.3: Common receiver gather for the first receiver (left) and its tau-p transformed gather (right). There are 200 ray parameters in the plane-wave gather from -0.6 to 0.6 s/km.

Stacking of the crosscorrelogram in Figure 2.4 results in a reconstructed trace with source and receiver at the same surface location (3km, 0). From the crosscorrelogram, I can observe that with $\tau - \mathbf{p}$ domain interferometry, there is some diffracted energy at the edge of the panel because of the finite aperture of seismic acquisition. This is because in reality, ideal interferometric geometries, which require that the medium be completely surrounded by physical sources (Wapenaar, 2004), are never satisfied. After stacking, these diffractions cancel out and the zero-offset trace using $\tau - \mathbf{p}$ domain interferometry and t-x domain interferometry produce nearly identical results, both in travel time, normalized amplitude and frequency content (Figure 2.5).
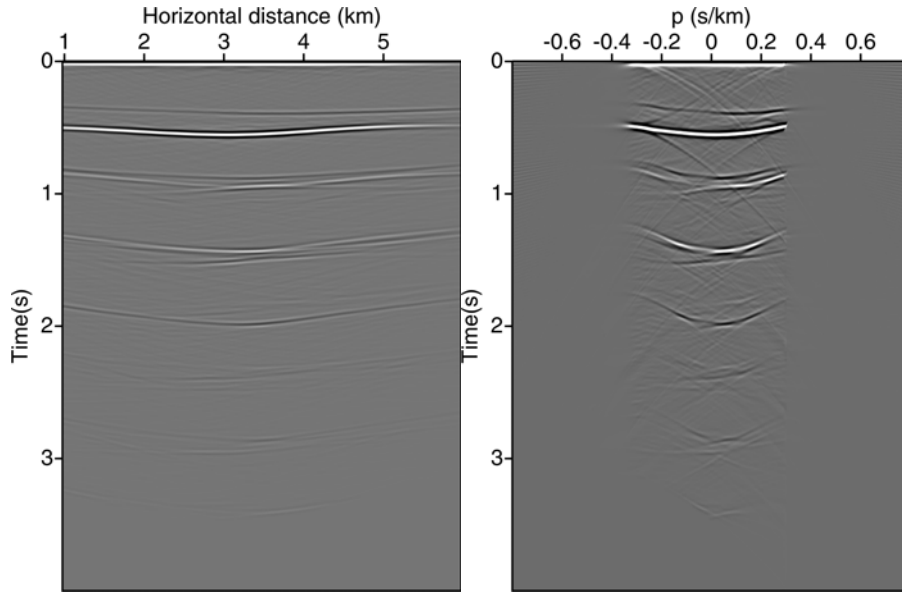
Figure 2.4: The crosscorrelogram for a master trace recorded at (3km,0) with a slave trace recorded at the same location for each subsurface position (only positive time is shown). Left: with time-space domain interferometry. Right: with plane-wave domain interferometry.

Figures 2.6a and 2.6b show the retrieved surface reflection data for one virtual shot gather using interferometry in these two different domains. The virtual source is

20

located at (3km, 0) and virtual receivers are located along the surface so that I can see the symmetric two-sided seismic response. The two plots are generated using the same gain. There are no noticeable differences between the two except for the absolute amplitude. This suggests that when the distribution of sources and receivers is satisfied for interferometric redatuming, time-space domain interferometry and plane-wave domain interferometry should produce similar results. When the acquisition condition is satisfied, correlating the data using all available ray information at one receiver location with another receiver location (time-space domain) should give similar result as those obtained by correlating data that has the same ray parameter (plane-wave domain approach).
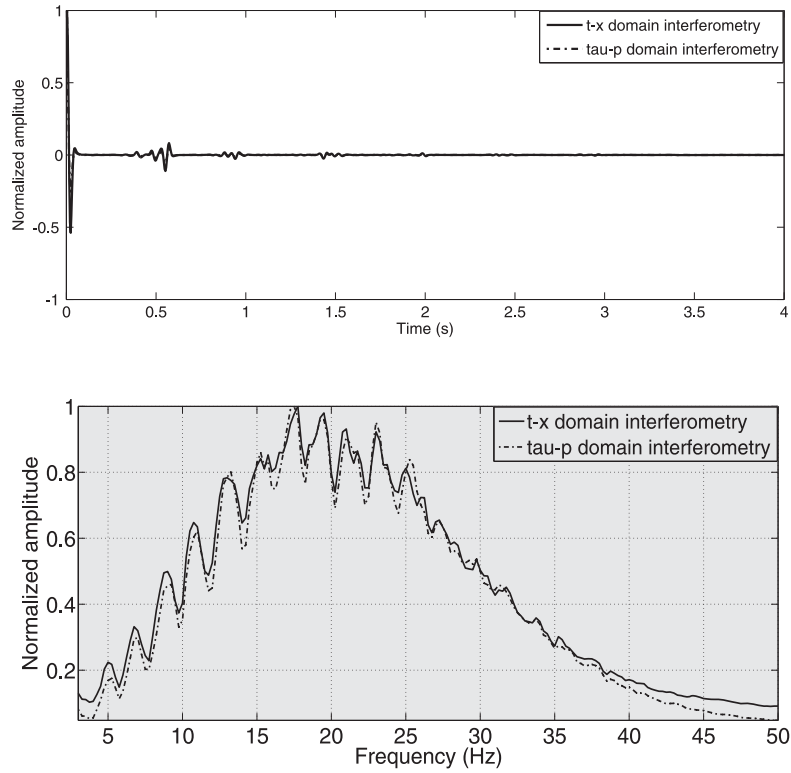
Figure 2.5: Comparison of a virtual zero-offset event with both source and receiver located at (3km, 0). Top: wiggle comparison. Each wiggle is obtained by a summation of the cross-correlogram in Figure 2.4. Bottom: the corresponding frequency spectrum.

21
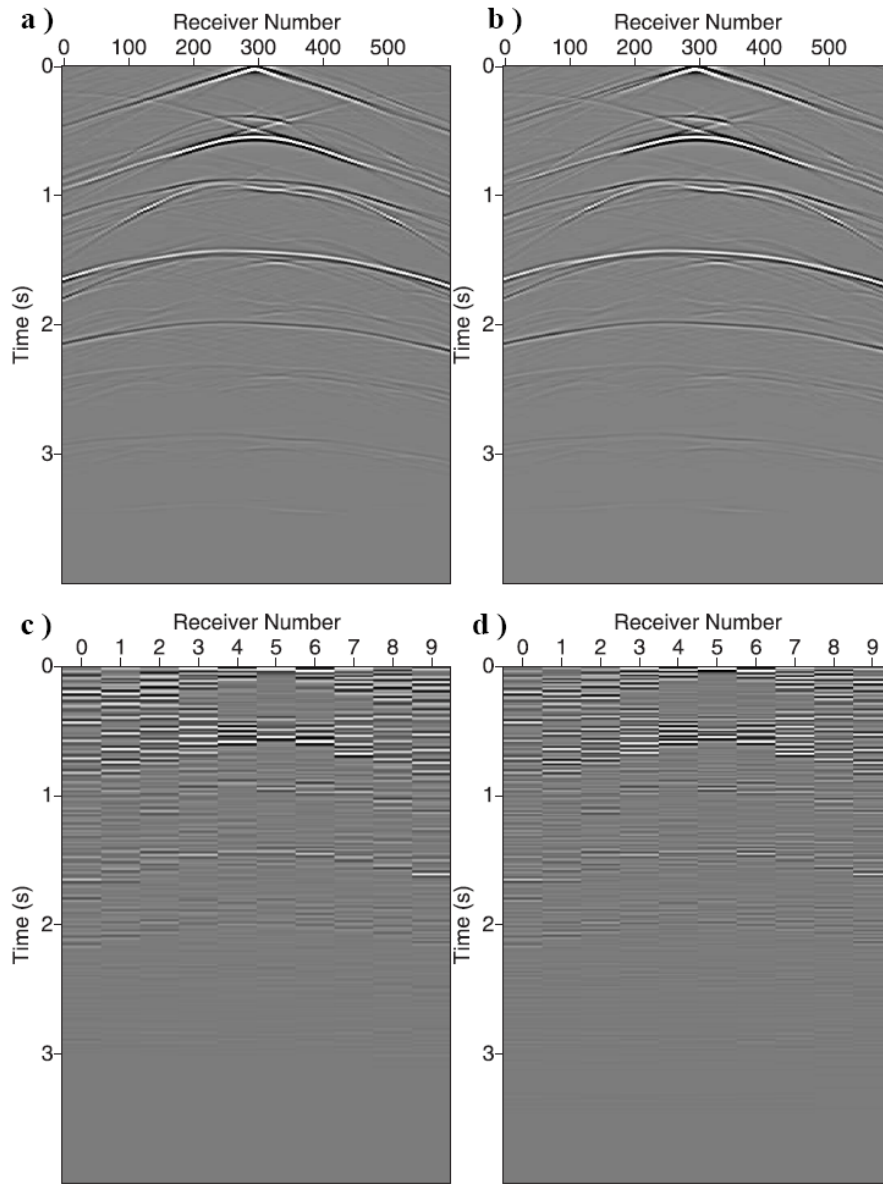
Figure 2.6: (a) Redatumed shot record using time-space domain interferometry for a virtual source located at (3km, 0) and receivers along the surface. (b) redatumed shot record using plane-wave domain interferometry. (c) redatumed shot record using time-space domain interferometry for only 10 real transient sources and 10 real receivers. (d) redatumed shot record using plane-wave domain interferometry. (c) and (d) has the same geometry as (a) and (b), except that the data is sparsely sampled.

When seismic data do not have adequate spatial sampling and sufficient acquisition aperture, $\tau - \mathbf{p}$ transform is known to suffer from edge effects and numerical aliasing. However, seismic interferometry requires a range of ray parameters and when stacked together, such artifacts are not noticeable (Figures 2.6 (c) and (d)).

One particular use of plane-wave domain interferometry is to select the range of ray parameters, which has the most dominant energy. This is useful when the recorded seismic data consist of significant amount of noise, e.g., random noise and surface-wave noise in land data. These noises should be suppressed because they might contribute to spurious arrivals in the redatumed events. Similar idea of retrieving seismic data using only dominant energy was also proposed by Vidal et al., (2011). However, their approach is based on the plane-wave transform of the virtual-source panel, which differs from my approach that is based on direct transform of the recorded data.

I demonstrate this approach by simply adding random noise to the original recorded data. The synthetic shot gather is shown in Figure 2.7a. By analyzing the correlogram in the plane-wave domain (Figure 2.7b), we can observe that most of the dominated ray parameters are between -0.2 s/km to 0.2 s/km. Figure 2.7c shows the redatumed result using full range of ray parameters. This retrieval results in reflections and diffractions. While for this case the diffraction arrivals are true in the real modeled gather, in reality, we may be interested in retrieving reflections only. Figure 2.7d shows the results using a set of range-selected ray parameters. Reflections are correctly retrieved while the diffractions are suppressed.
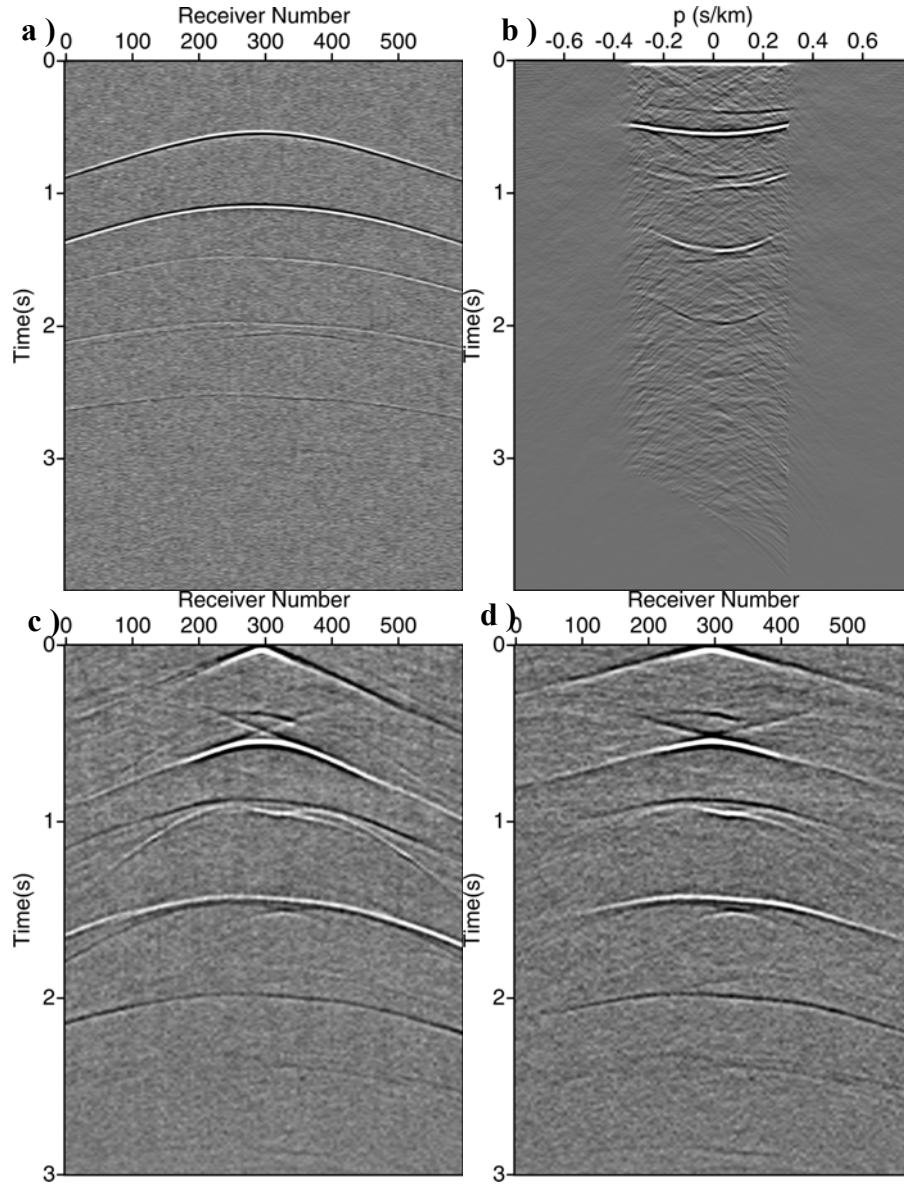
Figure 2.7: (a) Synthetic noisy shot gather. (b) crosscorrelogram in the plane-wave domain. (c) redatumed virtual shot gather using full range of ray parameters. (d) redatumed virtual shot gather using only ray parameters between -0.2 s/km and 0.2 s/km.

2.4.2 VSP redatuming

Another successful application of seismic interferometry is to redatum VSP data to SSP data or single well profile (SWP) data (Schuster 2009). For depth sampled data, $\tau - \mathbf{p}$ transform can be used to map the data $d(t,z)$ from time-depth domain to the plane-wave domain similar to surface seismic data. The forward and inverse transforms can be defined as:

$$\tilde{d}(\mathbf{p},\omega) = \int d(z,\omega)\exp(i\omega\mathbf{p}z)dz, \tag{2.14}$$

$$d(z,\omega) = \omega^2 \int \tilde{d}(\mathbf{p},\omega)\exp(-i\omega\mathbf{p}z)d\mathbf{p}. \tag{2.15}$$
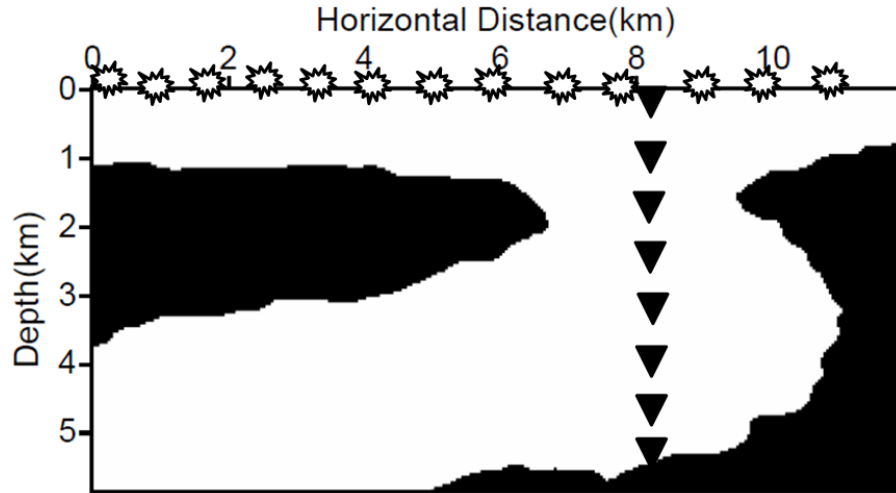
Figure 2.8: Walk-away VSP acquisition geometry for a synthetic test. The model is comprised of a background velocity of 2000 m/s with two salt bodies of velocity 4480 m/s.

Using this transform, seismic interferometry can be performed in the plane-wave domain without any further assumptions to redatum the VSP data to either SSP or SWP data. Here I use a 2-D synthetic model (Figure 2.8) modified from Lu et al., (2008) to redatum walk-away VSP data to SWP data (borehole sources and borehole receivers). For simplicity, I remove the background velocity gradient of the original velocity profile

and use a constant background velocity instead. The objective here is to image the synthetic salt dome with velocities similar to those in Gulf of Mexico salt domes.

Figure 2.9 compares retrieved SWP data using time-space domain interferometry and plane-wave domain interferometry with $\mathbf{p}$ values ranging from -0.6 s/km to 0.6 s/km. I can observe that despite some filtering effects present in the $\tau - \mathbf{p}$ domain approach, these two methods in general produce nearly identical results.
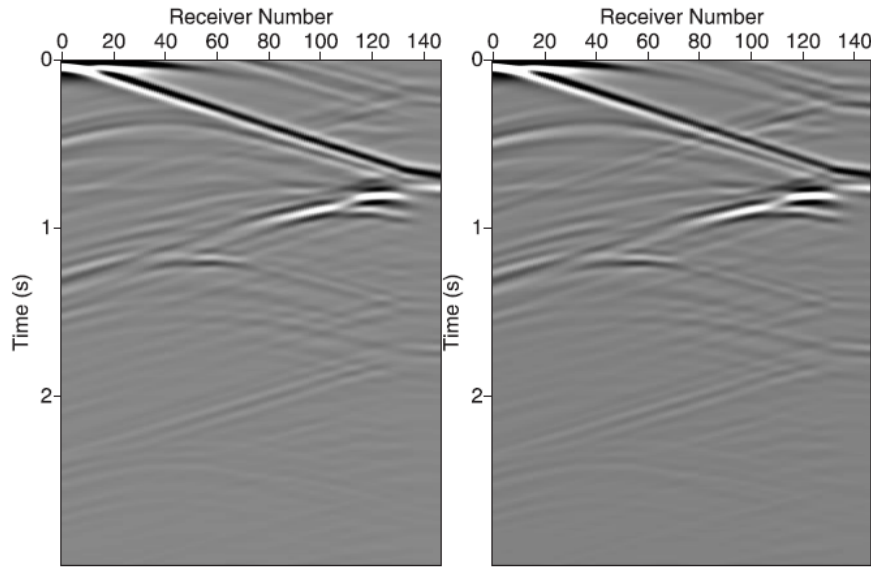
Figure 2.9: Comparison of redatumed downhole common shot gather using time-space domain interferometry (left) and plane-wave domain interferometry (right) for a virtual source located at the surface and receivers in the borehole.

Another application of $\tau - \mathbf{p}$ domain interferometry for VSP geometry is to separate left and right propagating waves using positive and negative ray parameters. Wavefield separation has already been used in seismic interferometry with up-down separation (Mehta et al., 2007). For VSP geometry in Figure 2.8, if we want to image the salt dome on the right of the model, then most of the stationary-phase points are contributions by negative (or positive depending on definition) $\mathbf{p}$ values. For the model shown in Figure 2.10, redatuming using full wave represents all the reflected data from

the whole subsurface geology. Waves coming from different parts of the model could intertwine with each other, and correspondingly affect the migrated image. With interferometry in the plane-wave domain, I can easily select only the ray parameters of interest to produce directional redatuming.
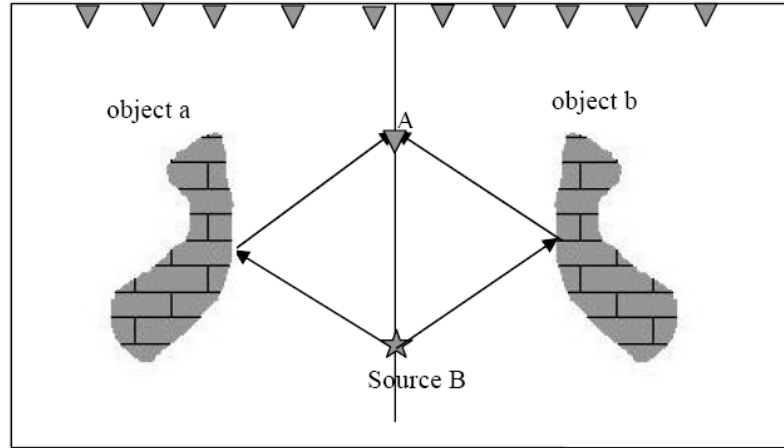
Figure 2.10: A model shows the limitation of conventional interferometry. There are different ray paths from A to B, which can be either from the left object a or the right object b. If we migrate this data, we could not effectively delineate both of the objects.

Figure 2.11 (left) compares redatumed shot gather using only negative **p** values. The result is similar to that using the entire range of **p** values and is also similar to the result obtained by directly putting a source and receivers in the borehole (Figure 2.11 right). After redatuming, I performed reverse time migration of the SWP data. The migration results with the full range of ray parameters and only negative ray parameters are shown in Figure 2.12 left and Figure 2.12 right, respectively. Both of these approaches could capture the edge of the salt dome correctly. I cannot observe many differences for this model, suggesting that I can effectively delineate the salt body with only a few ray parameters. Also, migration with only negative ray parameters might get a

slightly better image, e.g., it can suppress some anomalies at a depth of about 5km when using full range of ray parameters.
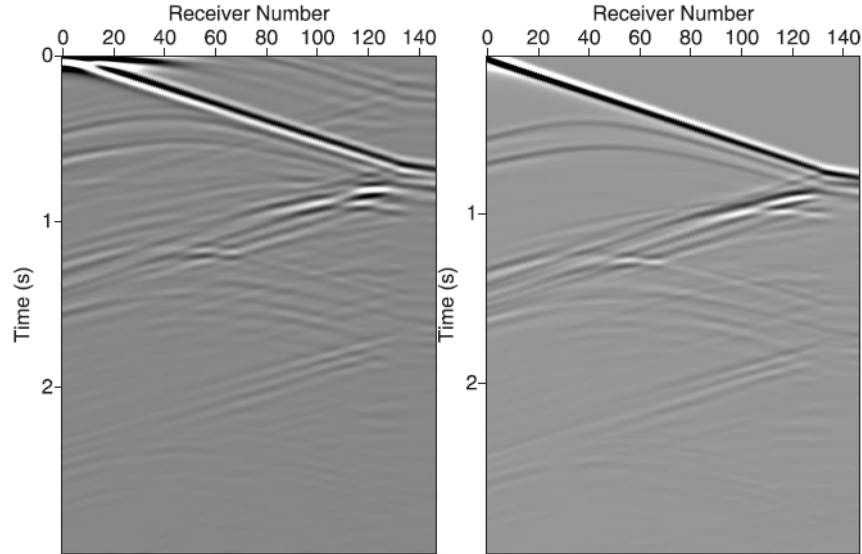
Figure 2.11: Comparison of plane-wave domain interferometric redatumed record using only negative ray parameters (left) and directly putting source at surface and receivers in the borehole (right).
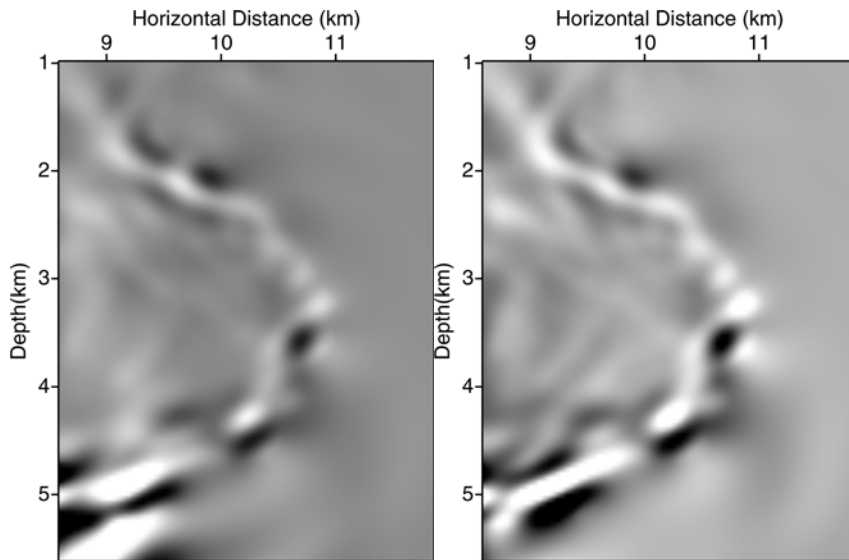
Figure 2.12: Comparison of depth migrated redatumed data using full range of ray parameters (left) and using only negative ray parameters (right).

28

2.4.3 Super-virtual interferometry

Super-virtual interferometry is a novel technique to increase the signal-to-noise ratio (SNR) of far-offset refracted waves (Bharadwaj et al., 2012). With this technique, the receiver spread of the refraction survey can be fully utilized if we are only interested in the travel time information. Virtual far-offset refraction arrivals are generated by first a crosscorrelation between adjacent refracted wave traces and summation to produce a virtual trace, and then by a convolution with the actual refraction traces. Mathematically, without retrieving the absolute amplitude, these two steps can be written as follows

$$\psi(\mathbf{x_A} \mid \mathbf{x_B}, t)^{virt} \approx \int\limits_{source} \psi^*(\mathbf{x_A} \mid \mathbf{x}, t) \otimes \psi(\mathbf{x_B} \mid \mathbf{x}, t) d\mathbf{x}, \tag{2.16}$$

$$\psi(\mathbf{x_B} \mid \mathbf{x_A}, t)^{super} \approx \int\limits_{receiver} \psi(\mathbf{x_B} \mid \mathbf{x'}, t)^{virt} \psi(\mathbf{x_A} \mid \mathbf{x'}, t) d\mathbf{x'}. \tag{2.17}$$

where $\psi(\mathbf{x_A} \mid \mathbf{x_B}, t)$ denotes the head wave contribution in the recorded data $d(\mathbf{x_A} \mid \mathbf{x_B}, t)$, $\psi(\mathbf{x_B} \mid \mathbf{x_A}, t)^{virt}$ is the virtual data by stacking the common receiver pair gather (CPG) (Dong et al., 2006), $\psi(\mathbf{x_B} \mid \mathbf{x_A}, t)^{super}$ is the processed super-virtual head wave data. Note that the above relation works both for P-wave and shear wave.

Virtual traces in the CPG gather are useful to distinguish between head wave and diving wave. This is because head wave corresponds to flat events while the diving waves are not flat. Equation 2.16 has a formula similar to that used in a conventional interferometry relation (equation 2.7). Thus, the plane-wave domain approach is also applicable for generation of the CPG gather.

Figure 2.13 shows synthetic data in three panels containing a head wave, a reflection and both. In the super-virtual interferometry, only head wave arrivals are treated as correlated useful information. While I want to avoid artifacts by windowing only head wave arrivals, for real problems, presence of other arrivals such as reflections

and noise will result in spurious arrivals of the processed data. Thus, I generate synthetic data for four different scenarios; corresponding CPG gather using a time-space domain approach is shown in Figure 2.14. In comparison, CPG gather using a plane-wave based approach is shown in Figure 2.15.
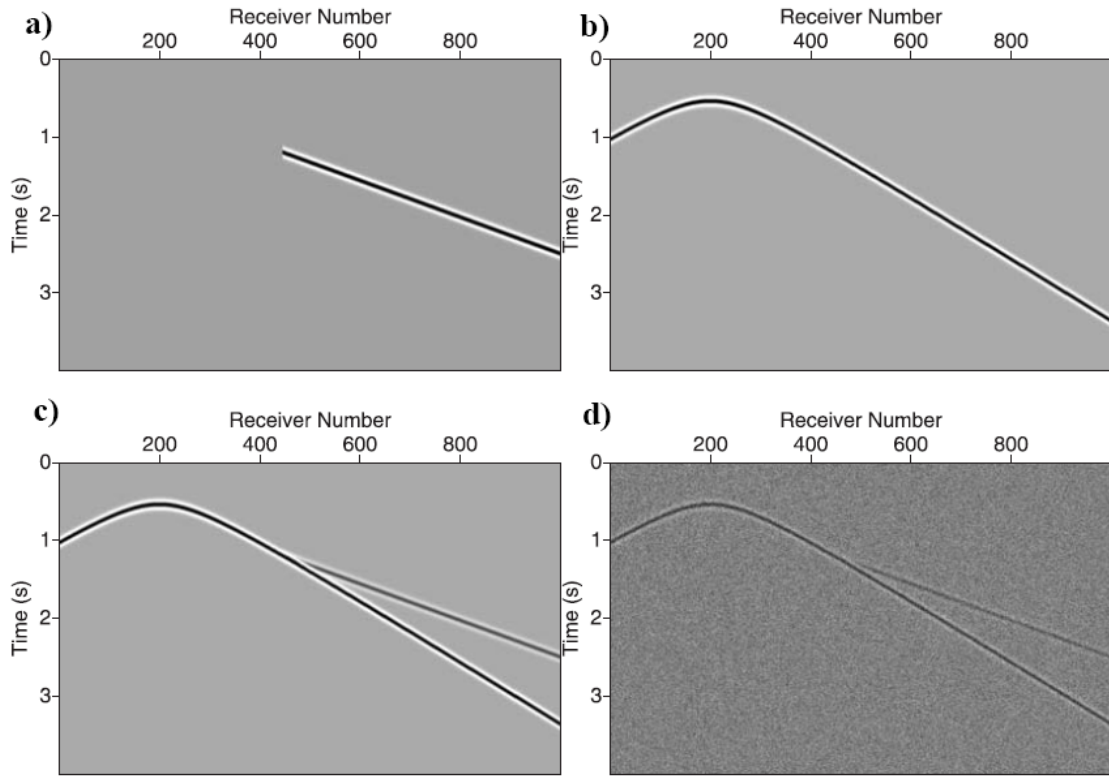
Figure 2.13: (a) Synthetic head wave. (b) synthetic reflection. (c) synthetic reflection plus head wave with a smaller amplitude. (d) similar to (c), but random noise is added.

For the pure head wave case (Figure 2.13a), the CPG gather shows a perfectly flat event with a time-space domain approach (Figure 2.14a) and a focused point with a plane-wave domain approach (Figure 2.15a). For pure reflection case in Figure 2.13b, the far-offset reflection in the CPG gather shows a nearly flat event (Figure 2.14b), which

30

may be falsely identified as a refraction arrival. However, with a plane-wave domain approach, there is no focusing point for the reflection arrival (Figure 2.15b). For the case with both reflection and refraction (Figure 2.13c), with a time-space domain approach, the refraction event in the CPG gather (Figure 2.14c) is only in the white box area. Thus, it requires careful filtering and muting to get rid of reflection arrivals.
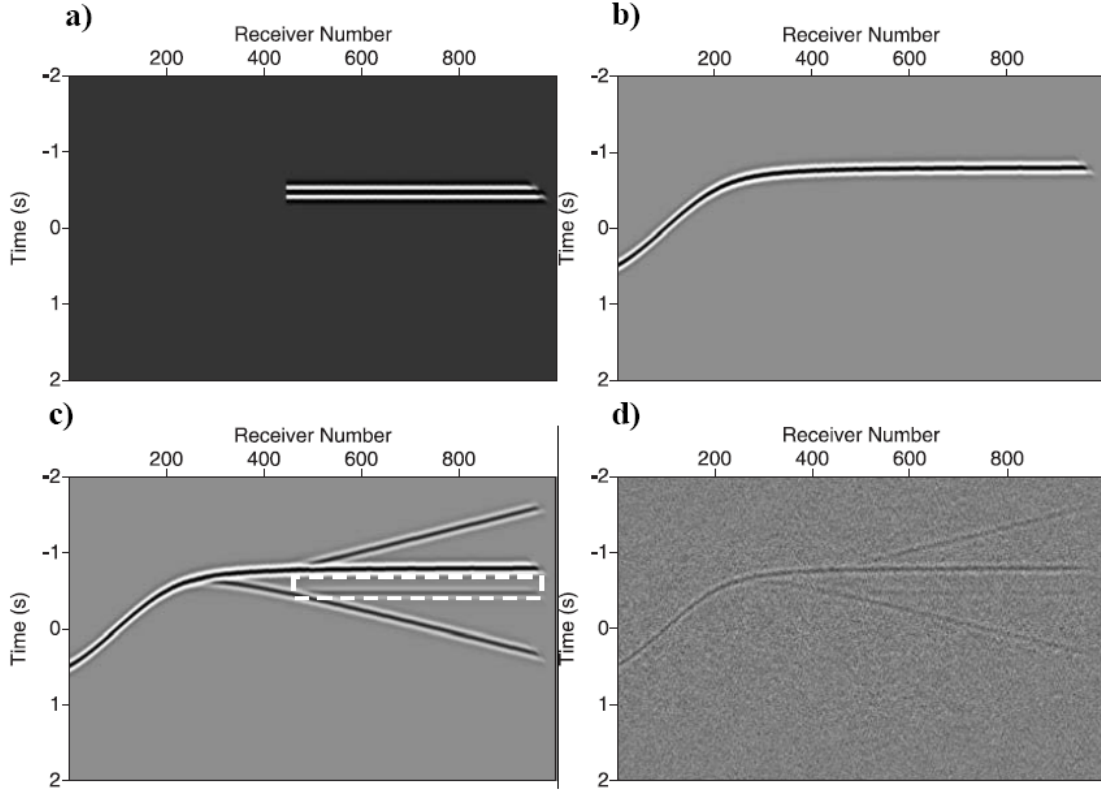
Figure 2.14: CPG gather for the four models in Figure 13. It is generated by correlating two different common receiver gathers in the time-space domain. (a) head wave. (b) reflection. (c) head wave plus reflection. Head wave contribution is denoted by a white box. (d) similar to (c), but with random noise.

In comparison, with a plane-wave domain approach, we could easily identify the focusing point in Figure 2.15c. Thus, we can easily suppress the influence of reflection arrivals. In addition, if we zoom in, we can see a flat event passing through the focusing point. Therefore, we can choose a range of ray parameters to stack the CPG gather. When

31

the data are noisy (Figure 2.13d), refraction events in the CPG gather (Figure 2.14d) may be difficult to identify since the amplitude is very small. In comparison, however, with a plane-wave domain approach, we can suppress the noise in the CPG gather (Figure 2.15d). This is because plane-wave transform involves stacking along ray parameters, which can attenuate random noise.
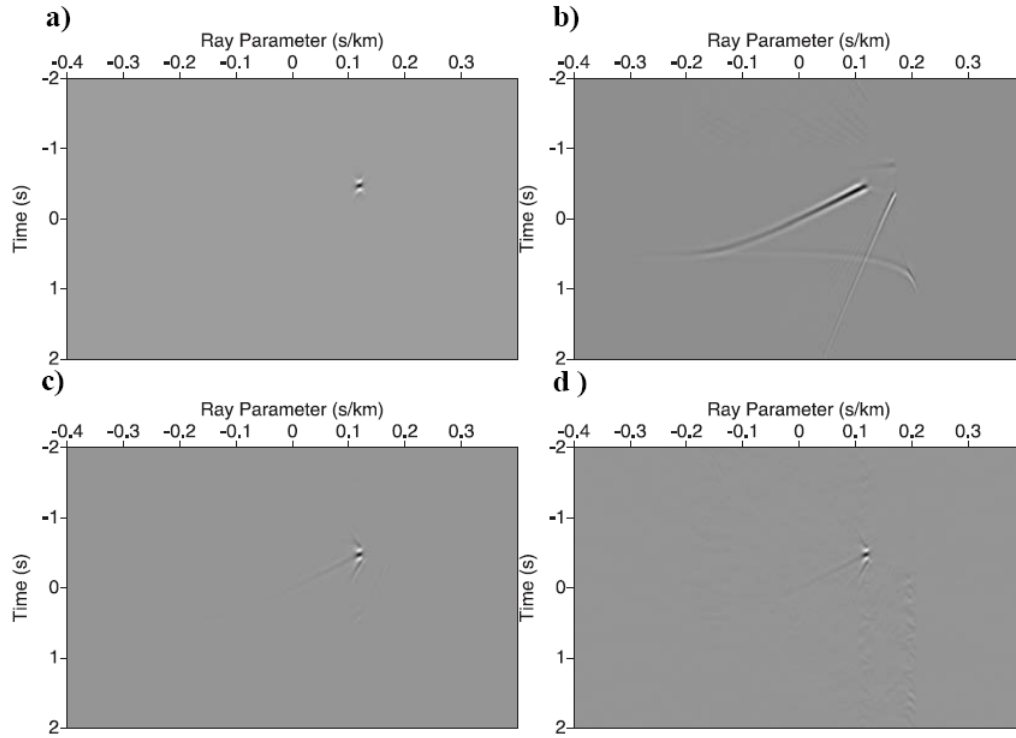
Figure 2.15: CPG gather generated with a plane-wave based approach. No filtering is applied. (a) head wave. (b) reflection. (c) head wave plus reflection. (d) similar to (c), but with random noise.

I apply this technique to one line (line 23) of Ocean Bottom Seismometer (OBS) survey data across Taiwan and the western Philippine Sea (McIntosh et al., 2005). The data are collected in four components (three component gimbaled 4.5 Hz geophones and a hydrophone). The seismic lines are shot with 100m spacing with a maximum of about

1500 shots. Initial data processing includes clock-drift correction, band-pass filtering and rotation of the horizontal component to radial and transverse components.
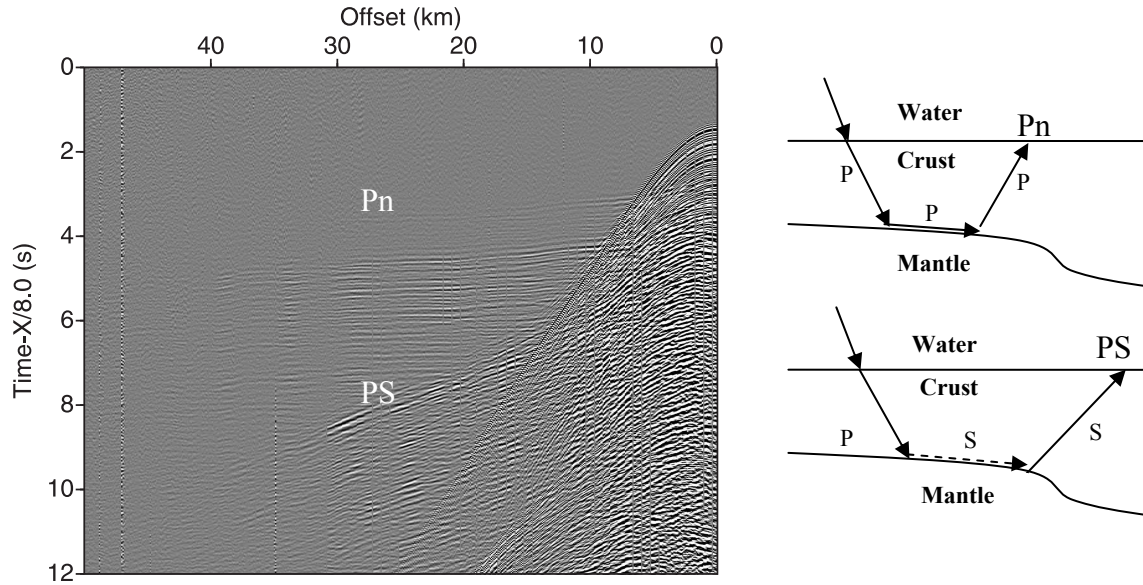
Figure 2.16: An example OBS data. Horizontal component is shown. Two major refraction arrivals are Pn (refracted in the subducting slab) and PS (refracted shear wave). Right diagram shows the the ray-paths of the two different arrivals.

Figure 2.16 shows a vertical geophone component at one OBS station (line 23, station number 32). The seismograms are plotted with a reduction velocity of 8.0 km/s, which flattens the Pn arrival. Another notable arrival at near offset is PS arrival, which is partially converted as shear arrival. The P-wave mode of PS arrival has likely converted into shear-mode before arriving at the crust-mantle boundary. However, we cannot make a conclusive judgment because we do not know the shear-wave velocity structure of this area. We can observe that shear-wave data is weaker and thus noisier than primary head waves. This is typical for most OBS surveys.

33

I then window the data around the PS arrivals in each component in each OBS gather with a muting window of about 3.0 s. After that, I cross-correlate traces recorded at two different OBS stations (station 32 and station 33) for each component to generate a CPG gather. A comparison of the radial horizontal component CPG gather obtained with a time-space domain approach and a plane-wave domain approach is shown in Figure 2.17. To increase the coherency, I use the super-virtual gather at one location, which is suggested by Bharadwaj et al., (2011). CPG gather in the plane-wave domain shows a focused refraction event at about 2.5 s and also a flat event because of imperfect focusing.
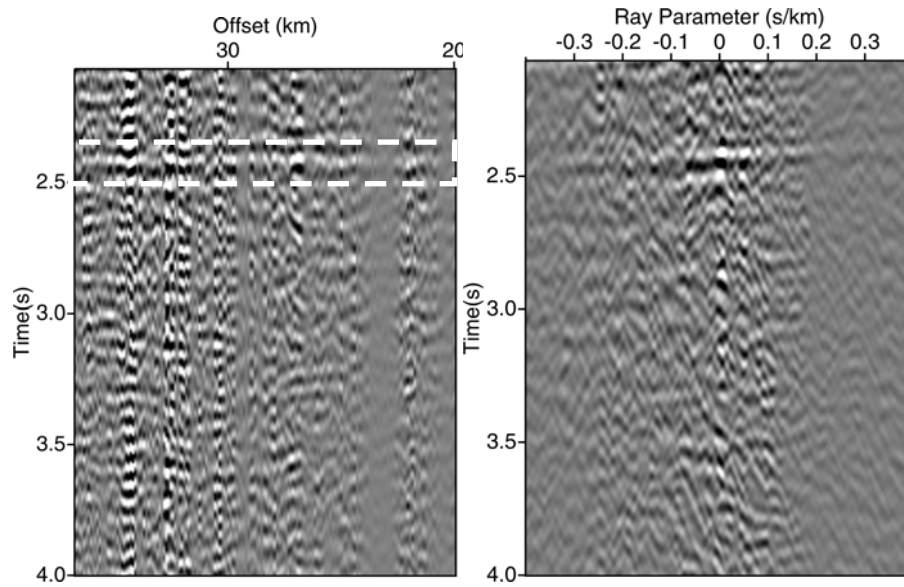
Figure 2.17: Comparison of the CPG gather obtained by correlating the radial horizontal component of the data at two stations. Super-virtual gather at one location is used to increase the coherency. Left: time-space domain approach. The flat event is marked in white box. Right: Plane-wave domain approach.

Figure 2.18 shows a comparison of the original data (station number 32) with data processed with super-virtual interferometry for the horizontal component. The refraction

at far-offset (35km-50km) before processing (Figure 2.18a) is not clear, while in the super-virtual gather it is clearly visible (Figures 2.18b and 2.18c). Results obtained using a time-space CPG gather and plane-wave CPG gather show similar improvements at far offsets. However, the plane-wave approach shows slightly better improvement in terms of the signal-noise ratio.
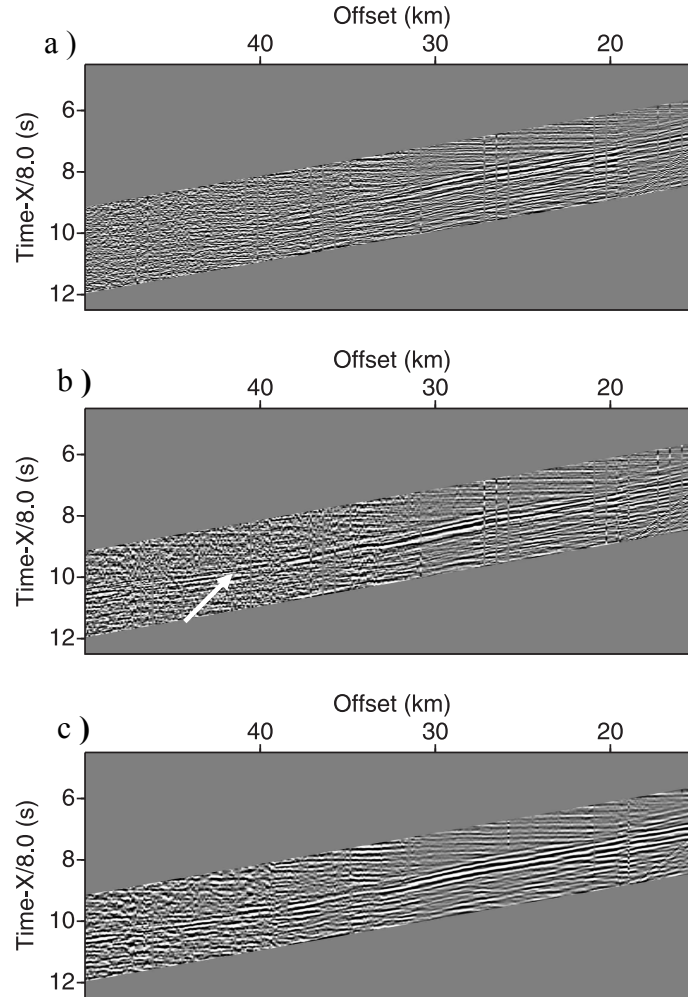
Figure 2.18: Comparison of the original data and processed data with super-virtual interferometry. (a) original data. (b) super-virtual data with a time-space domain CPG gather. (c) super-virtual data with a plane-wave domain approach.

35

2.5 CONCLUSIONS

In this chapter I present a new redatuming method in the plane-wave domain for seismic interferometry. This method can be used in many applications of controlled source interferometry if plane-wave transform is applicable to that data. Mathematically, it can be easily explained by the cancelation of waves that have common ray parameters. Using synthetic data generated with finite differences, we demonstrate that our method is effective in reconstructing virtual seismic responses from SSP and VSP data. We also demonstrate the effectiveness of this method with synthetic refraction data and a real OBS data.

Compared with interferometry in the time-space domain, our method offers several advantages. It transforms irregularly acquired seismic data into a regular coordinate system. It allows us to flexibly choose ray parameters so that we can selectively redatum the events we are mostly interested in and perform directional redatuming. A CPG gather generated with this approach for super-virtual study can help to reduce the effects of undesired phases and random noise. It also can reduce the computational cost if a large number of traces need to be crosscorrelated.

Chapter 3: Frequency-domain full waveform inversion with plane-wave data

3.1 ABSTRACT

In this chapter, I propose an efficient frequency-domain full waveform inversion (FWI) method using plane-wave encoded shot records. The forward modeling involves application of position dependent linear time shifts at all source locations. This is followed by propagation of wavefields into the medium from all shot points simultaneously. The gradient of the cost function needed in the FWI, is calculated first by transforming the densely sampled seismic data into frequency-ray parameter domain and then back-propagating the residual wavefield using an adjoint-state approach. I use a Gauss-Newton framework for model updating. The approximate Hessian matrix is formed with a plane-wave encoding strategy, which requires a summation over source and receiver ray parameters of the Green's functions. Plane-wave encoding considerably reduces the computational burden and cross-talk artifacts are effectively suppressed by stacking over different ray parameters. It also has the advantage of directional illumination of the selected targets. Numerical examples show the accuracy and efficiency of our method.

3.2 INTRODUCTION

During the past several decades, full waveform inversion (FWI) has become increasingly popular for velocity analysis in seismic exploration applications (e.g., Tarantola, 1984; Pratt et al., 1998; Pratt, 1999; Shin and Cha, 2008). This is because FWI utilizes both amplitude and travel time information from band-limited seismic wavefields. The forward problem is solved by the two-way wave equation, which can model full wave propagation effects including internal multiples. The FWI overcomes several limitations of some of the common imaging techniques such as those based on ray

theory and one-way wave equation and can provide high-resolution estimates of the medium parameters.

Mathematically, FWI can be formulated as a nonlinear optimization problem that iteratively minimizes a misfit functional between observed seismic data and synthetic data. Up to now, FWI is implemented in a gradient-type inversion approach instead of a Monte Carlo random search framework. In the frequency domain, the misfit is a summation of residuals over all sources, receivers and frequencies. Thus, although gradient-based methods generally provide fast convergence with a good starting model, they are still computationally expensive mainly because a large number of sources need to be simulated at each iteration. Modern acquisitions with high-fold, short source and receiver interval and wide azimuthal coverage make the size of the dataset very large and therefore, the cost of computation is increased significantly.

The simultaneous source encoding approach has been proposed to reduce the computational burden for the shot profile simulations required for time-domain prestack depth migration and FWI (Romero et al., 2000; Capdeville et al., 2005; Krebs et al., 2009; Boonyasiriwat et al., 2010). This approach is also implemented in the frequency-domain FWI (Herrmann et al., 2009; Ben-Hadj-Ali et al., 2011). The key step of this approach is forming super-shots by summing densely distributed individual shots. A random phase delay is applied to each of the shot records before summation. This, however, results in strong incoherent cross-talk artifacts. The artifacts can be suppressed with stacking the super-shots. However, the performance of this suppression is dependent on the choice of the random shots. This approach is also more sensitive to random noise than conventional FWI method, with the degree of sensitivity dependent on the number of shots assembled (Ben-Hadj-Ali et al., 2011).

An alternative approach to reduce the computational cost is to decompose the densely acquired data into plane-wave domain and to select a limited number of ray parameters for modeling and inversion (Whitmore, 1995; Zhang et al., 2005; Liu et al., 2006; Stoffa et al., 2006). In this approach, a linear phase delay is applied to each shot record to form the super-shot. The process of transforming the raw data into plane-wave gather is also known as slant-stacking since it sums amplitudes along a line of constant slope in the time-space domain (e.g., Stoffa 1989). Cross-talk artifacts are suppressed reasonably well with a sufficient number of ray parameters. The ray parameter is directly related to the take-off angle at the surface location and it can be selected based on an illumination analysis of the targets to be imaged, resulting in further saving of computational time. Plane-wave transformation also has the advantage of easily incorporating anisotropic analysis into the imaging process (Sen and Mukherjee 2003; Sil and Sen, 2009). Better imaging results can be obtained with split-spread receiver gathers, which are generated by the reciprocity principle (Liu et al., 2004).

Vigh and Starr (2008) applied a time-domain plane-wave FWI with applications to synthetic data sets and a real 3D marine data. Compared with the standard FWI, computational time is reduced significantly with comparable quality of the inverted velocity model. Zhang and Wang (2009) proposed a different propagation scheme with a time-domain wave-equation approach applied to a similar plane-wave dataset. This approach has similar computation cost per iteration as the approach proposed by Vigh and Starr (2008).

In this dissertation, we focus on the frequency domain methods and investigate some practical issues. A frequency-domain implementation of FWI remains an attractive approach since the coefficients of the allocated matrix can remain the same for different

shots. This domain also favors a natural choice of different groups of frequencies to avoid cycle skipping and to obtain multi-scale solutions (Virieux and Operto, 2009).

Preconditioning FWI with the Hessian matrix can provide considerable improvement of convergence by scaling the perturbed model (e.g., Pratt et al., 1998). Inversion with the use of the Hessian matrix is known as a Gauss-Newton optimization procedure, which generally takes fewer iterations to attain quadratic convergence than pure gradient methods. Constructing the full Hessian matrix is prohibitively expensive for very large models. An approximate Hessian is normally used for practical applications (Shin et al., 2001). Further reduction in the computational cost can be attained by using the diagonal part of the Hessian only. However, even the diagonal Hessian requires $Ns + Nr$ forward modeling steps, where Ns equals the number of shots and Nr equals the number of receivers. A phase encoding scheme (Tang, 2009; Tang and Lee, 2010) has been proposed to save computational time and disk storage for Hessian calculation.

To improve on the efficiency and accuracy of this full-wave inverse problem, we propose performing FWI with frequency domain plane-wave data. With limited selected ray parameters of plane-wave data, computation time is saved both in forward modeling and back-propagation of residuals for gradient calculation. This approach simplifies the inversion procedure with an effective suppression of cross-talk artifacts. The space-domain Hessian is formed using a plane-wave phase encoding approach, allowing efficient calculation with comparable results as those obtained with the standard method. Unlike random phase-encoding approach, this method is not sensitive to random noise. This approach can also be easily applied to marine geometry, which requires special treatment with a random phase encoding method (Krebs et al., 2009).

40

This chapter is organized as follows. We first review the conventional FWI approach and introduce our plane-wave encoding scheme for gradient and Hessian calculations. Next, we show some numerical examples to demonstrate the validity and efficiency of our method.

3.3 METHODOLOGY

Three primary steps comprise our method. The first involves slant stacking and temporal Fourier transform, which provides a frequency-domain plane-wave gather from recorded seismic data. This step uses data from different sources and compresses them into a regularized super-shot. Second, the synthetic data are simulated with a distance dependent linear time shift at all receiver locations. This delay is determined by the ray parameter in consideration, which can be determined by the take-off angle at the surface location. Similar approach is adopted for back propagation of the residuals. Third, the diagonal Hessian is formed with a plane-wave encoding approach. These are then used in model update in which the optimal step length is determined by a line search. The above process is repeated for multiple groups of frequencies until convergence is reached.

3.3.1 Frequency-domain plane-wave data

For a recorded time-domain shot gather, frequency-domain plane-wave data is obtained first through a plane-wave transformation, also known as $\tau - p$ transform (e.g., Stoffa 1989; Foster and Mosher, 1992). This transform can be implemented either in the time-domain or in the frequency-domain. In the frequency-domain, it involves a phase shift to each trace that is dependent on the offset and ray-parameter. This is followed by a stacking operator as follows

$$\tilde{u}(\mathbf{p}\,|\,\mathbf{x},\omega) = \int u(\mathbf{x_s}\,|\,\mathbf{x},\omega)\exp(i\omega\mathbf{p}(\mathbf{x}-\mathbf{x_s}))d\mathbf{x}\,, \qquad (3.1)$$

where $u(\mathbf{x_s} \mid \mathbf{x}, \omega)$ is the frequency-domain recorded data, $\tilde{u}(\mathbf{p} \mid \mathbf{x}, \omega)$ is the transformed plane-wave gather, $\mathbf{x}$ is the receiver position for a shot gather.

Note that in the time-domain, this transform needs to be modified by applying a time-shift to transform the data from $\tau - p$ domain to "physical" time-ray parameter ($t - p$) domain to match the modeling strategy discussed next (Whitmore, 1995). Also, application of the reciprocity principle can produce a split-spread receiver gather, which helps to overcome the aliasing problem in plane-wave composition (Liu et al., 2006).

In a conventional frequency-domain FWI framework, the forward modeling step is a simulation of the wavefield propagated from a surface location $\mathbf{x_s}$, which for a variable density acoustic medium, is governed by the following equation

$$\nabla \cdot \left(\frac{1}{\rho(\mathbf{x})} \nabla u(\mathbf{x}, \omega) \right) + \omega^2 \kappa(\mathbf{x}) u(\mathbf{x}, \omega) = f(\omega) \delta(\mathbf{x} - \mathbf{x_s}), \qquad (3.2)$$

where $u(\mathbf{x}, \omega)$ represents the scalar pressure wavefield in the frequency domain, $\rho(\mathbf{x})$ is the mass density, $\kappa(\mathbf{x})$ is the compressibility, and $f(\omega)$ is the frequency domain source signature.

After discretizing the model with a finite difference or finite element scheme, equation 3.2 can be rewritten in a matrix form as follows (e.g., Marfurt 1984)

$$\mathbf{A}(\mathbf{x}, \omega) u(\mathbf{x}, \omega) = f(\omega) \delta(\mathbf{x} - \mathbf{x_s}), \qquad (3.3)$$

where $\mathbf{A}(\mathbf{x}, \omega)$ is the complex impedance matrix. This matrix is typically sparse and symmetric, though it can be non-symmetric for many boundary conditions (Hustedt et al., 2004). Because of the pattern of the matrix, equation 3.3 can be solved with a scientific linear solver package such as the Multifrontal massively parallel solver (MUMPS) (Amestory et al., 2006, Operto et al., 2007) and Portable Extendable Toolkit for Scientific Computation (PETSc) (Balay et al., 2001). We adopt the direct solver approach with MUMPS to speed up the computation.

The modeled plane-wave gather is obtained with a linear source, which is obtained by using all the sources simultaneously each with a phase shift along the surface seismic line (Whitmore, 1995). Correspondingly, the forward problem can be formulated as

$$\mathbf{A}(\mathbf{x},\omega)u(\mathbf{x},\mathbf{p},\omega) = \sum_{\mathbf{x_s}} \varphi(\mathbf{x},\mathbf{x_s},\mathbf{p},\omega)f(\omega)\delta(\mathbf{x}-\mathbf{x_s}), \qquad (3.4)$$

with the ith component of $\varphi(\mathbf{x},\mathbf{x_s},\mathbf{p},\omega)$ given by

$$\varphi(x_i,x_{s_i},p_i,\omega) = \begin{cases} \exp\left(i\omega p_i(x_{s_i} - x_{0_i})\right) & p_i \geq 0 \\ \exp\left(i\omega p_i(x_{s_i} - x_{\max_i})\right) & p_i < 0 \end{cases}, \qquad (3.5)$$

where x_{0_i} and $x_{\max_i}$ are the ith component of the start and end receiver locations, respectively.

Mathematically, equation 3.4 is the same as the time-domain approach. However, a frequency-domain implementation does not need temporal interpolation and has the flexibility to include frequency scaling and phase rotation, which is useful when the source wavelet is not well estimated at first and when frequency-dependent processing is needed to obtain optimal results (Whitmore, 1995).

3.3.2 Gauss-Newton inversion strategy

Like other inverse problems, FWI is formulated as an optimization problem searching for the minimum of a suitably defined misfit functional that measures the difference between the observed and calculated data. A regularization term is necessary to make the inversion better posed (Menke, 1984). In the least-square sense, where L-2 norm is used for the misfit functional, this problem is stated as

$$\min_{\mathbf{m}} E(\mathbf{m}) \quad \text{such that} \quad F(\mathbf{m}) = 0, \qquad (3.6)$$

with

43

$$E(\mathbf{m}) = \frac{1}{2} \sum_{\omega} \sum_{\mathbf{x_s}} \sum_{\mathbf{x_r}} \left[\Delta \mathbf{d}^T \mathbf{W}_d \Delta \mathbf{d}^* + \lambda \left(\mathbf{m} - \mathbf{m}_{prior} \right)^T \mathbf{W}_m \left(\mathbf{m} - \mathbf{m}_{prior} \right)^* \right], \quad (3.7)$$

and

$$\Delta \mathbf{d} = \mathbf{d}_{obs} - \mathbf{d}_{cal}, \qquad (3.8)$$

where $\mathbf{m}$ is the model, $\Delta \mathbf{d}$ is data residual between the observe data $\mathbf{d}_{obs}$ and calculated data $\mathbf{d}_{cal}$, $\Delta \mathbf{d}^T$ and $\Delta \mathbf{d}^*$ are the transpose and complex conjugate of the data residual, respectively; $\mathbf{W}_d$ and $\mathbf{W}_m$ are the inverses of the data and model covariance matrices respectively; $\mathbf{m}_{prior}$ is the prior information about the model known from travel time tomography, for example; λ is a damping factor to weigh the data space and model space errors.

The above nonlinear problem is linearized by expanding the misfit function in a generalized Taylor's series about a reference model $\mathbf{m}_0$. Assuming that the Hessian (second partial derivative of the data with respect to model parameters) can be constructed easily, the model perturbation can be obtained by the following Gauss-Newton framework formula (Pratt et al., 1998)

$$\begin{aligned} \Delta \mathbf{m} &= -\left(\left(\frac{\partial^2 E}{\partial \mathbf{m}^2} \right) + \lambda \mathbf{I} \right)^{-1} \mathbf{W}_m^{-1} \frac{\partial E}{\partial \mathbf{m}} \\ &= -\frac{\mathbf{W}_m^{-1} \operatorname{Re}\left[\mathbf{J}^T \mathbf{W}_d \Delta \mathbf{d}^* \right]}{\operatorname{Re} \mathbf{H} + \lambda \mathbf{I}} \end{aligned}, \qquad (3.9)$$

where $\mathbf{J}$ is the Fréchet derivative of the data with respect to model parameters, $\mathbf{H}$ is the Hessian matrix and $\mathbf{H} = \mathbf{J}^T \mathbf{W}_d \mathbf{J}^*$, $\mathbf{I}$ is the identity matrix (Re means the real part), and $\mathbf{W}_m^{-1}$ acts as a smoothing operator (Virieux and Operto, 2009).

Under Born approximation, the gradient is obtained by taking the partial derivative of the perturbed misfit functional with respect to the perturbed model. It can be explicitly expressed as (Sirgue and Pratt, 2004, Tromp et al., 2005)

$$g(\mathbf{x}) = \frac{\partial E}{\partial \mathbf{m}} = \sum_{\omega} \sum_{\mathbf{x_s}} \sum_{\mathbf{x_r}} \mathrm{Re} \left\{ \frac{-2\omega^2 f(\omega)}{v^3(\mathbf{x})} G(\mathbf{x}, \mathbf{x_s}, \omega) G(\mathbf{x}, \mathbf{x_r}, \omega) W_d(\mathbf{x}, \mathbf{x_r}) \Delta d^*(\mathbf{x_r}, \mathbf{x_s}, \omega) \right\},$$

(3.10)

where $G(\mathbf{x}, \mathbf{x_s}, \omega)$ is the Green's function at the space location $\mathbf{x}$ with a source term at $\mathbf{x_s}$, $W_d(\mathbf{x}, \mathbf{x_r})$ and $\Delta d^*(\mathbf{x_r}, \mathbf{x_s}, \omega)$ are the elements of the matrix $\mathbf{W}_d$ and $\Delta \mathbf{d}^*$, $\mathbf{W}_d \Delta \mathbf{d}^*$ is the weighted data residual and '*' is complex conjugation.

Efficient calculation of the gradient is based on the well-known adjoint state approach, where the forward-propagated source wavefield is crosscorrelated with the back-propagated data residual. In practice, the source wavefield is usually simulated for every source location and the back-propagated residual wavefield is usually simulated simultaneously for all the receiver locations. Under such a framework, the gradient can be written as (e.g., Plessix and Mulder, 2004)

$$g(\mathbf{x}) = \frac{\partial E}{\partial \mathbf{m}} = \sum_{\omega} \sum_{\mathbf{x_s}} \mathrm{Re} \left\{ \frac{-2\omega^2 f(\omega)}{v^3(\mathbf{x})} S(\mathbf{x}, \mathbf{x_s}, \omega) \Delta R^*(\mathbf{x}, \mathbf{x_s}, \omega) \right\},$$

(3.11)

where $S(\mathbf{x}, \mathbf{x_s}, \omega)$ is the forward simulated source wavefield at a particular shot location and $\Delta R(\mathbf{x}, \mathbf{x_s}, \omega)$ is the reverse-time propagated data residual for all the receiver locations. The gradient can be expressed as the crosscorrelation between the forward simulated wavefield and back propagated data residual at the zero lag time.

The full Hessian matrix is never formed explicitly for practical implementations because of the prohibitive computation cost. Instead, the second-order term, which accounts for multiple scattering effects, is often neglected. The first-order term (approximate Hessian), which is due to single scattering, can be explicitly written as (Plessix and Mulder 2004; Tang, 2009)

45

$$H_\alpha(\mathbf{x}_i, \mathbf{x}_j) = \sum_\omega \sum_{\mathbf{x}_s} \sum_{\mathbf{x}_r} \mathrm{Re}\left\{ \frac{4\omega^4 |f(\omega)|^2}{v^3(\mathbf{x}_i)v^3(\mathbf{x}_j)} G^*(\mathbf{x}_i, \mathbf{x}_s, \omega) G(\mathbf{x}_j, \mathbf{x}_s, \omega) G^*(\mathbf{x}_i, \mathbf{x}_r, \omega) G(\mathbf{x}_j, \mathbf{x}_r, \omega) \right\}$$

$$(3.12)$$

where $H_\alpha(\mathbf{x}_i, \mathbf{x}_j)$ is the approximate Hessian. The rows of the Hessian can be interpreted as source and receiver illumination energy. Preconditioning the gradient with the approximate Hessian can remove the geometrical amplitude decay of the Green's functions. Therefore, it can scale the deep perturbations constrained by far-offset data and shallow perturbations constrained by near-offset data (Pratt et al., 1998). The diagonal part of the Hessian can be interpreted as the resolution of individual model parameters. The off-diagonal part of the Hessian can be viewed as the trade-off between different parameters. Large off-diagonal elements indicate large uncertainties for parameter estimation since the misfit functional can remain unchanged when changes of some parameters are compensated with changes of other parameters.

Computing this Hessian is still expensive for present real data applications since it requires $Ns \times Nr$ forward simulations to obtain the source-side and receiver-side Green's functions, with Ns and Nr representing the number of shots and number of receivers, respectively. To further reduce the computational burden, the diagonal part of the Hessian ($\mathbf{x}_i = \mathbf{x}_j$) can be used. It is given by

$$H_0(\mathbf{x}) = \sum_\omega \sum_{\mathbf{x}_s} \sum_{\mathbf{x}_r} \mathrm{Re}\left\{ \frac{4\omega^4 |f(\omega)|^2}{v^6(\mathbf{x})} |G(\mathbf{x}, \mathbf{x}_s, \omega)|^2 |G(\mathbf{x}, \mathbf{x}_r, \omega)|^2 \right\}, \qquad (3.13)$$

where $H_0(\mathbf{x})$ is used to represent the diagonal Hessian, the L-2 norm of the complex scalar vector $|f(\omega)|^2 = f^*(\omega)f(\omega)$, the same relation holds for $G(\mathbf{x}, \mathbf{x}_s, \omega)$ and $G(\mathbf{x}, \mathbf{x}_r, \omega)$. The diagonal Hessian is the autocorrelation of the partial derivative of the source-side and receiver-side Green's functions. This can still provide a good approximation to scale the gradient when the off-diagonal elements are relatively small.

46

The computation cost of the diagonal Hessian requires $Ns + Nr$ forward simulation, which is equivalent to the calculation of the gradient.

3.3.3 Plane-wave encoding strategy

The computational cost of FWI can be reduced significantly by incorporating the plane-wave forward modeling strategy because of reduction in the number of wavefield simulations. The source wavefield, back-propagated residual wavefield and the source and receiver Green's functions are calculated using equations 3.4 and 3.5. After numerical simulations, plane-wave encoded gradient can be written as

$$\tilde{g}(\mathbf{x}) = \sum_{\omega}\sum_{\mathbf{p_s}} \text{Re}\left\{\frac{-2\omega^2|\omega|}{v^3(\mathbf{x})}f(\omega)\tilde{S}(\mathbf{x},\mathbf{p_s},\omega)\Delta\tilde{R}^*(\mathbf{x},\mathbf{p_s},\omega)\right\}. \tag{3.14}$$

where $\mathbf{p_s}$ is the source ray parameter. It corresponds to the take-off angle of the receiver gather. Equation 3.14 is equivalent to shot-profile gradient (Appendix B). However, the summation in plane-wave gradient is over ray parameters instead of shot locations, which can considerably reduce the number of simulations when a large number of sources are simulated in a conventional FWI.

Similar to the gradient calculation, the diagonal Hessian can also exploit a plane-wave phase encoding strategy to reduce computational burden. Using an algorithm similar to wave-equation shot-profile migration, we can obtain the receiver-side phase-encoded Hessian as follows

$$\tilde{H}_0(\mathbf{x}) = \sum_{\omega}\sum_{\mathbf{x_s}}\sum_{\mathbf{p_r}} \text{Re}\left\{\frac{4\omega^4|\omega||f(\omega)|^2}{v^6(\mathbf{x})}|G(\mathbf{x},\mathbf{x_s},\omega)|^2|\tilde{G}(\mathbf{x},\mathbf{p_r},\omega)|^2\right\}. \tag{3.15}$$

We further consider a simultaneous encoding of source and receiver-side Green's functions with plane-wave simulations. This computing process can be easily implemented. It requires only two forward wavefield simulations: one for source plane-

wave wavefiled and the other for receiver plane-wave wavefield. Mathematically, it can be written as

$$\tilde{\tilde{H}}_0(\mathbf{x}) = \sum_\omega \sum_{\mathbf{p_s}} \sum_{\mathbf{p_r}} \mathrm{Re} \left\{ \frac{\omega^4 |\omega f(\omega)|^2}{v^6(\mathbf{x})} |\tilde{G}(\mathbf{x}, \mathbf{p_s}, \omega)|^2 |\tilde{G}(\mathbf{x}, \mathbf{p_r}, \omega)|^2 \right\}. \tag{3.16}$$

Details of the derivation of equations 3.15 and 3.16 are given in Appendix C. Equations 3.13, 3.15 and 3.16 have similar forms, but they differ from each other in the propagation scheme and in the number of forward simulations required. Assuming that a survey has *Ns* shots and *Nr* receivers, Table 3.1 compares the number of simulation steps needed for approximate Hessian and diagonal part of the approximate Hessian, respectively. *Nw* is the number of frequencies required for the simulation. Unlike random phase encoding approach, where the number of simulations varies for marine geometry and for land-acquisition or ocean bottom seismometer (OBS) geometry, plane-wave encoding approach provides the same number irrespective of the number of shots and receivers used. The number of simulations, however, depends on how many ray parameters are selected. The number of plane-waves can be determined by the imaging targets and by the maximum dips present in the shot and receiver gathers.

Any kind of inversion approach in a Gauss-Newton framework can be applied for optimization. For simplicity, we use a steepest decent approach to update the velocity model. This method exploits a parabolic or cubic search algorithm with a modification of the fixed length model perturbation of equation 3.9. The model update can be written as

$$\mathbf{m}_k(\mathbf{x}) = \mathbf{m}_{k-1}(\mathbf{x}) + \alpha_k \left(H_{0_{k-1}}(\mathbf{x}) + \lambda \mathbf{I} \right)^{-1} \mathbf{W}_m^{-1} g_{k-1}(\mathbf{x}). \tag{3.17}$$

where α_k is a scalar for the kth iteration. It is defined by line searching along the preconditioned gradient direction.

48

Table 3.1: Cost comparison of the number of wavefield simulations for different scenarios.

	Direct computation	Plane wave, receiver side	Plane wave, simultaneous
Approximate Hessian	$Nw \times Ns \times Nr$	$Nw \times Ns \times Np_r$	$Nw \times Np_s \times Np_r$
Diagonal Hessian	$Nw(Ns + Nr)$	$Nw(Ns + Np_r)$	$Nw(Np_s + Np_r)$

A multi-scale solution is obtained through the inversion of discrete frequencies. Groups of frequencies are selected and two successive loops are formed. The first loop is over a group of frequencies, and the second loop is over the frequencies of the frequency group (Brossier et al., 2009). Inverted model of the previous frequency loop is used as the starting model for the new loop. The stopping criterion for each iteration is given by

$$\varepsilon = \frac{\left\| \mathbf{d}_{obs}(\mathbf{x}, \omega) - \mathbf{d}_{cal}(\mathbf{x}, \omega) \right\|}{\left\| \mathbf{d}_{cal}(\mathbf{x}, \omega) \right\|} \quad . \tag{3.18}$$

If ε is big, computational time can be saved. On the other hand, if ε is small, a more accurate data fitting can be obtained. In this dissertation, we use $\varepsilon = 0.001$ and this value is fixed for all iterations and for each frequency group.

3.4 NUMERICAL EXAMPLES

In this section, I will demonstrate the numerical applications of our approach with synthetic examples. The synthetic examples are based on a 2D profile of the SEG/EAGE Overthrust model (Figure 3.1). Analysis of effects of different ray parameters on the gradient is presented to demonstrate the dip-selective advantage of our method. In other words, the gradient calculation can be benefited from plane-wave decomposition by focusing the image on selected targets. Calculation of the Hessian with plane-wave encoding is more powerful since only approximated Hessian with a few ray parameters are necessary to ensure the illumination compensation of the gradient. I analyze the

effectiveness of the encoding schemes with selective ray parameters. I finally show the inversion results from our examples with different strategies on the selection of the ray parameters.

3.4.1 Dip-selective gradient
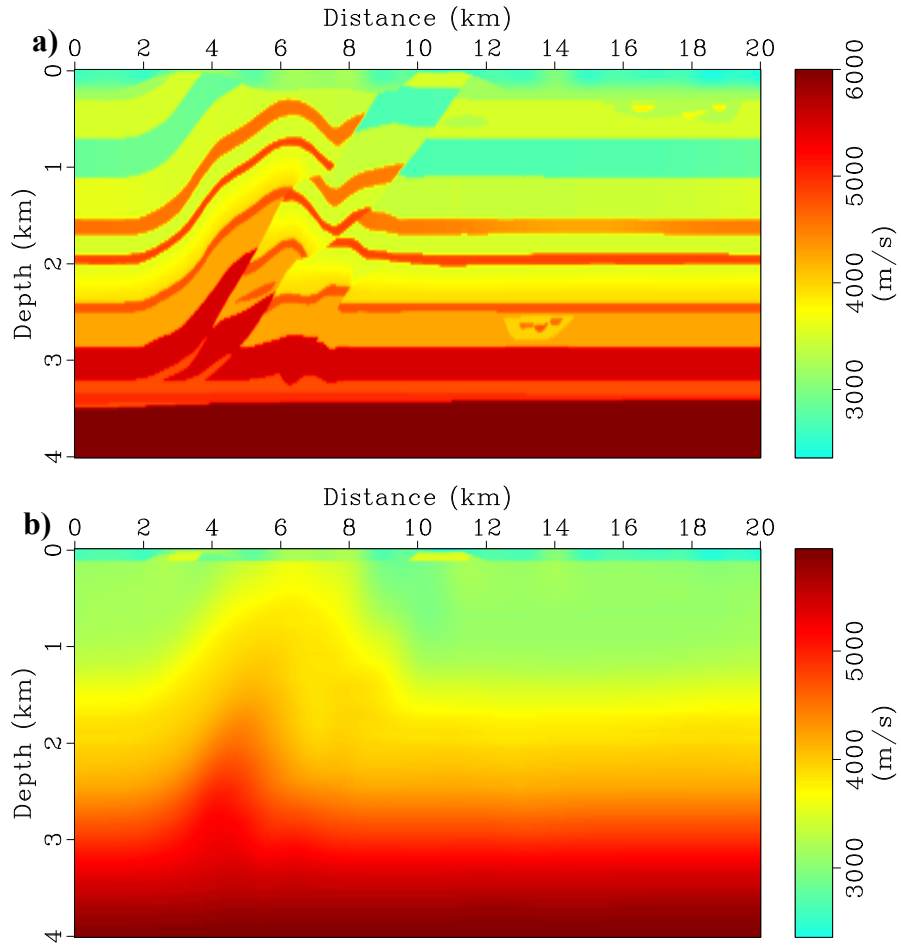
Figure 3.1: A 2-D profile of the Overthrust model. (a) P-wave velocity of the true model. (b) Initial model used to generate the gradient and for inversion in this dissertation.

The P-wave velocities of the true model and initial model (a smoothed version of the true model) are shown in Figure 3.1. This model has 187×801 grids with 25m of

horizontal and vertical grid intervals. While a 2-D example is presented for simplicity, extension of all the following analysis to a bigger model and to 3-D is trivial. Similar to Ben-Hadj-Ali et al. (2011) we use seven different frequencies ranging from 3 Hz to 20 Hz. The range of ray parameters is defined by the maximum dip of the recorded data. The range for the ray parameters used in this dissertation is $[-0.4 \text{ s/km}, 0.4 \text{ s/km}]$.
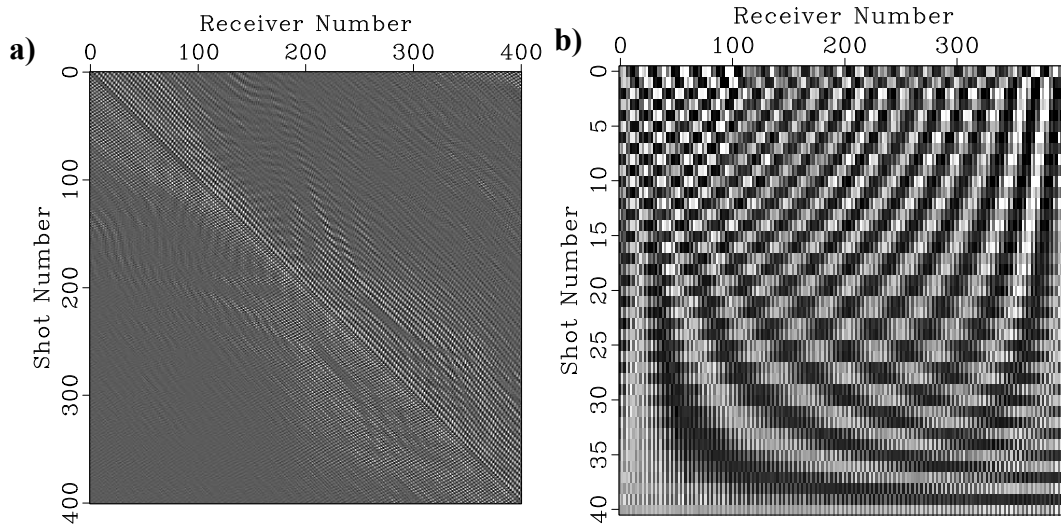
Figure 3.2: Monochromatic frequency (f=20 Hz) seismic data (a) in the source-receiver domain. (b) ray parameter-receiver domain.

A single frequency data for shot-domain modeling and plane-wave domain is shown in Figure 3.2. The sparsity in the shot-domain does not exist in the plane-wave domain. The effect of a single ray parameter to the gradient is shown in Figure 3.3. The gradient shown here is the first iteration with the initial model in Figure 3.1b. Though a sequential calculation for each frequency or each frequency band is typically required for the inversion process, the gradient is summed over all the frequencies for a better display. Each ray parameter is related to the take-off angle at the surface location, it contributes to the gradient with directional illumination. More energy is focused on the left side of the

51

model for a negative ray parameter (Figure 3.3b) and more illumination on the right side of the model for a positive ray parameter (Figure 3.3d). A zero ray parameter (Figure 3c) tends to provide balanced illumination on both sides of the model. For reference, the gradient with 81 ray parameters is shown in Figure 3.3a.
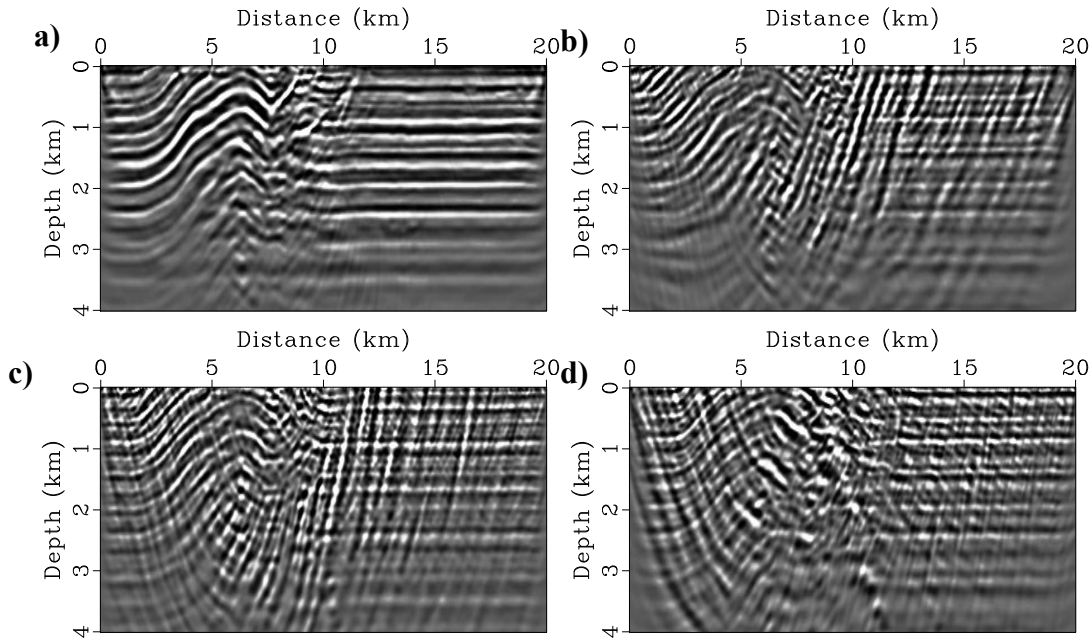
Figure 3.3: Plane-wave gradient at the first iteration with different choice of parameters. (a) 81 source and receiver ray parameters; (b) single source ray parameter $p_s = -0.3$ s/km; (c) $p_s = 0$ s/km; (d) $p_s = 0.3$ s/km.

Figure 3.4 shows how the number of ray parameters affects the gradient calculation. For only zero ray parameter (zero offset section), there are plenty of artifacts that are generated by the footprints of plane-wave encoding (Figure 3.4a). For 9 ray parameters (Figure 3.4b), the artifacts reduce substantially. For 41 ray parameters (Figure 3.4c), the artifacts become unnoticeable compared with 81 ray parameters (Figure 3.4d). This provides a qualitative way to define the number of ray-parameters to be used in the inversion. For this problem, we assess that 41 ray parameters are adequate for inversion,

52

provided that the 41 ray parameters are enough for Hessian calculation. Another quantitative requirement of the number of ray parameters can be written as (Zhang et al., 2005)

$$N_p \geq \frac{L f_{max}(\sin \alpha_{max} - \sin \alpha_{min})}{v}, \tag{3.19}$$

where N_p is the number of ray parameters. L is the distance of the receiver gather, f_{max} is the maximum frequency for inversion, α_{min} and α_{max} are the minimum and maximum take-off angles, respectively, v is the velocity at the surface location.
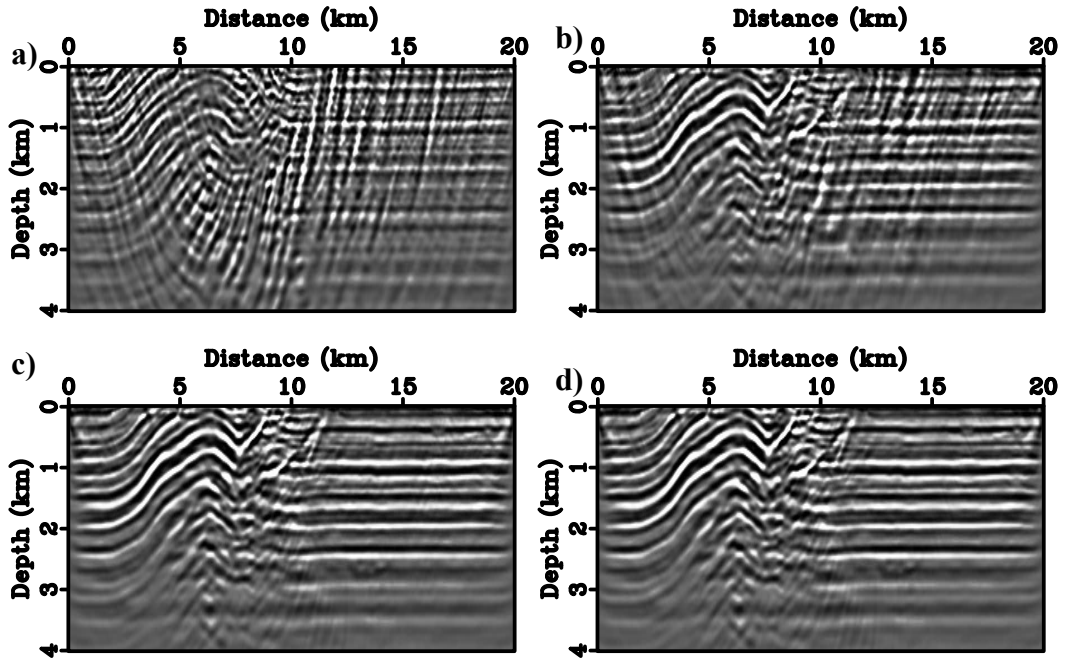
Figure 3.4: The effect of the ranges of ray parameters on the plane-wave gradient. (a) 1 source ray parameter ($p_s = 0$ s/km); (b) 9 source ray parameters equally distributed from -0.4 s/km to 0.4 s/km ; (c) 41 source ray parameters; (d) 81 source ray parameters.

3.4.2 Plane-wave encoded Hessian

To test the effect of directional illumination for the source and receiver plane-wave encoded Hessian, we first apply this strategy to a medium with a constant velocity gradient. The background velocity is described by $v(z) = 2000 + 0.4z$. The diagonal Hessian for a single frequency at 3 Hz is shown in Figure 3.5.

In addition, the Hessian for the true Overthrust model is shown in Figure 3.6. Unlike the gradient calculation where only a single ray parameter is used, the Hessian is calculated with one source ray parameter and all receiver ray parameters. Similar to the gradient, single ray parameter contributes to directional illumination of the model. This decomposition process can be used for directional choice of the ray parameters for the inversion.
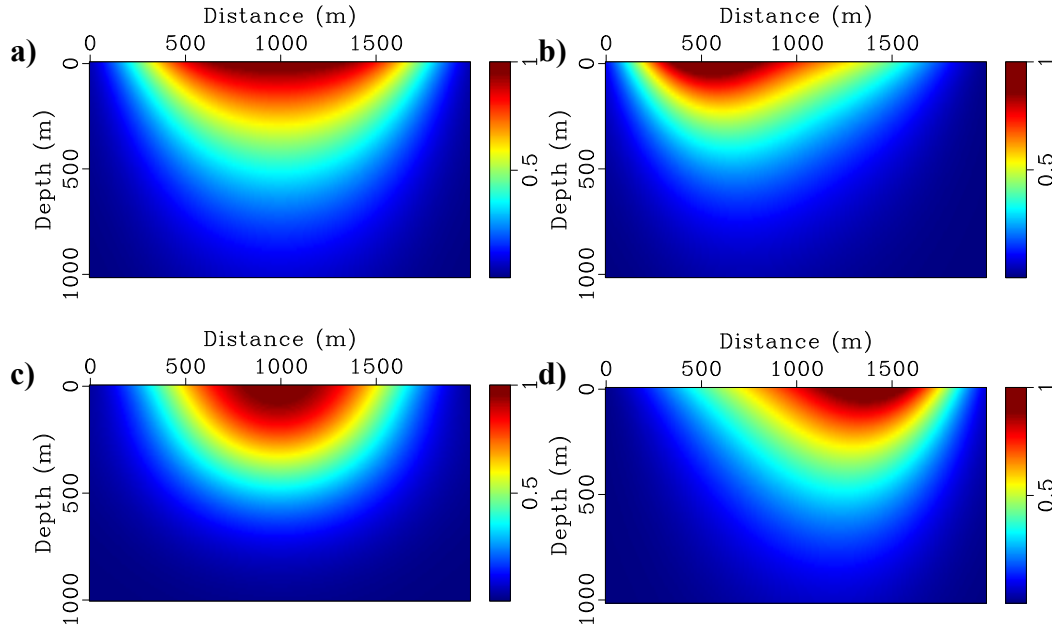
Figure 3.5: Diagonal Hessian for a constant velocity gradient model with different selection of parameters. (a) 101 source and 101 receiver ray parameters; (b) single source ray parameter $p_s = -0.3$ s/km and 101 receiver ray parameters; (c) $p_s = 0$ s/km; (d) $p_s = 0.3$ s/km.

54
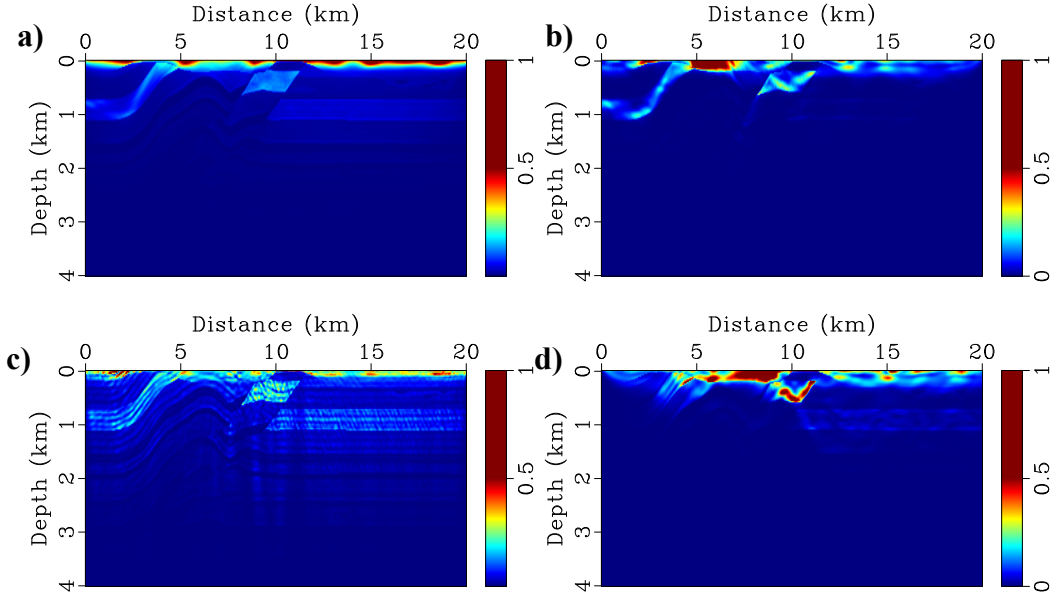
Figure 3.6: Diagonal Hessian for the true model with different selection of parameters. (a) 81 source and receiver ray parameters; (b) single source ray parameter $p_s = -0.3$ s/km and 81 receiver ray parameters; (c) $p_s = 0$ s/km; (d) $p_s = 0.3$ s/km.

Plane-wave encoding provides a good approximation of the Hessian (Figure 3.7). For the plane-wave 9 source and receiver ray parameters (Figure 3.7b), this approximation has notable artifacts. For 41 source and receiver ray parameters, theses artifacts are very small (Figure 3.7c). 81 source and receiver ray parameters (Figure 3.7d) produce visually identical results as the true Hessian calculated with 401 sources and 401 receivers (Figure 3.7a). This also confirms that 41 ray parameters may be enough for this problem. Note that if the off-diagonal parts of the Hessian are needed for the accuracy requirement, plane-wave encoding is more appealing because significant amount of computational time can be saved.
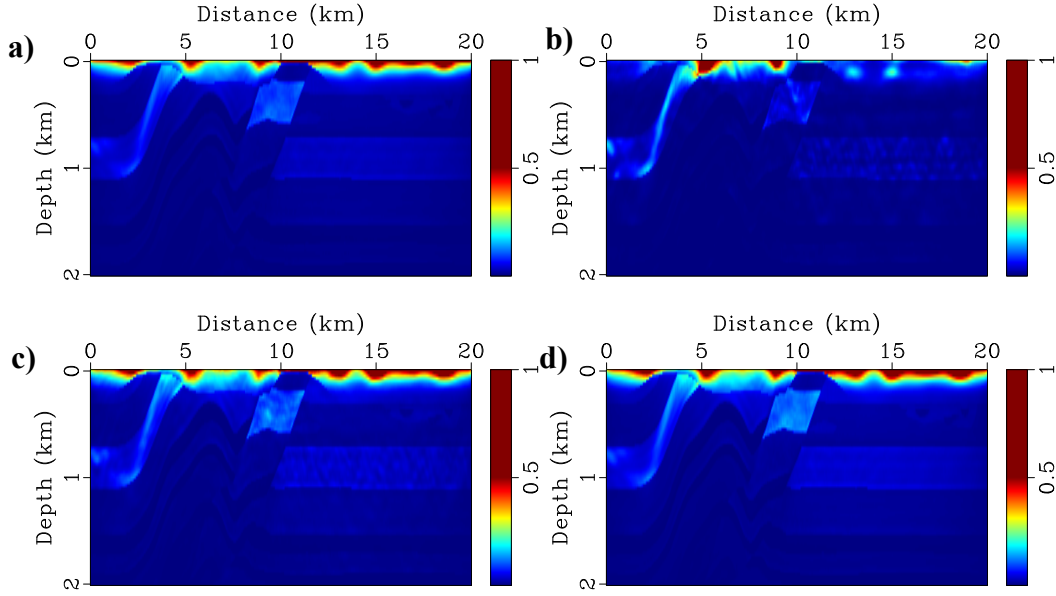
Figure 3.7: Comparison of the diagonal Hessian for the true model with different parameters. (a) true Hessian with 401 sources and 401 receivers; (b) plane-wave Hessian with 9 source and receiver ray parameters; (c) plane-wave Hessian with 41 source and receiver ray parameters; (d) plane-wave Hessian with 81 source and receiver ray parameters.

3.4.3 Inversion of the Overthrust model

We employ a sequential inversion approach for each frequency component. Inverted result from the previous step is used as the starting model for the next step and so on. Figure 3.8b shows the inverted model with 9 source and receiver ray parameters. The gradient only requires decomposing the receiver gather into source ray parameters. The receiver ray parameters are used for Hessian calculation. For each iteration, only 9 wavefield simulations are needed for the gradient and also 9 simulations for the Hessian when the reciprocity is considered. From the inverted result, we can see that with only 9 ray parameters, promising result can be obtained. However, it still has some errors (more on the shallow part) because the gradient is not preconditioned very well. With 41 ray parameters (Figure 3.8c), the inverted result is almost comparable with the result using

56

401 sources and 401 receivers (Figure 3.8a). Ben-Hadj-Ali et al. (2011) used 199 sources and 200 receivers and showed similar results as Figure 3.8a.
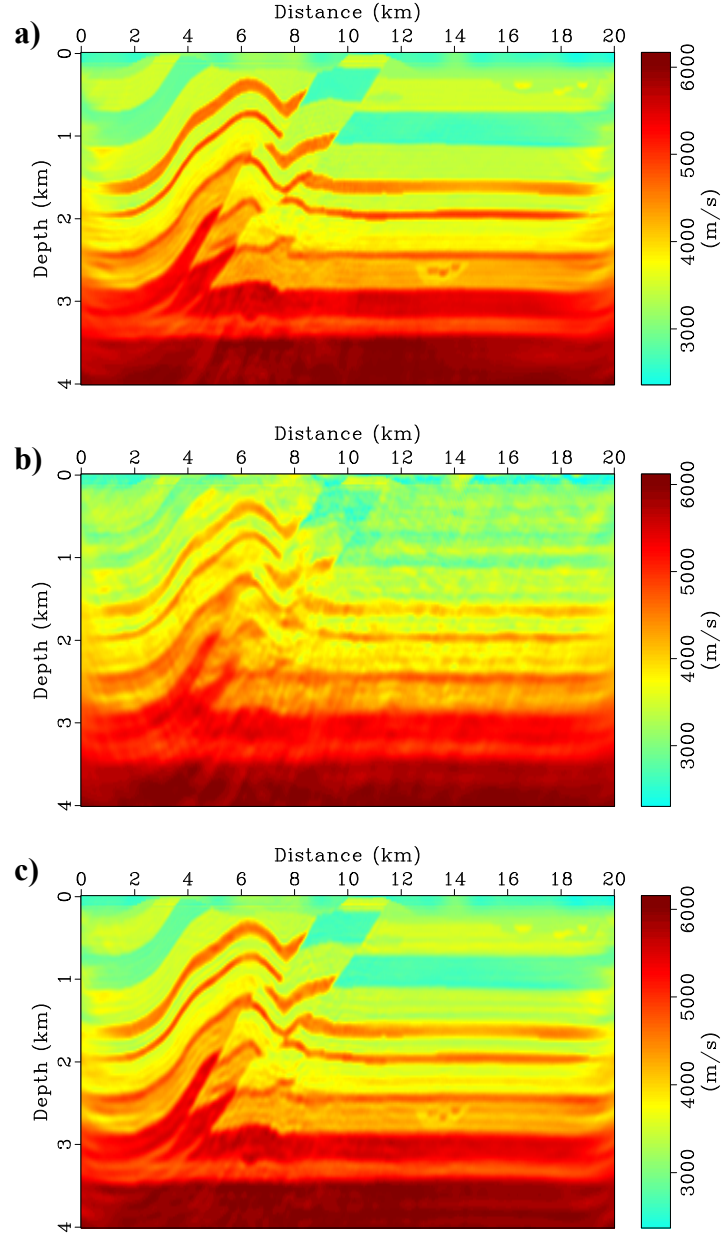
Figure 3.8: Inverted P-wave velocity model with (a) shot domain approach with 401 sources and 401 receivers; (b) plane-wave domain with 9 source and receiver ray parameters; (c) plane-wave Hessian with 41 source and receiver ray parameters.

Figure 3.9 plots the normalized data residuals as a function of the iteration number. For a lower frequency, the error converges faster and using more iterations does not reduce the error much. For a higher frequency, since the starting model is already close to the true model, a slow search process is shown in this plot. Figure 3.10 shows the time domain plane-wave gather for $p_s = 0$ s/km and $p_s = 0.3$ s/km, respectively. This is obtained by an inverse Fourier Transform of the frequency domain data. While little structure can be observed from the starting model, the inverted model shows features that are similar to those of the true model.
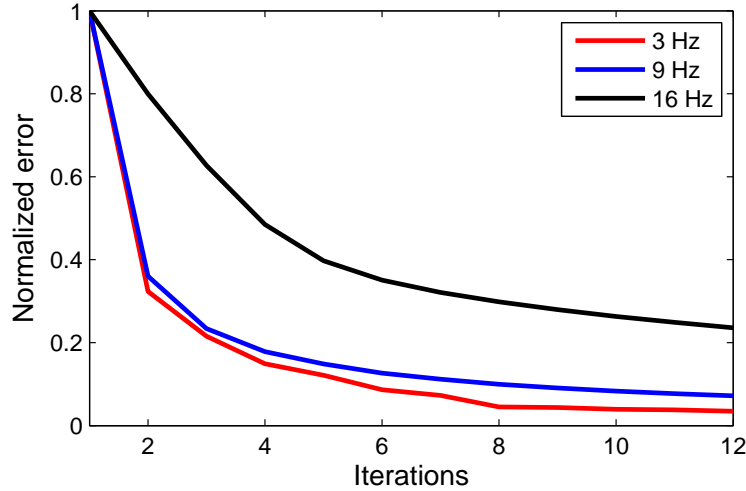
Figure 3.9: Normalized error versus iterations for three different frequencies.

3.5 CONCLUSIONS

In this chapter, I have presented a plane-wave encoding strategy for efficient frequency-domain full waveform inversion (FWI). This approach employs a simultaneous-shot approach for gradient calculation and a source and receiver encoding scheme for Hessian calculation. Compared to random phase encoding approach, it does

not introduce strong cross-talk artifacts. Efficiency of this method is achieved by the use of a limited range of ray parameters for wavefield simulations. This encoding is robust with respect to noise. Directional illumination can be obtained with the selection of ray parameters, which may result in more savings of the computational time. Compared to time domain methods, frequency domain forward modeling can model multiple frequencies efficiently. Thus, this approach provides an efficient and robust tool for full waveform inversion.
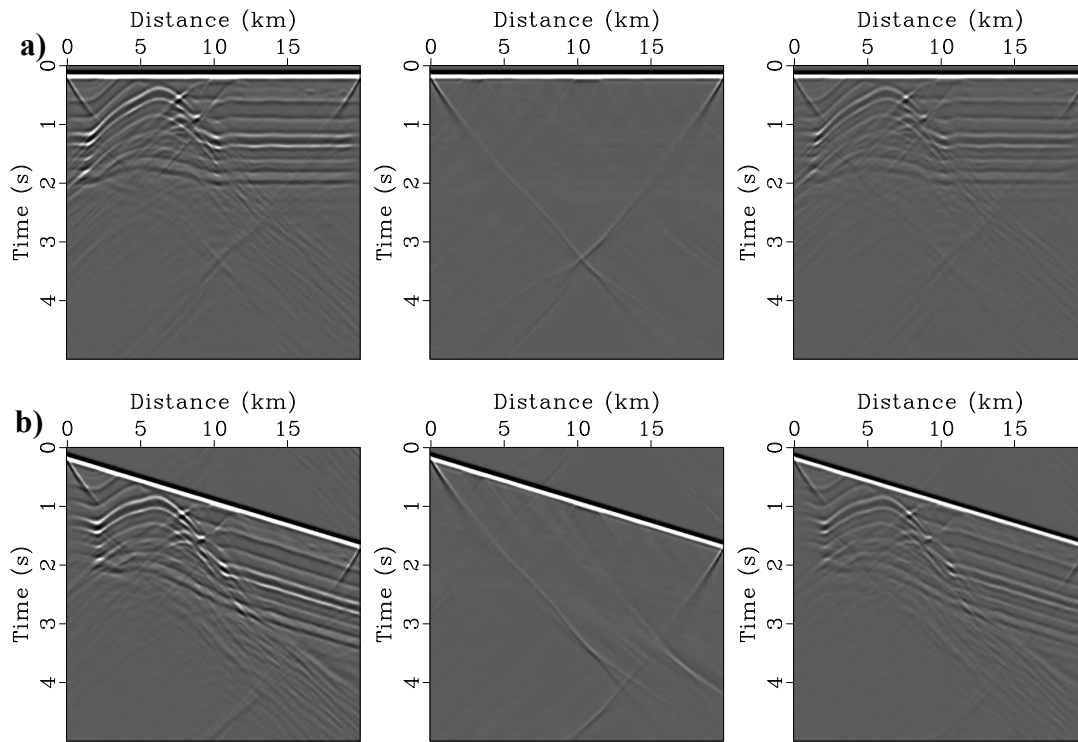
Figure 3.10: Time domain plane-wave gather for (a) $p_s = 0$ s/km and (b) $p_s = 0.3$ s/km. Left: true model. Middle: initial model. Right: Inverted model.

59

Chapter 4: Suppressing non-Gaussian noises with scaled receiver wavefield for reverse time migration

4.1 ABSTRACT

Numerical implementation of the gradient of the cost function in a gradient-based full waveform inversion (FWI) is essentially a migration operator used in wave equation migration. In FWI, minimizing different data residual norms results in different weighting strategies of data residuals at receiver locations prior to back propagation into the medium. In this dissertation, we propose different scaling methods to the receiver wavefield and compare their performances. Using time-domain reverse time migration (RTM), I show that compared to conventional algorithms, this type of scaling is able to significantly suppress non-Gaussian noise, i.e., outliers. Numerical tests also show that scaling by its absolute norm produces better results than other approaches.

4.2 INTRODUCTION

Migration/inversion using the wave equation honors the physical basis of seismic wave propagation. Unlike the ray-based migration/inversion, they retain band-limited frequency effects instead of using infinite-frequency rays and inherently include multi-arrivals. Migration/inversion with two-way wave equation that takes advantage of the full wave information is known as reverse time migration (RTM) (e.g., Baysal et al., 1983; Symes, 2007; Zhang and Zhang, 2009) and full waveform inversion (FWI) (e.g., Tarantola, 1987; Virieux and Operto, 2009). Both can handle challenging geological structures with no limit of dips or overhangs, and thus are able to provide an accurate high-resolution migrated image and velocity model for quantitative imaging of the Earth's subsurface.

Despite the general success of RTM, it is also known for producing artifacts when uninterested arrivals (e.g. head waves, diving waves and backscattered waves) correlate (Guitton et al., 2007). In addition, field data usually contains outliers such as non-Gaussian noise and high-amplitude anomalies. With a conventional back-propagation approach, the noise and the signal have equal chance to enter the medium. Thus, outliers will affect the quality of the image obtained by wave equation imaging methods. To enhance the migrated image, the noise footprint is usually reduced by stacking multi-fold data and by careful preprocessing. However, enhancing the image by novel imaging methods offers better chances to produce high quality results.

Common approaches to suppress noise in the migrated image typically rely on applying different imaging conditions. Fletcher et al. (2005) suggested application of a directional damping factor to remove the imaging noise. Yoon and Marfurt (2006) proposed a Poynting-vector imaging condition. Costa et al. (2009) further utilized the Poynting-vector information for obliquity correction. Zhang and Sun (2009) proposed to obtain true-amplitude imaging and mute artifacts with prestack angle gathers. Yan and Xie (2009) used wavefield separation to decompose the data into plane-wave components in the angle domain. Liu et al. (2011) decomposed the extrapolated wavefields into one-way wavefileds and selected different components to crosscorrelate. However, all of these methods are concentrated on removing unwanted arrivals such as diving waves and backscattered waves, and not on outlier noise.

Some reports reveal that improved noise attenuation in geophysical inversion can be obtained from L1-norm (Tarantola, 1987), and a hybrid L1/L2-norm (Huber 1973, Bube and Langan, 1997, Guitton and Symes, 2003). Recently, these norms have successfully been incorporated in frequency domain FWI (Pyun et al., 2009, Ha et al, 2009 and Brossier et al., 2010). It is known that the process to obtain the gradient of the

misfit function with an adjoint-state approach is equivalent to one-step wave equation migration (Tarantola, 1984; 1988, Shin et al., 2001). The only difference is that FWI back-propagates residual receiver wavefield and wave equation migration back-propagates the receiver wavefield (Mora 1998). Based on this, Lee et al. (2011) proposes a frequency domain RTM with an L-1 norm gradient calculation procedure of FWI to weigh the receiver wavefield.

In this chapter, I extend Lee et al.'s approach to include different weight factors and to apply these factors using a time-domain RTM. This procedure closely resembles conventional imaging procedures, except that the back-propagated receiver wavefield is scaled. This approach does not introduce significant additional cost except for a mapping from time domain data to frequency domain at the correlation or imaging stage only, which is negligible compared to the cost of calculation of source and receiver Green's functions. The synthetic examples show that this method can indeed suppress outlier noise much better than the conventional approaches.

4.3 METHODOLOGY

4.3.1 Conventional time-domain reverse time migration

RTM is typically performed on individual common-shot records and then a stacking operator is applied on each individual migrated data to produce the final image. For simplicity, we consider the following acoustic wave equation for source wavefield and receiver wavefield extrapolation

$$\nabla \cdot \left(\frac{1}{\rho(\mathbf{x})} \nabla u(\mathbf{x},t) \right) - \kappa(\mathbf{x}) \frac{\partial^2 u(\mathbf{x},t)}{\partial t^2} = f(t)\delta(\mathbf{x}-\mathbf{x_s}), \qquad (4.1)$$

62

where $u(\mathbf{x},t)$ is the pressure wavefield in the time domain, $\rho(\mathbf{x})$ is the mass density, $\kappa(\mathbf{x})$ is the compressibility, i.e., the inverse of bulk modulus, and $f(t)$ is the time domain source signature.

After constructing the source wavefield forward in time and the recorded receiver wavefield backward in time with equation 4.1, an imaging condition is applied to combine these two different wavefields (e.g., McMechan, 1983, Whitmore, 1983, Baysal et al., 1983). Like other migration methods, various imaging conditions can be applied to retrieve the reflectivity information of the subsurface. For simplicity, we utilize a conventional crosscorrelation imaging condition (Claerbout, 1971)

$$I(\mathbf{x}) = \sum_{\mathbf{x_s}} \sum_{t} R(\mathbf{x}, \mathbf{x_s}, t) S(\mathbf{x}, \mathbf{x_s}, t), \qquad (4.2)$$

where $I(\mathbf{x})$ is the migrated image at a space location $\mathbf{x}$ in the subsurface, $S(\mathbf{x}, \mathbf{x_s}, t)$ and $R(\mathbf{x}, \mathbf{x_s}, t)$ are the forward propagated source wavefield and backward propagated receiver wavefield, respectively.

In equation 4.2, both source and receiver wavefields at the subsurface location $\mathbf{x}$ depend on the source location $\mathbf{x_s}$. Equation 4.2 also indicates that the recorded wavefield at each receiver location in the survey should be back-propagated simultaneously to obtain $R(\mathbf{x}, \mathbf{x_s}, t)$. Using this propagation scheme, the contribution of the amplitude in the image space for each trace is based on the recorded amplitude in the data space. Because no division is required, this approach in general is robust. Random noises can be suppressed with a stacking operator of the migrated gathers. However, it is not effective in suppressing the non-Gaussian type noises, i.e., outliers.

4.3.2 Inversion for data with outliers

As mentioned, FWI is closely related to migration by the fact that each iteration can be viewed as a migration process with a model updating strategy. Similar to the work

reported by Lee et al. (2011) for frequency domain RTM with an L-1 norm type scaling, we begin our scaling approaches by revisiting the FWI approaches for the data with outliers.

A standard FWI procedure is a minimization of the least squares norm cost function (Pratt 1998, Virieux and Operato, 2009). This norm is the square product of difference between the observed seismic data and calculated data. To introduce the FWI methods for outliers, we use the following general L_p norm of the misfit function

$$E_{L_p} = \|\Delta \mathbf{d}\|_p = \|\mathbf{d}_{obs} - \mathbf{d}_{cal}\|_p, \tag{4.3}$$

where E is the misfit function, $\mathbf{d}_{obs}$ and $\mathbf{d}_{cal}$ are the recorded and synthetic data, respectively. To effectively invert the data with outliers, p is typically smaller than two. Under Born approximation, the gradient of equation 4.3 can be obtained by differentiating the misfit function with respect to model parameters. For L_2 norm ($p=2$), the gradient is

$$g_{L_2} = J^T \Delta \mathbf{d}, \tag{4.4}$$

where J is the Fréchet derivative matrix, T denotes a transpose operator, $\Delta \mathbf{d} = \mathbf{d}_{obs} - \mathbf{d}_{cal}$. For efficient calculation of the gradient, J is never formed explicitly. Instead, an adjoint-state approach (Tromp et al., 2005, Plessix, 2006) is routinely used. J^T is the migration operator which naturally leads to a crosscorrelation between the source wavefield and back propagated receiver wavefield at zero-lag time.

The gradient of L-1 norm misfit function can be calculated as follows (Tarantola, 1987, Brossier et al., 2010)

$$g_{L_1} = J^T F^{-1} \left(\frac{\Delta \mathbf{d}(\omega)}{|\Delta \mathbf{d}(\omega)|} \right), \tag{4.5}$$

where $\mathbf{d}(\omega)$ is frequency domain data, and F^{-1} means inverse Fourier transform. Note that treating the real and imaginary parts of $\mathbf{d}(\omega)$ in equation 4.3 separately results in the following formula (Pyun et al., 2009, Lee et al. (2011))

$$g_{L_1} = J^T F^{-1} \left(\text{sgn} \left(\text{Re} \left[\Delta \mathbf{d}(\omega) \right] \right) + i \, \text{sgn} \left(\text{Im} \left[\Delta \mathbf{d}(\omega) \right] \right) \right), \qquad (4.6)$$

where Re means the real part and Im means the imaginary part, sgn is the signum function.

A hybrid L1/L2 norm can be obtained using a Huber's criterion (Huber 1973). Another approach for a hybrid L1/L2 norm is given by Bube and Langan (1997). The gradient for a Huber norm can be written as follows

$$g_{L_{Huber}} = \begin{cases} J^T \delta \mathbf{d} & \left| \Delta \mathbf{d}(\omega) \right| \leq \varepsilon \\ J^T F^{-1} \left(\Delta \mathbf{d}(\omega) / \left| \Delta \mathbf{d}(\omega) \right| \right) & \left| \Delta \mathbf{d}(\omega) \right| > \varepsilon \end{cases} \qquad (4.7)$$

where ε is a threshold parameter.

4.3.3 Migration with scaled receiver wavefield

Incorporating the above ideas from FWI, a new approach for reverse time imaging is proposed as follows

$$I(\mathbf{x}) = J^T F^{-1} \left(\frac{\mathbf{SR}(\omega)}{\left| \mathbf{SR}(\omega) \right|^\alpha + \xi} \right), \qquad (4.8)$$

where α is a parameter to scale the amplitude of the receiver wavefield $\mathbf{R}(\omega)$, ξ is a regularization parameter to stabilize division, $\mathbf{S}$ is a smoothing operator necessary for noisy data to removing the random oscillations and I is the migrated image. This operator acts on the frequency domain data and we use a moving average smoothing for simplicity. For real data where noise attenuation is typically applied before migration, this operator is not necessary. When $\alpha = 0$, it corresponds to the conventional imaging procedure; when $\alpha = 1$, it is the L-1 norm case as in FWI; when combining the two above

cases together, we can also get a hybrid solution as Huber norm does. Also, when $\alpha = 2$, it corresponds to the case of a student's t-distribution of the *a posteriori* distribution from a statistical point of view of the FWI minimizing functional (Aravkin et al., 2011).

Logarithmic wavefield scaling used in frequency domain FWI (Shin and Min, 2006), can also be used here for RTM with the migration approach described as follows

$$I(\mathbf{x}) = J^T F^{-1}\left(\frac{\ln\left(\mathbf{SR}(\omega)\right)}{\left|\ln\left(\mathbf{SR}(\omega)\right)\right|^\alpha + \xi} \right).$$

$$(4.9)$$

From equations 4.8 and 4.9 we observe that the method proposed here differs from conventional methods with a scaling factor to the amplitude. In a conventional method, the outliers, typically presented by amplitude anomalies, have equal chance as other traces to enter into the image domain. However, after amplitude scaling, the anomalies are either normalized or somehow scaled; therefore our method is more outlier resistant.

Although the conventional crosscorrelation imaging condition is applied to form an image from the source wavefield and receiver wavefield, this method applies to any other imaging conditions such as deconvolution imaging condition and time-shifting imaging conditions (Sava and Fomel, 2006). This approach also applies to angle gathers (Xu et al., 2011), plane-wave data (Vigh and Starr, 2008), and one-way wave equation migration methods (Muder et al., 2004).

4.4 EXAMPLES

4.4.1 Non-Gaussian random noise

Our first application is that of a simple layered (1-D) model (Figure 4.1). A single source and single receiver numerical experiment is performed to analyze the sensitivity of the anomaly of the recorded data. This will generate a migration kernel, which represents

the contribution of individual source and receiver to the migrated image. The source is located at (750m, 30m) and the receiver is located at (2250m, 30m). We follow the approach of Roth and Korn (1993) to generate non-Gaussian random noise. This type of random media has zero mean and a particular variance related to its autocorrelation function. We use an exponential autocorrelation function defined as follows

$$\phi(k_x, k_z) = \varepsilon^2 \frac{4ab}{\left(1 + a^2 k_x^2\right)\left(1 + b^2 k_z^2\right)},$$

(4.10)

where ϕ is the spatial autocorrelation function, ε is the variance, k_x and k_z are the horizontal and vertical wavenumber, respectively, a and b are the horizontal and vertical correlation length, respectively.
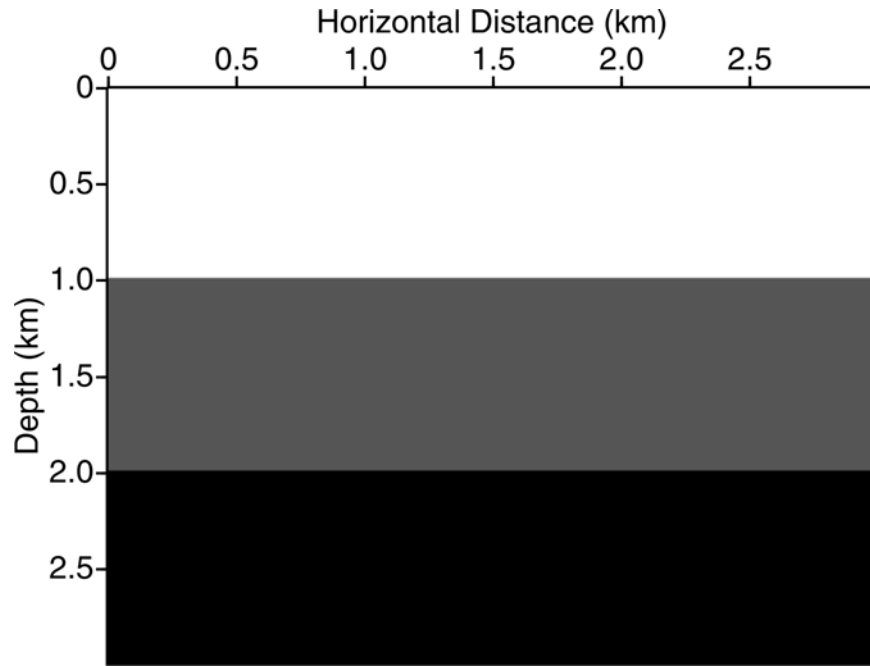
Figure 4.1: A layered medium used to generate synthetic data. The model has 300×300 grids with a uniform horizontal and vertical spacing. Velocity for each layer is 2500m/s, 2600m/s and 2650m/s, respectively.

Figure 4.2 shows the wiggle plot of the synthetic trace of two reflectors with direct wave removed. The amplitude of the random noise is comparable to that of the signal of the trace. This is used to simulate high anomalies of the real seismic data.
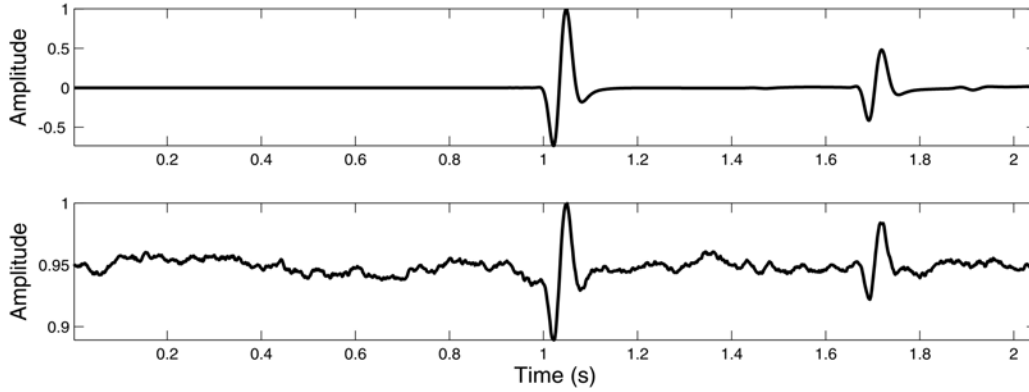
Figure 4.2: Wiggle plot of the normalized receiver data for two reflectors. Top: Noise-free data. Bottom: Noisy data.

Figure 4.3 shows the migration kernel of the reflections without any filtering. For a clean data (Figure 4.3a), the migrated image has two circular events. With a conventional approach (Figure 4.3b), where the back-propagated wavefield is the recorded receiver data, it generates low frequency artifacts plus high-amplitude spurious events. When the receiver wavefield is scaled for back-propagation by its absolute norm (L-1 norm) (Figure 4.3c), the artifacts shown in Figure 4.3b are eliminated. With a hybrid approach (Figure 4.3d) similar to hybrid norm used in FWI, we can obtain cleaner results than that obtained by the conventional method, but still there are some artifacts in the migrated image. Note that, however, the hybrid approach uses a parameter defined by the amplitude of the seismic data here. When the parameter changes, we can obtain similar results as that obtained with scaling by its absolute norm. Scaling with logarithmic wavefield (Figure 4.3e) produces better results than the conventional approach. However, the amplitude of the second reflector is affected by other two elliptical events.
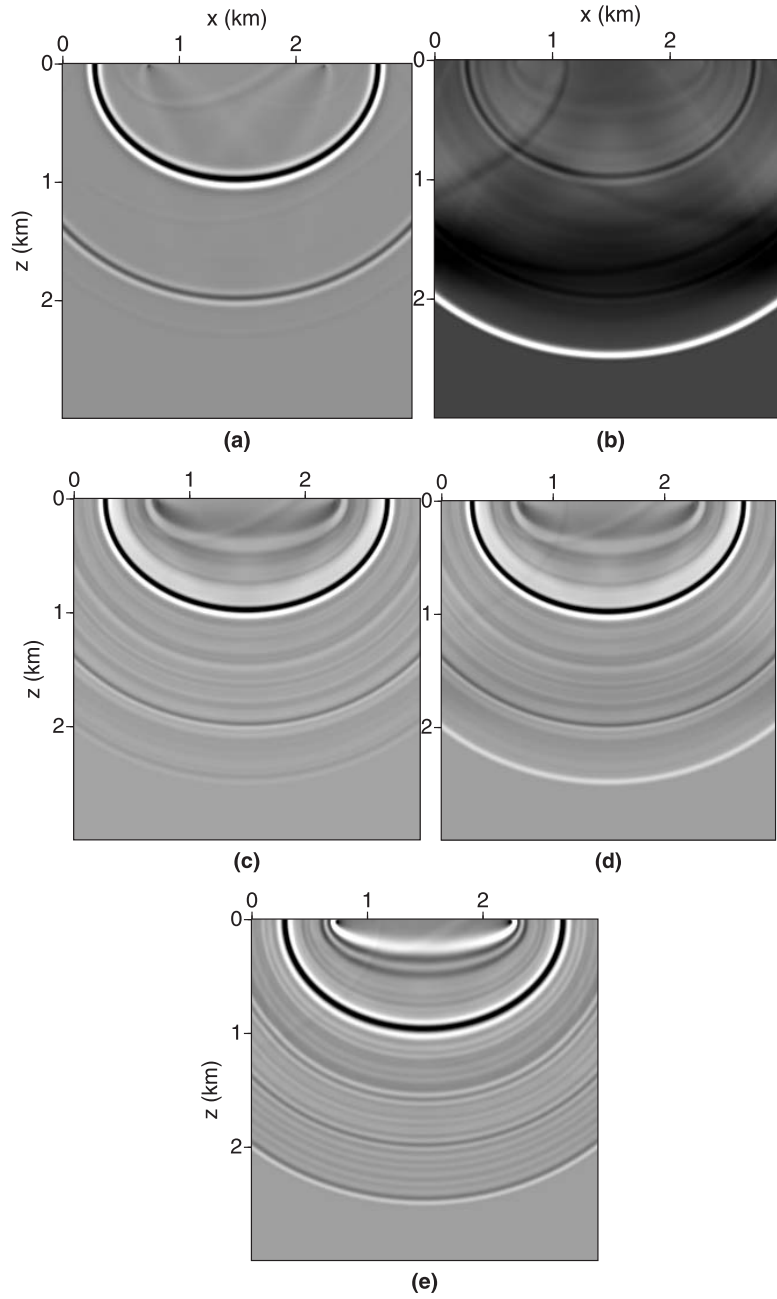
Figure 4.3: Migration kernel of the two reflections for the single source and single receiver experiment. (a) migration of clean data; (b) noisy data with conventional method; (c) noisy data with a L-1 norm scaling approach; (d) noisy data with a hybrid norm scaling approach; (e) noisy data scaling on logarithmic data.
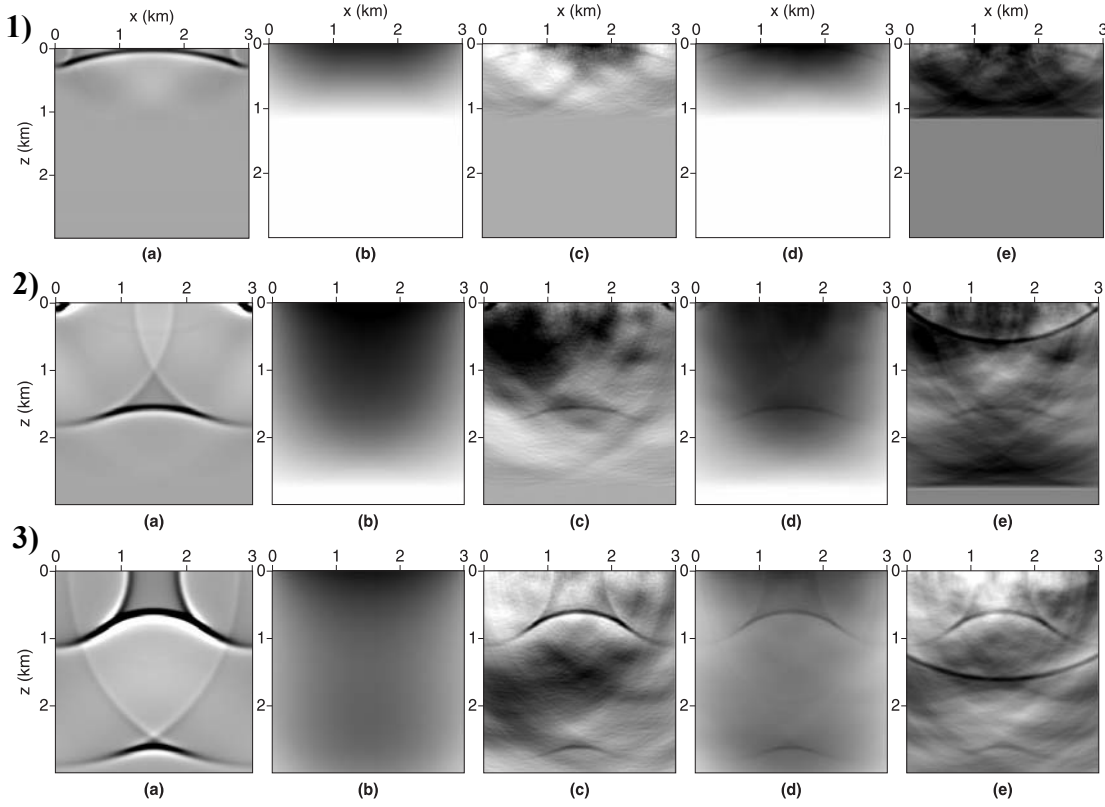
Figure 4.4: Snapshot of the back-propagated receiver wavefield for a source located at the middle of the surface. 1) t=0.6s; 2) t=1.5s; 3) t=2.1s. (a) migration of clean data; (b) noisy data with conventional method; (c) noisy data with a L-1 norm scaling approach; (d) noisy data with a hybrid norm scaling approach; (e) noisy data scaling on logarithmic data.

Figure 4.4 shows the snapshot of the back-propagated receiver wavefield for the forward propagated source located at the middle of the surface. Different scenarios of different scaling approaches are shown for t=0.6s, 1.5s and 2.1s. We observe that when non-Gaussian noise is added, standard back-propagation approach does not provide desired results because we cannot see obvious coherent events on the snapshots. On the other hand, coherent events are pronounced with an absolute norm scaling and hybrid approach, with slight difference in background amplitude. For this case, the logarithm

wavefield approach can retain coherent events, but also introduces some migration artifacts.
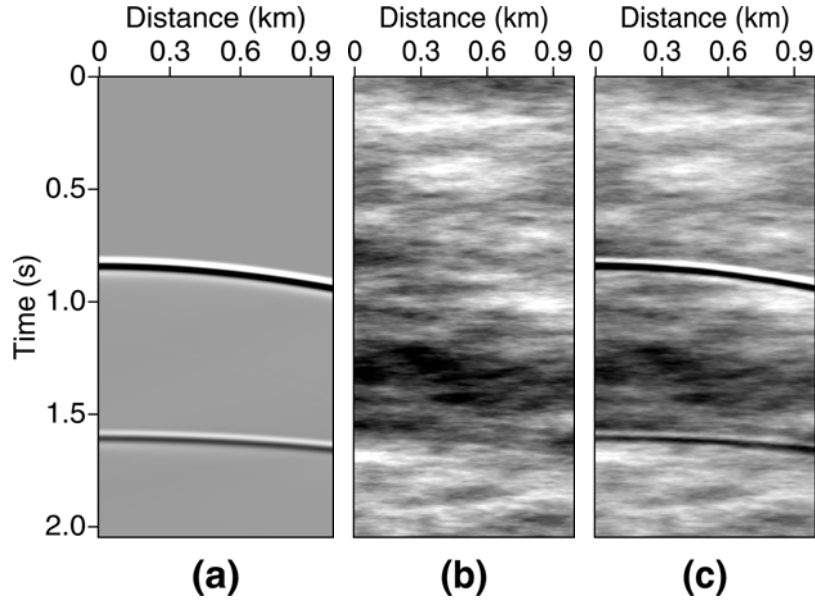
Figure 4.5: Synthetic seismogram for one shot gather with direct wave removed. (a) noise-free seismogram; (b) non-Gaussian noise; (c) noisy seismogram by adding (a) and (b).

The common shot gather for this 1-D model is shown in Figure 4.5. Two parameters for the random perturbations of the data are the horizontal correlation length and vertical correlation length, which construct the anisotropic structure. Correspondingly, the noise does not obtain Gaussian statistics and can be considered as outliers here. The results of the migrated image after a Laplace filter are shown in Figure 4.6. Unlike the example in Figure 4.3, where the wavefield at a single receiver is used, here the data at all the receiver locations are back-propagated simultaneously. This is the practical implementation strategy of RTM for a full acquisition geometry.
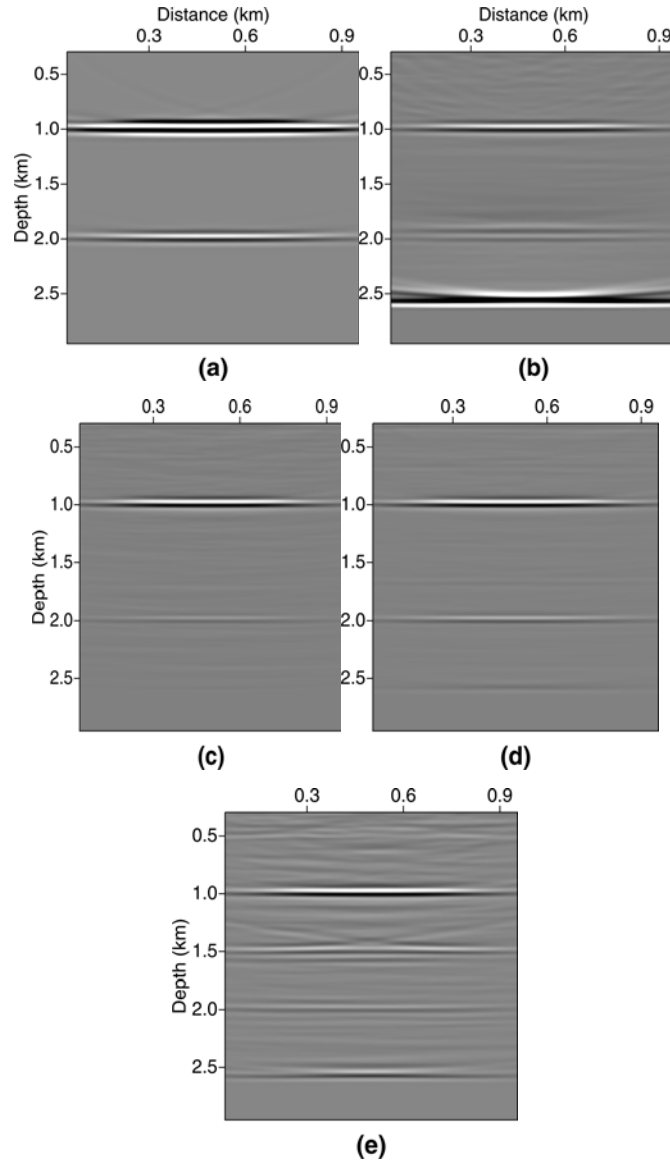
Figure 4.6: Migrated image for a layered medium. (a) migration of clean data; (b) noisy data with conventional method; (c) noisy data with a L-1 norm scaling approach; (d) noisy data with a hybrid norm scaling approach; (e) noisy data scaling on logarithmic data.

Conventional approach (Figure 4.6b) introduces high amplitude artifacts after the second reflector and significantly changes the wavelet of the second reflector. When scaled by its absolute norm (Figure 4.6c), we can obtain cleaner results. A hybrid

approach (Figure 4.6d) can also remove the artifacts but still shows some anomaly at a depth about 2.5 km. Spurious reflectors are created when scaling with logarithmic wavefield (Figure 4.6e).
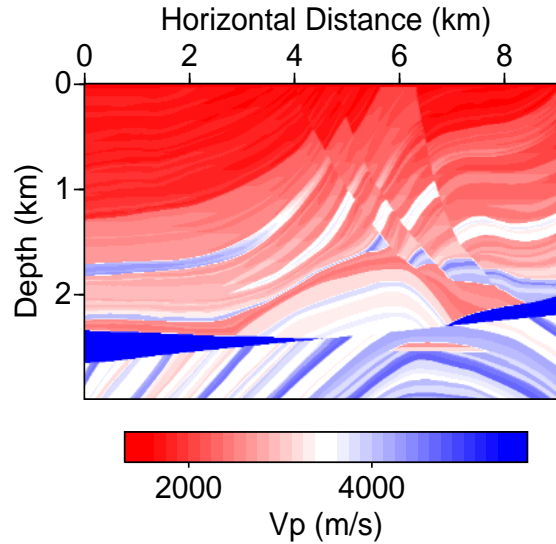
Figure 4.7: The acoustic velocity profile for the Marmousi model.

4.4.2 Random noise with isolated outliers

We then illustrate the use of our approach with random noise and isolated outliers for the data generated with Marmousi model (Figure 4.7). Figure 4.8 shows a shot record with the shot located at the middle of the surface. The outliers for the shot record are generated first by adding non-Gaussian random noise and then by adding an isolated rectangular hole with high amplitude anomaly (Figure 4.8b). This is to simulate extremely poorly processed raw data. Noise is added to each shot record and common image gathers are formed by migrating the shot record separately followed by a stacking to produce the final image.

73
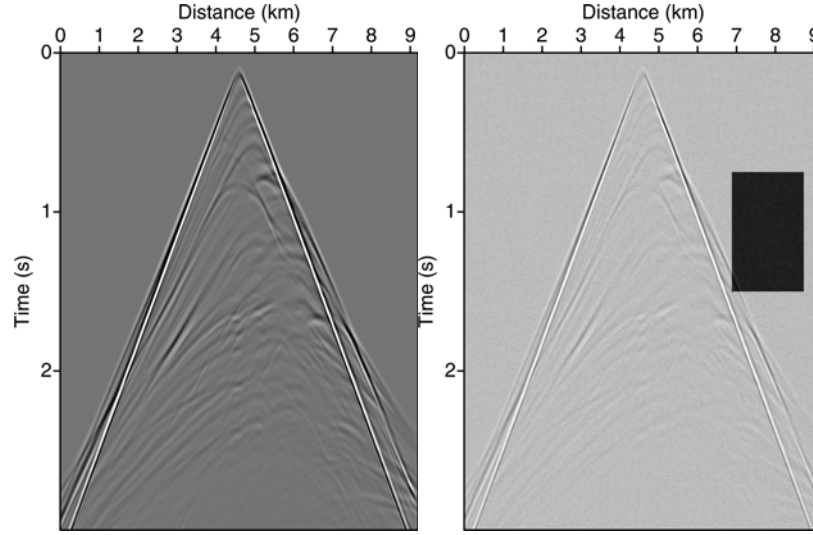
Figure 4.8: A common shot gather for the Marmousi model. Left: clean data; Right: noisy data with non-Gaussian random noises and a rectangular outlier.

Snapshot for the receiver wavefield of the Marmousi model is shown in Figure 4.9. The forward source location is also at the middle of the surface. Since the amplitudes of different snapshots are different, we apply different gains for each group of snapshots. For this type of outliers, the four different methods provide similar results near the beginning and near the end of back-propagation process, with slight differences of the random noise (Figures 4.9.1 and 4.9.3). However, significant difference appears when the outlier record enters the medium (Figure 4.9.2). With a standard approach, most of the signal is contaminated by the noise. With a hybrid approach, the amplitude of the noise is reduced, but it still behaves as a high amplitude anomaly. With the strategy of scaling by its absolute norm, the amplitude is suppressed reasonably well. The logarithm wavefield approach also can suppress the outlier amplitude, but introduces some artifacts.

74
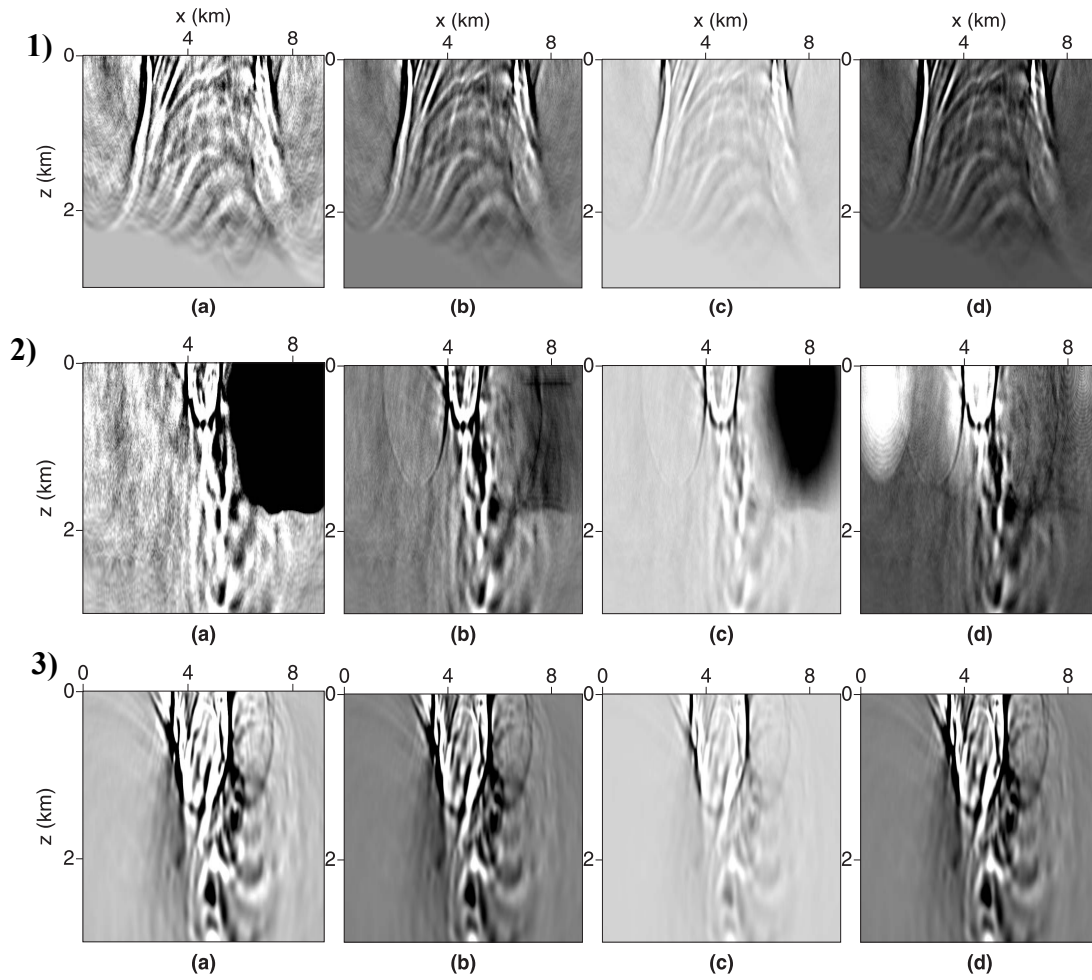
Figure 4.9: Snapshot of the back-propagated receiver wavefield for the Marmousi model. The source is located at the middle of the surface. 1) t=0.8s; 2) t=1.6s; 3) t=2.4s. (a) with conventional method; (b) with scaling by receiver data's absolute norm (L-1 norm approach); (c) with a hybrid norm scaling approach; (d) with scaling on the logarithmic wavefield.
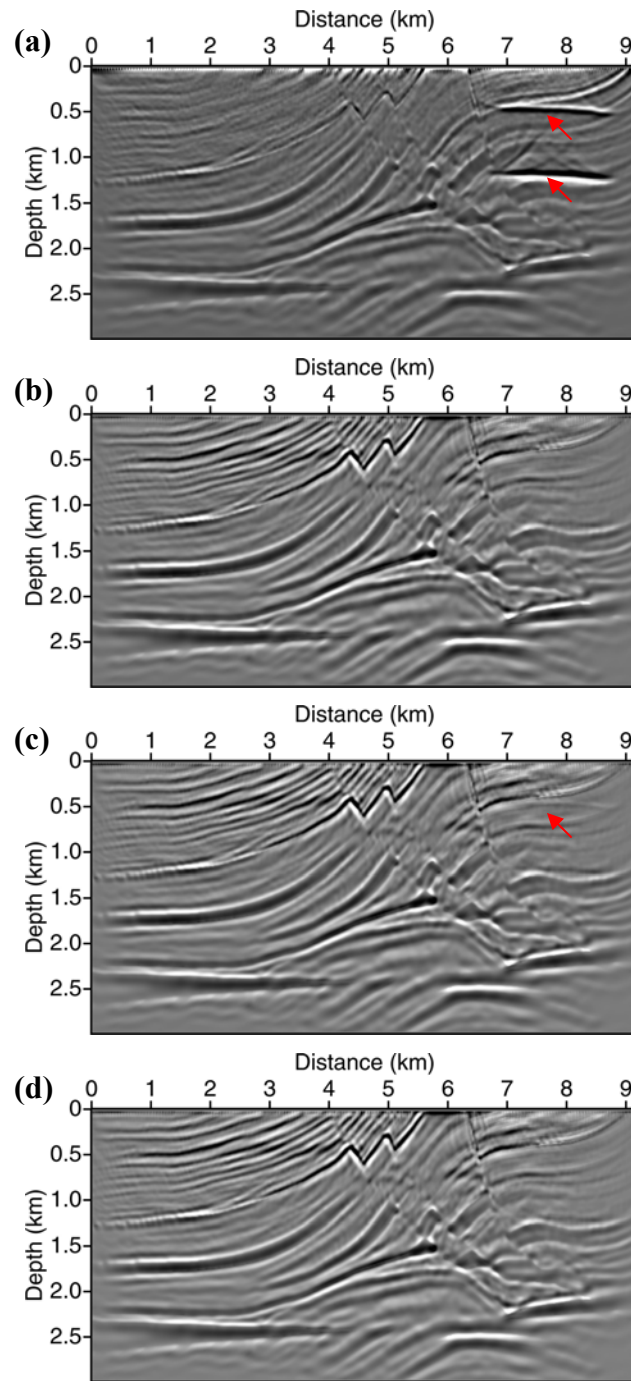
Figure 4.10: Reverse-time migrated image for the Marmousi model. (a) with conventional method; (b) with scaling by receiver data's absolute norm (L-1 norm approach); (c) with a hybrid norm scaling approach; (d) with scaling on the logarithmic wavefield.

Figure 4.10 shows the migrated image from the noisy shot records. With a conventional approach (Figure 4.10a), we can observe anomalies in the right upper part of the model, namely, two dipping events at depths of about 0.5 km and 1.2 km shown as red arrows. Also, upper part of the image has been affected by the random noise. When the receiver wavefield is scaled to its absolute norm (Figure 4.10b), the high amplitude anomaly and the random noise have been removed. The Migrated image with a hybrid scaling approach (Figure 4.10c) shows similar benefits to remove the outliers, but still shows little energy of the artifacts shown with the red arrow. With scaling on a logarithmic wavefield (Figure 4.10d), comparable results as L-1 norm scaling is obtained after stacking all the shots. Logarithmic scaling works better for Marmousi model than for a layered model. This is because the scaling depends on the complexity of the model and characteristics of the noises. Scaling with any approach can produce better image than conventional approach. This suggests that scaling the receiver wavefield can be a good alternative to conventional approach for migration of poorly processed data.

4.5 CONCLUSIONS

I have applied a robust imaging approach for prestack reverse time depth migration. This method borrows ideas from full waveform inversion and is based on back-propagation of the amplitude scaled receiver wavefield. While synthetic examples show that a conventional approach is very sensitive to non-Gaussian outliers of the seismic data, migration artifacts can be significantly reduced with scaled wavefield. Amongst all these methods, scaling by its absolute norm, in general, is most resistant to outliers. However, a quantitative comparison depends on the problems we solve and the parameters we choose.

77

The additional cost here is forward transforming the data into frequency-domain and inverse transforming them back to time-domain. This is negligible considering the cost of wavefield simulation. I restricted the analysis here to 2-D acoustic problems; however, extending this approach to 3D and to elastic cases is trivial.

The discussion in this chapter is based on a regular reverse time migration approach. However, for plane-wave RTM (can be implemented similar to the plane-wave gradient calculation in Chapter 3), I expect that the sensitivity of non-Gaussian noise to the scaling of receiver wavefield should give similar results.

Chapter 5: A robust stochastic inversion workflow for time-lapse data: hybrid starting model and double-difference inversion

5.1 ABSTRACT

Non-uniqueness presents challenges to seismic inverse problems, especially for time-lapse inversion where multiple inversions are needed for different vintages of seismic data. For time-lapse applications, the focus typically is to detect relatively small changes in seismic attributes at limited locations and to relate these differences to changes in the underlying physical properties. In this chapter, I propose a robust inversion workflow where the baseline inversion uses a starting model, which combines a fractal prior and the low-frequency prior from well log data. This starting model provides an estimate of the null space based on fractal statistics of well data. To further focus on the localized changes, the inverted prior from the baseline model and the difference between two time-lapse data are summed together to produce the virtual time-lapse seismic data. This is known as double-difference inversion, which focuses primarily on the areas where time-lapse changes occur. The misfit function uses both data and model norms so that the ill-posedness of the inverse problem can be regularized. The seismic data are pre-processed using a local correlation based warping algorithm to register the time-lapse datasets. Finally, very fast simulated annealing, a non-linear global search method, is used to minimize the misfit function. I demonstrate the effectiveness of our method with synthetic data and field data from Cranfield site used for CO_2 sequestration studies.

5.2 INTRODUCTION

Time-lapse (4D) seismic data is an important tool to monitor temporal changes of reservoir properties associated with hydrocarbon production and environmental engineering. With the improvement of the data fidelity and repeatability in seismic

acquisition, processing and interpretation, 4D seismic data is now used routinely for many reservoir management tasks. Successful applications of time-lapse data have been reported worldwide for different types of reservoirs (Koster et al., 2000; Landrø et al., 2001; Lumley 2001; Hall et al., 2006; Vedanti and Sen, 2009). Future 4D seismic technology is expected to be more sensitive to small dynamic reservoir changes induced by fluid flow and to provide more quantitative interpretations (Calvert, 2005).

Seismic inversion is an important tool for quantitative interpretation of time-lapse seismic data. It minimizes the misfit between observed and modeled seismic data and converts the seismic information into elastic properties such as P-wave velocity, S-wave velocity and density (e.g., Sen, 2006). There are many different types of seismic inversion. They differ mainly from each other by the forward modeling strategy and the usage of travel time and/or amplitude information. Travel time is sensitive to smooth changes or low-frequency variations that are important in defining the geometrical structures at a large scale. Seismic amplitude, on the other hand, is affected by small-scale heterogeneities or high-frequency variations within the seismic resolution range. Travel time tomography (e.g., Ivansson, 1985; Grand, 1987) uses ray tracing as its forward modeling method and uses only travel time information from seismic data. Full waveform inversion (e.g., Tarantola, 1984; Pratt et al., 1998; Sen and Roy 2003) uses wave equation modeling approaches and both travel time and amplitude information. Despite the general success of this method within the seismic imaging community in the last several decades, it is still not applicable at the reservoir characterization scale because of its high computational demands and the local minimum issue. Up to now, seismic inversion with a convolution model as its forward modeling scheme (e.g., Rowbotham et al., 2003; Hampson et al., 2005), is still the most effective and efficient approach for reservoir characterization. This approach typically uses NMO corrected

angle gathers and focuses primarily on the amplitude variation with angle in pre-defined seismic horizons. Under this framework, high-frequency well log data are combined with seismic data to obtain considerably high resolution and geologically continuous elastic properties.

Time-lapse seismic inversion requires robust algorithms because accurate differences in elastic properties are needed to map dynamic reservoir parameters. High-resolution inversion algorithms (Zhang and Castagna, 2011) are typically necessary because small elastic changes are helpful for further quantitative petrophysical analysis. In general, a robust inversion algorithm depends on the starting model (e.g., Tao and Sen, 2012). A fractal starting model extracts the self-similar or self-affine statistics of well logs. It is an effective approach to estimate the high-frequency components not constrained by band-limited seismic data (Srivastava and Sen, 2009 and 2010). This approach overcomes the limitation of random Gaussian statistics in building a high-frequency starting model such as Bayesian search criterion (Tarantola 2005) and Markov-chain Monte Carlo methods (Sen and Stoffa, 1991, 1995; Hong and Sen, 2011). Some aliased estimates, especially at the beginning and end of the well logs can be improved with a combination of the fractal model and a low-frequency starting model (Tao et al., 2011). When a stochastic optimization approach, such as very fast simulated annealing (VFSA) is used, reliable high resolution elastic properties can be obtained from seismic data. This approach can be directly extended to the inversion of 4D seismic data.

A conventional time-lapse inversion workflow requires two separate inversions and a subtraction of the two different inverted datasets to obtain a difference image. Because of noise characteristics and different non-linear search processes, this inversion method can introduce spurious structures instead of real time-lapse signatures in the difference image. Double-difference inversion (Watanabe et al., 2004, Denli and Huang,

2009, Zheng et al., 2011) uses the inverted results from the inversion of baseline data and the difference of the two time-lapse datasets for the inversion of the repeat data. In the second inversion process, inverted elastic properties are taken as prior constraints. Thus, it focuses mainly on the residual data where temporal changes have occurred while keeping other places constant.

In this chapter, we report a robust workflow for the stochastic inversion of time-lapse datasets. This workflow uses the hybrid starting model as a priori constraint for the baseline data and double difference inversion for the repeat data. A local correlation based warping algorithm is used to register the time-lapse datasets to improve on the consistency of time-lapse signatures. VFSA is used as the non-linear global search approach to find the minimum of the misfit function. This approach is justified by synthetic data and by comparison with the conventional inversion workflow. We further demonstrate the feasibility of this approach with a field dataset from Cranfield CO_2 sequestration site.

5.3 METHODOLOGY

5.3.1 Hybrid starting model

Synthesizing well logs with fractal statistics is justified by the fact that self-similar and self-affine statistics can be approximated by fractal heterogeneities (Hewett 1986; Stefani and Gopa, 2001; Browaeys and Fomel, 2009). Within a fractal framework, elastic and petrophysical properties of the subsurface can be represented by a smooth background trend at a large scale plus fluctuations at finer scales

$$f(\mathbf{r}) = f_0(\mathbf{r}) + \sigma(\mathbf{r}),$$
(5.1)

where $f(\mathbf{r})$ is any elastic or petrophysical parameter, $f_0(\mathbf{r})$ is the background parameter which can be defined deterministically, and $\sigma(\mathbf{r})$ is the spatial fluctuation which can be synthesized stochastically.

The fractal series $\sigma(\mathbf{r})$ has a zero expectation value, a determined spatial covariance and a power-law dependence of its Fourier spectrum

$$P(k) = k^{-\beta} \quad , \tag{5.2}$$

where $P(k)$ is the energy spectrum, k is the wavenumber and β is a scaling factor that is linearly related to the Hurst coefficient H (Hurst et al., 1965). Specifically, $\beta = 2H - 1$ is for fractal Gaussian noise and $\beta = 2H + 1$ is for fractal Brownian motion.

To examine if a time signal or a space signal has strong or weak fractal behavior, spectral analysis is used, and the value of β is examined. The correlation of adjacent points is measured as persistence. For a signal with $\beta = 0$, the energy spectrum is independent of frequency and the signal consists of uncorrelated random noise. When $\beta > 0$, the signal is positively correlated and the data is persistent. When $\beta < 0$, the signal is anti-correlated and anti-persistent.

Figure 5.1 shows four different velocity curves and Figure 5.2 shows the corresponding log-log Fourier spectrum. Although only velocity is shown here, spectra of other parameters, such as density, impedance, permeability and porosity, can also be depicted this way. Figure 5.1a is a blocky velocity model that represents layered strata that can generate seismic reflections. A blocky model consists of piecewise Heaviside functions while its amplitude spectrum is a superposition of *sinc* functions (Figure 5.2a). Figure 5.1b is a smoothed low-frequency background velocity trend that can be used as a low-frequency starting model for inversion. Its spectrum does not have perturbations and it typically cannot be accurately synthesized by a fractal series (Figure 5.2b). For a

typical well log shown in Figure 5.1c, fractal statistics provide a good approximation of it considering power-law dependence of its spectrum (Figure 5.2c). A random series (Figure 5.1d) can be considered as a special case of a fractal series when the least-square-fit slope of its log-log spectrum equals one (Figure 5.2d).
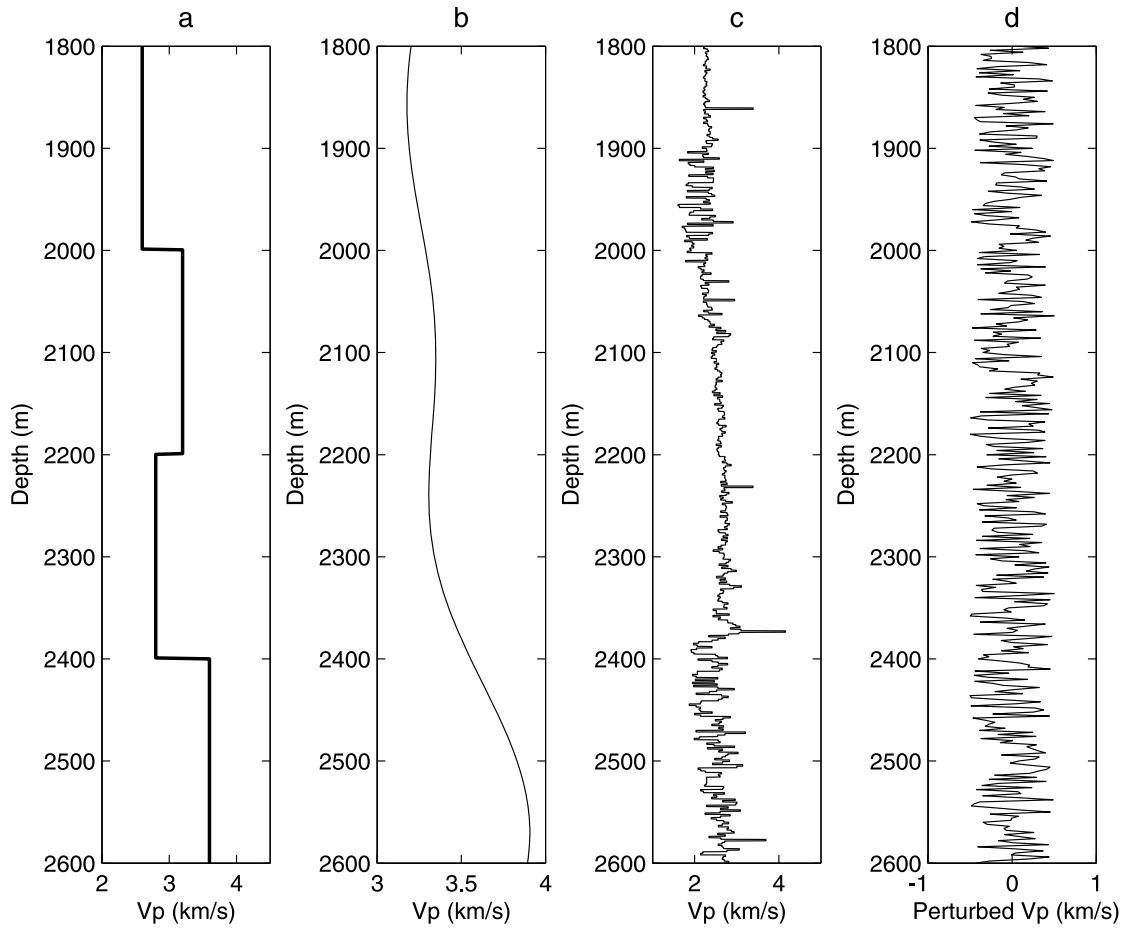
Figure 5.1: (a) A blocky heterogeneous velocity curve; (b) a smooth low frequency velocity trend derived from a well log; (c) P-wave velocity from a real well log data; (d) random velocity perturbations.
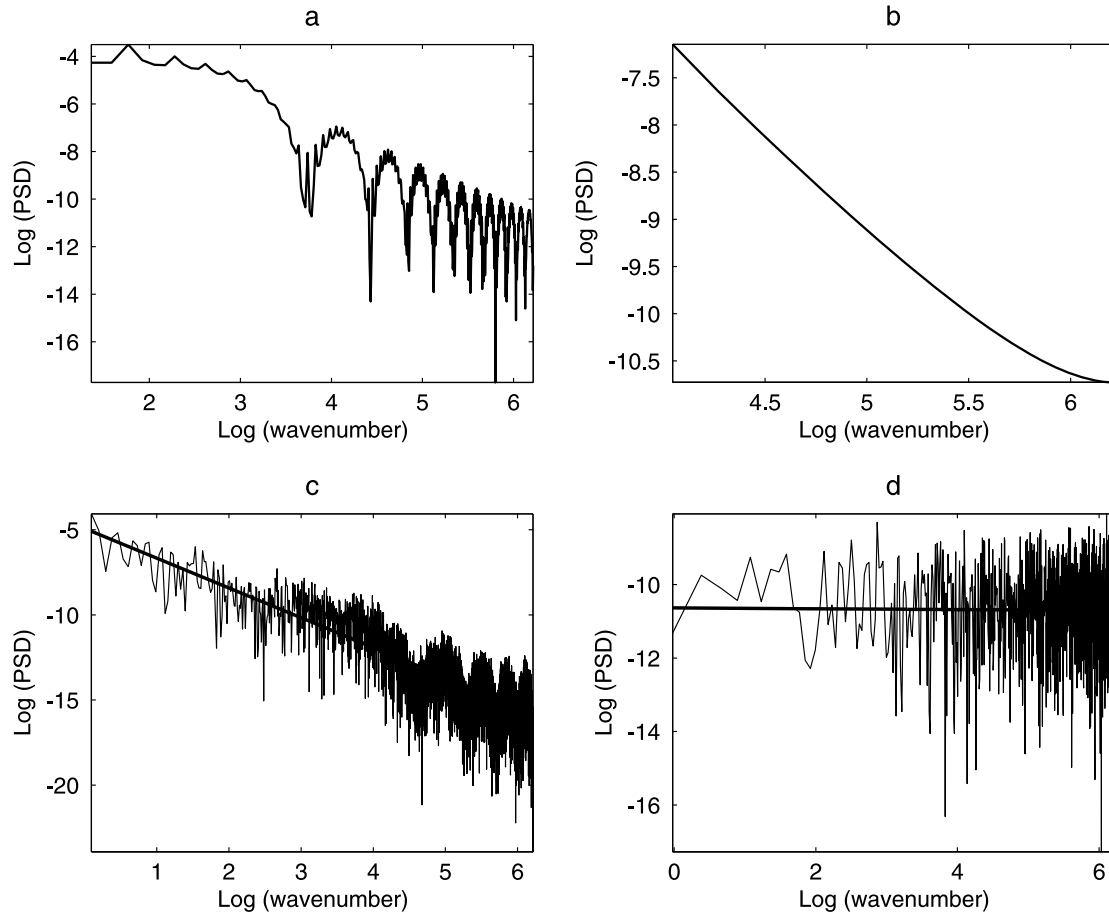
Figure 5.2: Log-log spectra for (a) A blocky velocity; (b) a smooth velocity; (c) P-wave
velocity from a real well log data; (d) random velocity perturbations. A
linear trend is estimated for (c) and (d).

To describe a given well log data using fractal statistics, three parameters should

be estimated from the well log data: the expectation, the standard deviation and the Hurst

coefficient. Expectation and standard deviation can be easily estimated. To estimate the

Hurst coefficient, several methods are available. Following Srivastava and Sen (2009),

we use a method known as rescaled range analysis

$$\frac{R}{S} = \left(\frac{N}{2}\right)^H \quad , \tag{5.3}$$

where R and S are the range and standard deviation, respectively, of the given dataset. N is the number of sampling points. If we plot log(R/S) versus log(N/2), we obtain a straight line with the slope that is the Hurst coefficient (H).

After H is estimated, I synthesize the log curve with a fractal Gaussian series. This involves a process to generate random perturbations with an estimated auto-covariance and H (Srivastava and Sen, 2010). The auto-covariance of the fractal Gaussian noise can be given by

$$\mathbf{A}(t) = 0.5\sigma^2 \left(|t+1|^{2H} - 2|t|^{2H} + |t-1|^{2H}\right) \quad , \tag{5.4}$$

where $\mathbf{A}$ is the auto-covariance, σ is the standard deviation of the well log data.

A fractal initial model typically has the same frequency range as that of the well log data. This provides a good estimate of the null space. However, this method can introduce spurious frequency components when the well log does not strictly satisfy fractal statistics. Random realizations of well data can also bring frequency uncertainties into the model space. This can be compensated for with a hybrid model approach. Suppose we have the low-frequency model $\mathbf{m}_1$ and the fractal model $\mathbf{m}_2$, the hybrid starting model $\mathbf{m}$ is given by

$$\mathbf{m} = \alpha \mathbf{m}_1 + (1-\alpha)\mathbf{m}_2 \quad , \tag{5.5}$$

where α is a weight number such that $0 \le \alpha \le 1$. Setting $\alpha = 0.5$ gives the low-frequency trend and high-frequency components equal weight.

Figure 5.3 shows a comparison between synthetic well logs generated using the fractal method and the hybrid method. Compared to the real impedance log, the pure fractal model has some deviations, especially at the start and end depths of the log. The

86

root mean square error is 19.4 km/s*g/cc for the pure fractal model, and is 10.5 km/s*g/cc for the hybrid model. Our hybrid model better matches the true answer.
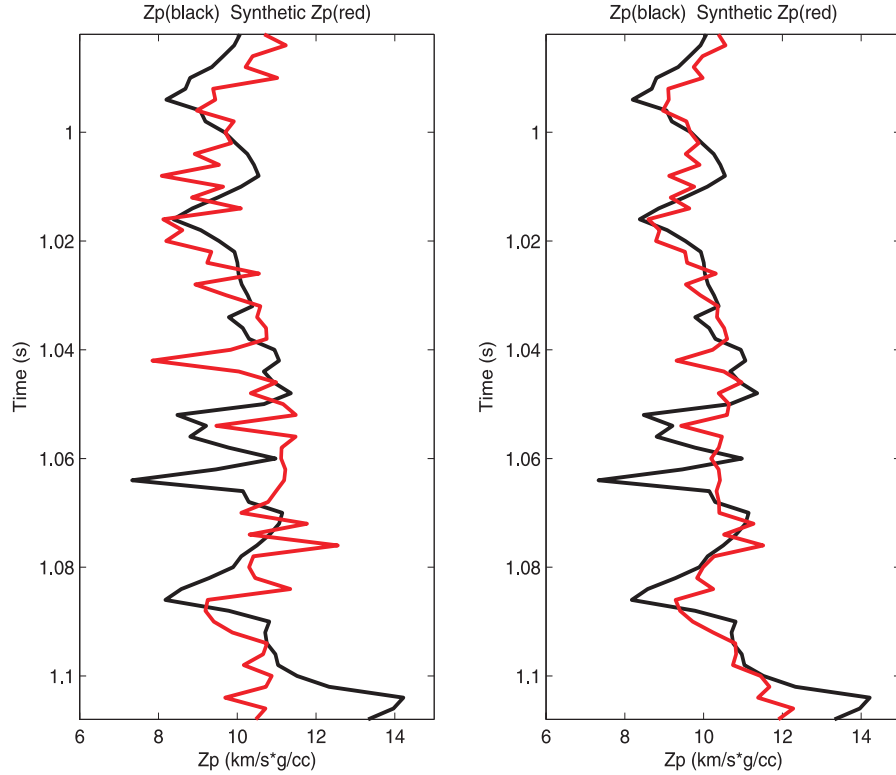
Figure 5.3: Comparison of synthetic well log P-impedance (red) with real well log data (black). Left: fractal model. Right: hybrid mode

5.3.2 Forward modeling and optimization scheme

This chapter uses a convolution model as its forward modeling scheme. The modeling assumes a layered Earth. Inversion of the seismic data is based on a trace-by-trace approach and can be applied to 3D datasets with horizon constraints. The forward model is given by

$$\mathbf{S}(\theta,t) = \mathbf{w}(t) * \mathbf{R}(\theta,t) + \mathbf{n}(t) \quad , \tag{5.6}$$

where $\mathbf{S}(\theta,t)$ is the observed seismic data, typically an angle gather in a prestack case; $\mathbf{w}(t)$ is the source wavelet, and $\mathbf{n}(t)$ is the noise; $\mathbf{R}(\theta,t)$ is the reflectivity, and the angle-dependent reflectivity can be given by a linear approximation of the Zoeppritz equations (e.g. Fatti et al., 1994). A full Zeoprittz equation with 1D reflectivity modeling can also be used (e.g. Sen and Roy, 2003). This chapter uses the convolution model for numerical simplicity and efficiency.

Most real inverse problems are ill-posed. Regularization is necessary to make the inversion process more robust. Similar to Srivastava and Sen (2010), we choose the misfit function between the observed and calculated data as

$$M = \frac{2\sum |\mathbf{d}_{obs} - \mathbf{d}_{cal}|}{\sum |\mathbf{d}_{obs} - \mathbf{d}_{cal}| + \sum |\mathbf{d}_{obs} + \mathbf{d}_{cal}|} + \varepsilon_1 \sum |\mathbf{m}_{new} - \mathbf{m}_0| + \varepsilon_2 \sum |\mathbf{A}_{obs} - \mathbf{A}_{cal}| \quad , \tag{5.7}$$

where $\mathbf{d}_{obs}$ and $\mathbf{d}_{cal}$ are the observed and calculated seismic data, respectively; $\mathbf{m}_{new}$ and $\mathbf{m}_0$ are the model parameter of new iteration and the starting model, respectively; $\mathbf{A}_{obs}$ and $\mathbf{A}_{cal}$ are the observed and calculated autocorrelation series; ε_1 and ε_2 are weighting factors.

The optimization method used in this chapter is VFSA. VFSA is an efficient global searching algorithm with its ability to identify optimal parameters for non-linear problems (Sen and Stoffa, 1995). This algorithm requires the model space to be subdivided into equally spaced intervals. Within different iterations, the temperature, i.e., the control parameter, is selected based on an exponentially decaying cooling schedule. Starting from an initial model, which is typically random but here it is a fractal based hybrid model, a new model is accepted by evaluating the probability of the misfit function. This search process is typically known as the Metropolis algorithm (Metropolis

88

et al., 1953). The best-fit model is accepted by repeating this inversion process to reduce the bias of sampling in the model space. In theory, global optimization based on Monte Carlo searching can be applied to other computationally intensive inverse problems such as full waveform inversion with two-way wave equation forward modeling (e.g., Tarantola, 1984). However, because the number of forward calculations is typically prohibitively large, global optimization methods are still considered impractical for those problems.

5.3.3 Double-difference inversion

Double-difference inversion for 4D seismic data has been applied in travel time tomography (Waldhauser and Ellsworth, 2000) and full waveform inversion problems (Watanabe et al., 2004; Denli and Huang, 2009). Here we extend this idea to the stochastic time-lapse inversion.

Conventional inversion strategy for time-lapse data involves two independent inversions. Changes in the underlying elastic properties are given by a subtraction between those model parameters. This can be written as

$$\delta \mathbf{m} = \mathbf{m}_{repeat} - \mathbf{m}_{base} \quad , \tag{5.8}$$

where $\delta \mathbf{m}$ is the time-lapse changes of the model parameters. $\mathbf{m}_{repeat}$ and $\mathbf{m}_{base}$ are the inverted model parameters for the repeat data and baseline data, respectively. The two different inversions typically require a same inversion workflow to ensure that the inversion process itself does not introduce time-lapse difference. However, because of non-uniqueness, independent inversions of these two different datasets may converge into different results that are not true time-lapse signatures.

Similar to joint 4D inversion proposed by Johnson et al. (2009), where geophysical parameters are used to constrain the inversion of hydrogeologic data,

inverted parameters from the baseline data is used as a constraint for the repeat data for double-difference inversion. The observed repeat data is replaced by the simulated data of the baseline model and the residual data

$$\mathbf{d}_{obs\_repeat} = f(\mathbf{m}_{base}) + \Delta \mathbf{d}_{obs} \quad, \tag{9}$$

where $f(\mathbf{m}_{base})$ is the calculated data for the baseline model. Δd_{obs} is the residual data between two different datasets and $\Delta \mathbf{d}_{obs} = \mathbf{d}_{obs\_repeat} - \mathbf{d}_{obs\_base}$.

This method is equivalent to conventional time-lapse inversion if the inversion performed on the baseline data does not truly match the observed data (Zheng et al., 2011). However, data matching cannot be perfect for real inverse problems. Correspondingly, the inverted time-lapse difference for double-difference inversion and conventional approach are typically different. Because $\Delta \mathbf{d}_{obs}$ corresponds to the localized changes in the time-lapse survey, this method only inverts for the actual differences in the seismic data. Therefore, it better constrains the inversion of the repeat data.

5.3.4 Local correlation-based warping

Seismic data for time-lapse purposes are typically processed through multiple independent workflows. At any step, if not processed properly, artifacts can be brought into the final image. Although modern acquisition techniques have increased the repeatability of time-lapse datasets, acquisition can still introduce unwanted artifacts. To correctly identify the time-lapse signature, we use a warping method to isolate the changes at the reservoir interval from its surrounding environments. Warping was firstly investigated as a cross equalization method on 4D seismic image registration (Rickett and Lumley 2001; Druzhinin and MacBeth, 2001).

In this chapter, the warping method is applied on the repeat data based on a maximum local correlation trend (Fomel, 2007; Fomel and Jin, 2009). This approach

90

involves the calculation of the local correlation, followed by squeezing and stretching the stacked image of the repeat data with reference to the baseline data. This is a trace-by-trace approach, which implements a shaping regularization to stabilize the local-correlation output. The maximum local-correlation trend is picked automatically as same as semblance picking in velocity analysis (Fomel, 2009).

5.4 NUMERICAL EXAMPLES

5.4.1 Synthetic validation

Synthetic examples are used to validate the effectiveness of our inversion workflow. Since this inversion requires an initial model with high-frequency components, we create 1D synthetic data and 2D synthetic data based on the hybrid initial model from Hampson-Russell's demonstration datasets (Srivastava and Sen, 2010). This impedance model also involves extrapolating the impedance well log based on picked horizons. Seismic dataset is created by convolving a wavelet with the reflectivity derived from the impedance model.

Figure 5.4 shows the 1-D P-impedance well log data and its corresponding inverted result. The time-lapse impedance model is created by putting an impedance difference of 1860 m/s*g/cc from 1.02s to 1.05s. This is used to simulate the effect of injected CO_2. We use a constant impedance difference for all the sampling points within a reservoir interval to simplify the fluid substitution and to quantitatively analyze how our inversion workflow resolves the wedge model in the following 2-D model. From the inverted result of the baseline model, we observe that, in general, stochastic inversion using a hybrid initial model can provide effective estimate of the true impedance, despite some mismatches because of the tuning effects of seismic data.
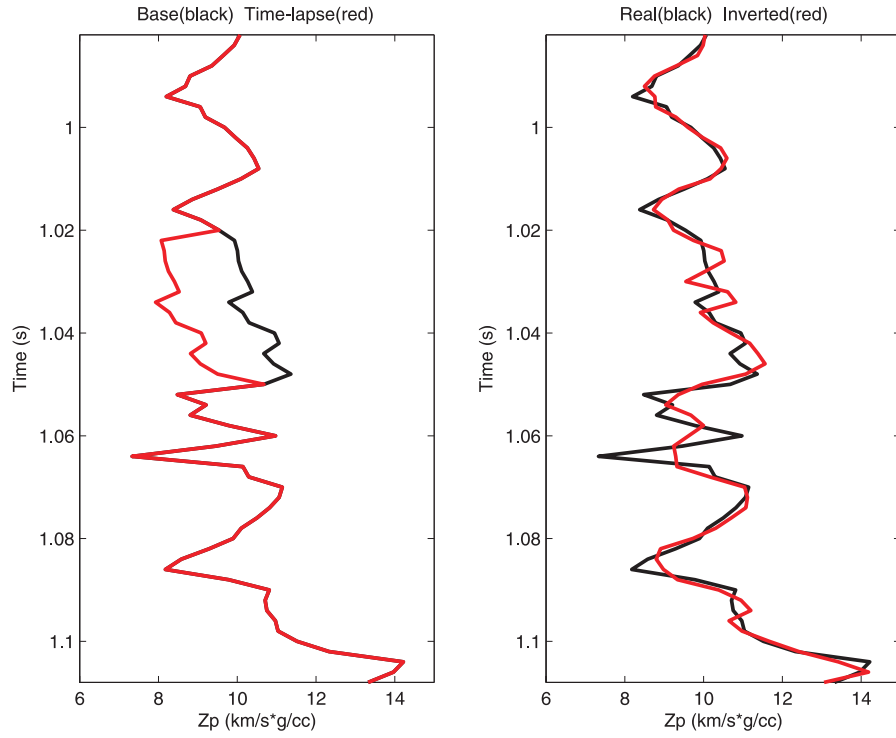
Figure 5.4: Left: P-impedance log data of the baseline model (back) and the time-lapse model (red); Right: Comparison of inverted impedance (red) for the baseline model with the true model (black).

Figure 5.5 shows the inverted best-fit P-impedance difference of the 1D model using a conventional approach (two separate independent inversions) and double-difference inversion. Both approaches detect the decreased impedance change due to CO_2 injection. However, the magnitude of the impedance difference with the conventional approach is smaller than the true difference, and there are some deviations outside the target interval. In comparison, inversion with the double-difference approach produces better matching of the target interval and fewer deviations outside of it. For stochastic inversion, multiple realizations produce slightly different inverted results because the optimization strategy randomly selects different model parameters. Figure 5.5b shows 20 realizations of the inverted differences. We observe that the deviations for different

92

inversions are quite small, especially for the double-difference inversion. This suggests that our inversion workflow is stable.
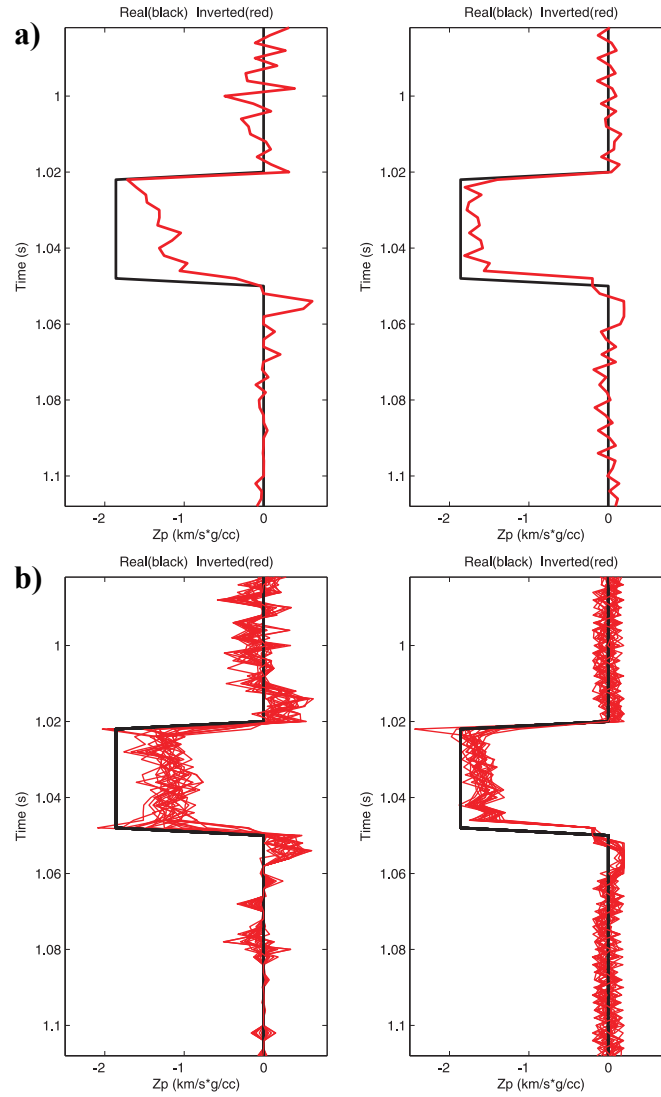
Figure 5.5: (a) Comparison of the inverted impedance difference with conventional approach (left) and the double-difference inversion approach (right); (b) multiple realizations of (a). For all of those plots, black denotes the true difference and red denotes the inverted difference.
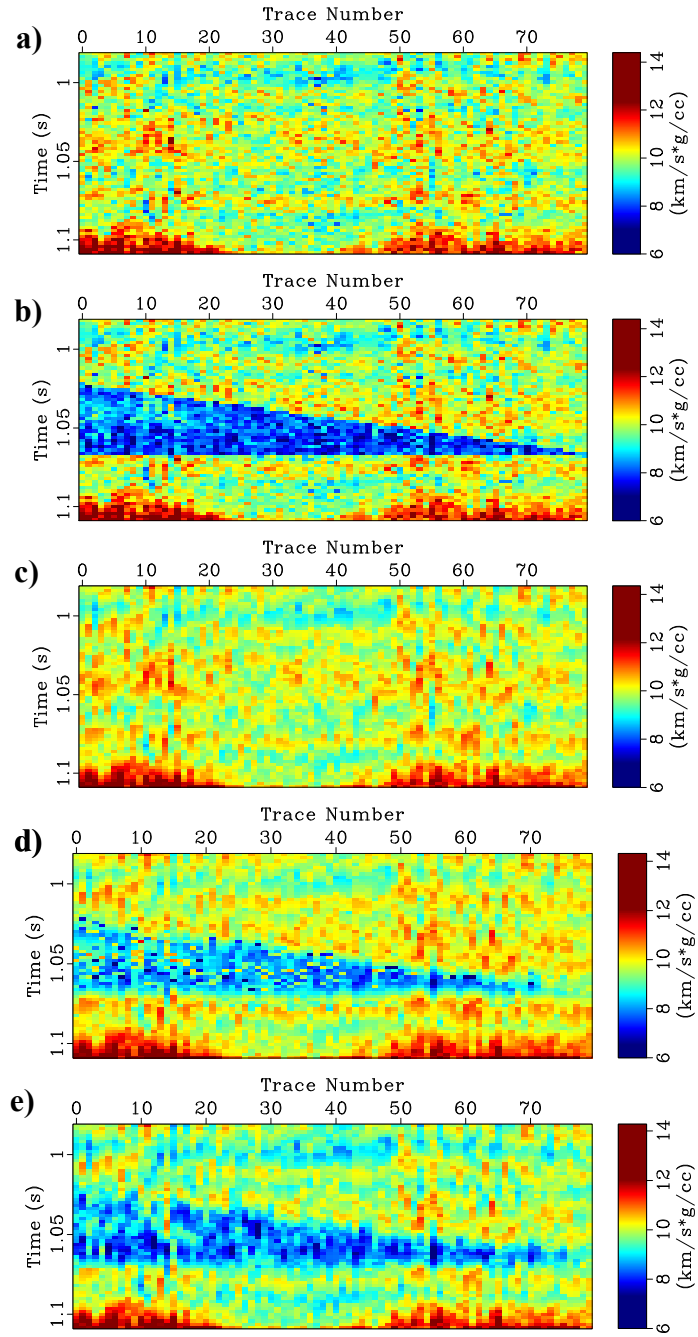
Figure 5.6: P-impedance plots for (a) baseline model; (b) time-lapse model; (c) inverted base line model; (d) inverted time-lapse model with conventional approach; (e) inverted time-lapse model with double-difference inversion.
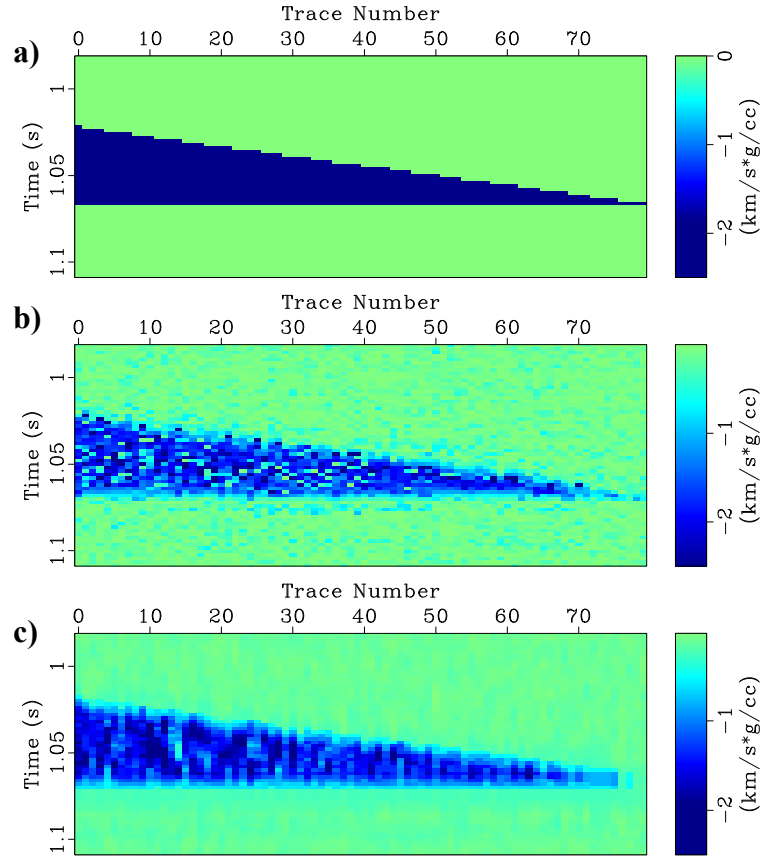
Figure 5.7: Comparison of the subtracted P-impedance with conventional approach and with double-difference inversion. (a) True difference; (b) inverted difference with conventional approach; (c) inverted difference with double-difference inversion.

Figure 5.6 shows 2D true impedance models and their corresponding inverted models with the two approaches. Figure 5.7 shows the impedance differences. The geometry is a wedge model to test how our algorithm can resolve the tuning effects of thin-bed layers. Similar to the 1D scenario, for each case, inverted impedance model is close to the true model. Although the inverted difference becomes smaller than the true answer when the layer becomes thinner, both of the approaches could resolve the wedge model very well. The inverted difference with a double-difference inversion approach matches the shape of the wedge better than conventional approach. Double-difference

inversion also shows better match in terms of amplitude of the difference inside the wedge model and better constrains the deviations outside the places where temporal changes occur.

5.4.2 Field data applications

The field data used in this study comes from Cranfield in southwest Mississippi. The Gulf Coast Carbon Center (GCCC) at the University of Texas used this site for CO_2 injection at twenty-three drilled well locations. The injection interval is the lower Tuscaloosa formation from 3012m to 3142m below the surface. From year 2008 to year 2010, a cumulative mass of 2.2 million metric tons was injected. To monitor this CO_2 injection process, pre-injection seismic data was acquired in the year 2007 before the CO_2 injection, and post-injection seismic data was acquired in the year 2010 (Hovorka et al., 2011). The injection interval of the lower Tuscaloosa formation at Cranfield appears as a thin bed layer with the thickness around 15 meters, and no overburden (low-velocity and high pressure) emission is detected from the in situ well-log measurements. This leads to an assumption that the time shifts at this Cranfield datasets are aliasing because injected CO_2 will not change the thin bed characteristics of the injection interval. This confirms that a warping method is necessary to identify the true time-lapse signature for this dataset.

Figure 5.8 shows a cross-section of the pre-injection seismic data and the calculated time-lapse differences. For comparison, the differences calculated using the warped and non-warped post-injection data are shown. Within the injection interval (the two picked horizons in Figure 8b and 8c), the original difference (Figure 8b) shows the opposite sign of the amplitude for different crossline locations. There is some mis-alignment of the time-lapse signatures for the previous processing. After warping,

96

seismic differences become laterally continuous for all the crossline locations. This is because the time-shift has been separated from amplitude changes (Fomel and Jin, 2009). We think warping might help improve the fidelity of time-lapse interpretation. However, we cannot rule out the possibilities that data without warping are more close to the true time-lapse signatures because of the heterogeneities in the injection interval.
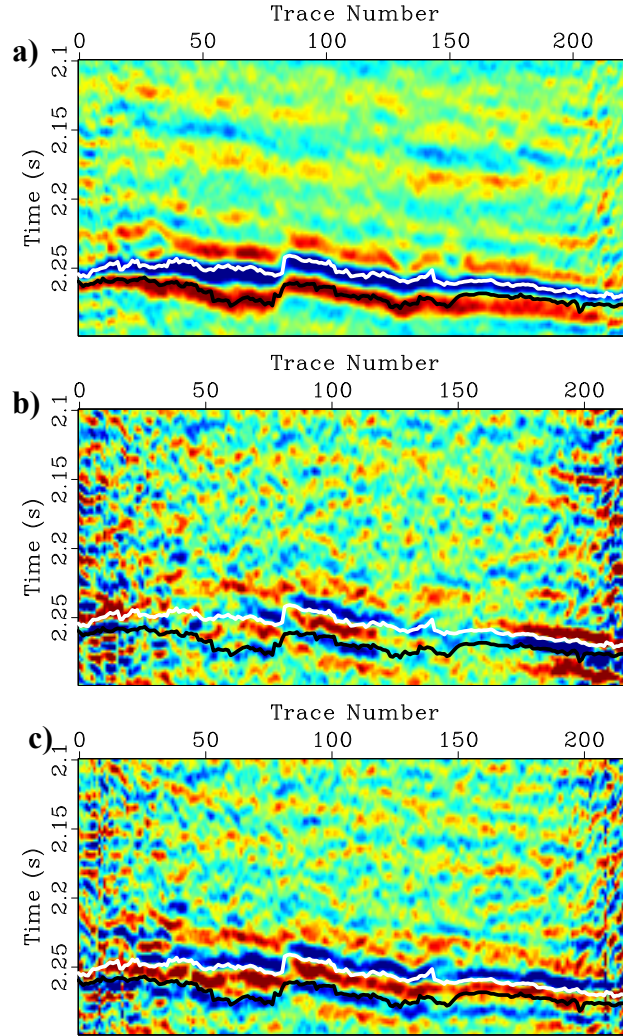
Figure 5.8: A cross-section of (a) Pre-injection seismic data; (b) Time-lapse difference without warping; (c) with local-correlation based warping. Two horizons are overlaid to this plot. The white denotes the top sand and base denotes the bottom sand of the injection interval.

Well-to-seismic analysis was performed to find optimal wavelets, which ensure well log data ties well with seismic section. The extracted wavelet and its corresponding spectrum are shown in Figure 5.9. This wavelet is close to zero phase. After pre-processing with warping and wavelet extraction, the well log data was resampled to the seismic sampling interval and extrapolated to the entire seismic section with the guidance of picked horizons. A hybrid starting model was then built based on the extrapolated well logs and used for the inversion of the pre-injection data.
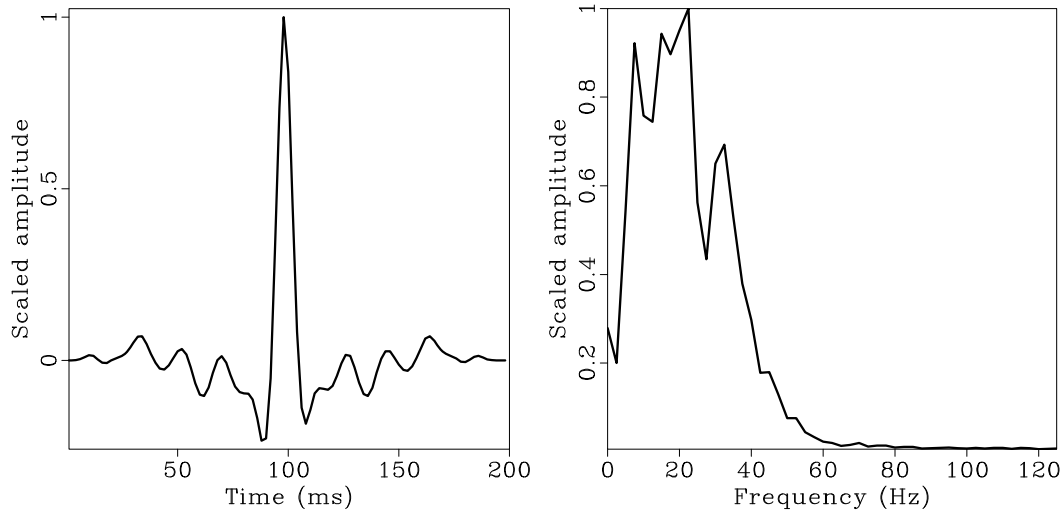
Figure 5.9: Wavelet used for the Cranfield datasets. Left: wiggle plot of the wavelet; Right: its corresponding spectrum.

Figure 5.10 shows the inverted P-impedance for the pre-injection data. There is one well located in this seismic section. A comparison of the inverted result at the well location is shown on the right. In general, the inverted impedance agrees with the true impedance. Within the injection interval, the impedance is apparently lower than the surrounding environment. The convergence history at the well location is shown in

Figure 5.11. This resembles a typical convergence history of the VFSA optimization scheme. At the beginning, the misfit has some fluctuations. After 400 iterations, this searching is close to the global minimum and begins a slow searching process. The observed seismic data and inverted seismic data are shown in Figure 5.12. Most of the key features of the observed seismic data are captured by the simulated data using the inverted impedance.
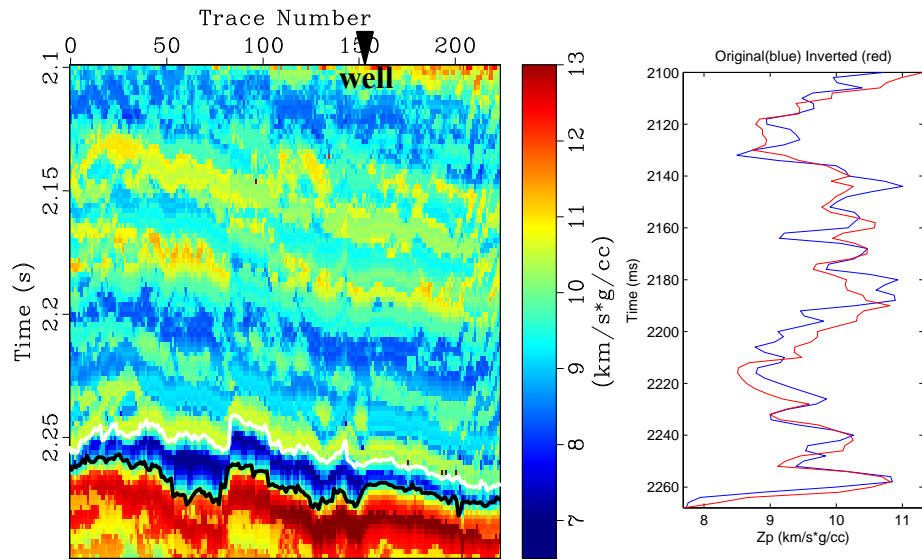
Figure 5.10: Inverted P-impedance for the pre-injection data (left). Right shows the true impedance (scaled to the seismic data sampling interval) and inverted result at the well location.

Figure 5.13 shows the subtracted P-impedance for different scenarios. Within the injection interval, there are some discrepancies between the conventional approach (Figures 5.13a and 5.13b) and double-difference inversion (Figures 5.13c and 5.13d). The most notable discrepancy is that the inverted difference for double difference inversion has a longer time span. As analyzed by the synthetic examples, we think that the inverted results with a double difference approach can better capture the true time-lapse

difference. Also, conventional inversion (Figures 5.13a and 5.13b) has some fluctuations of the inverted difference outside the CO_2 injection interval. Similar to the seismic data, misalignment of the impedance difference exist for the non-warped seismic data (Figures 5.13a and 5.13c). Inverted results show same negative change of P-impedance across the picked horizons. This is consistent with the existing geological knowledge of this area, which suggests that rock properties within the injection interval are laterally consistent. There is an impedance decrease below the picked top sand horizon on the inverted warped result (Figures 5.13b and 5.13d). This confirms the conjecture that most CO_2 accumulates at the top of the injection interval.
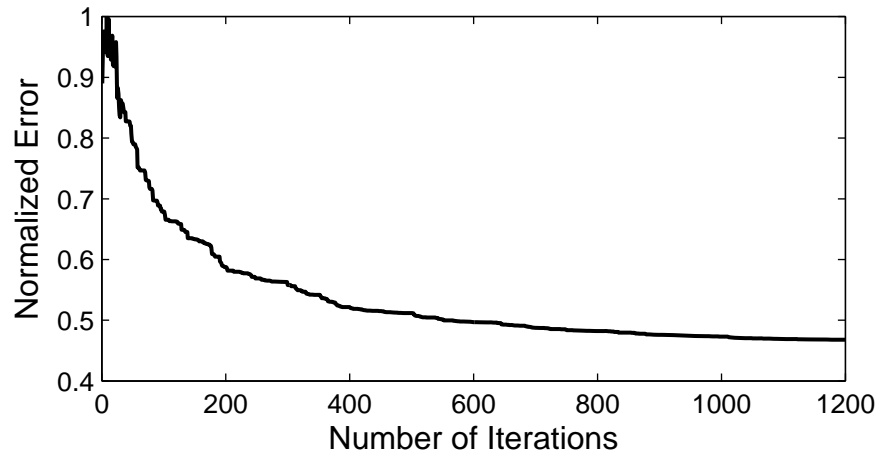
Figure 5.11: Normalized misfit versus number of iterations at the well location.

5.5 CONCLUSIONS

In this chapter, I have presented a robust stochastic time-lapse inversion strategy. This workflow involves two key steps. The first step is to invert the baseline model with a hybrid starting model which combines the fractal prior and low-frequency prior of the well log data. The second step is to use a double-difference inversion scheme to focus on the local areas where time-lapse changes have occurred. I find that double difference

100

inversion is able to capture the time-lapse differences better than conventional separate inversion approach. Synthetic and real data example shows that inversion with double difference data could focus on the places where time-lapse signatures have occurred. In addition, inversions with warped data show continuous change of the reservoir properties.
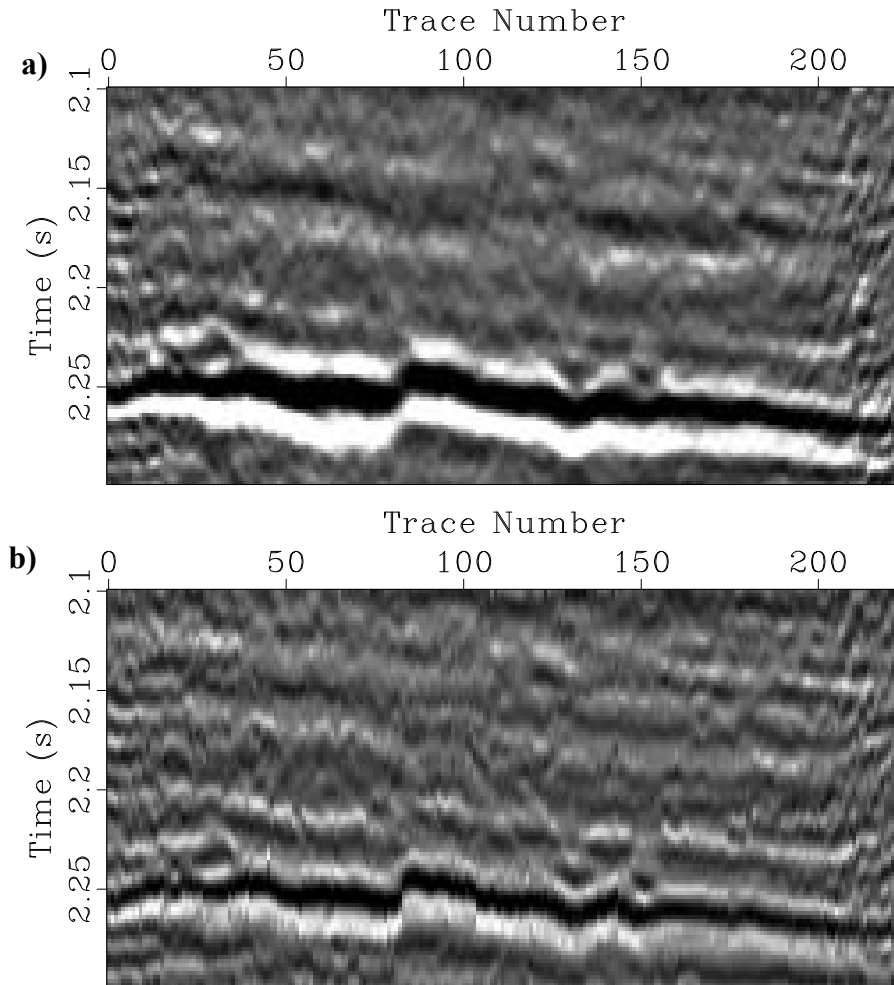
Figure 5.12: Comparison of the pre-injection seismic data and inverted result. (a) Observed seismic data; (b) simulated data with inverted P-impedance.
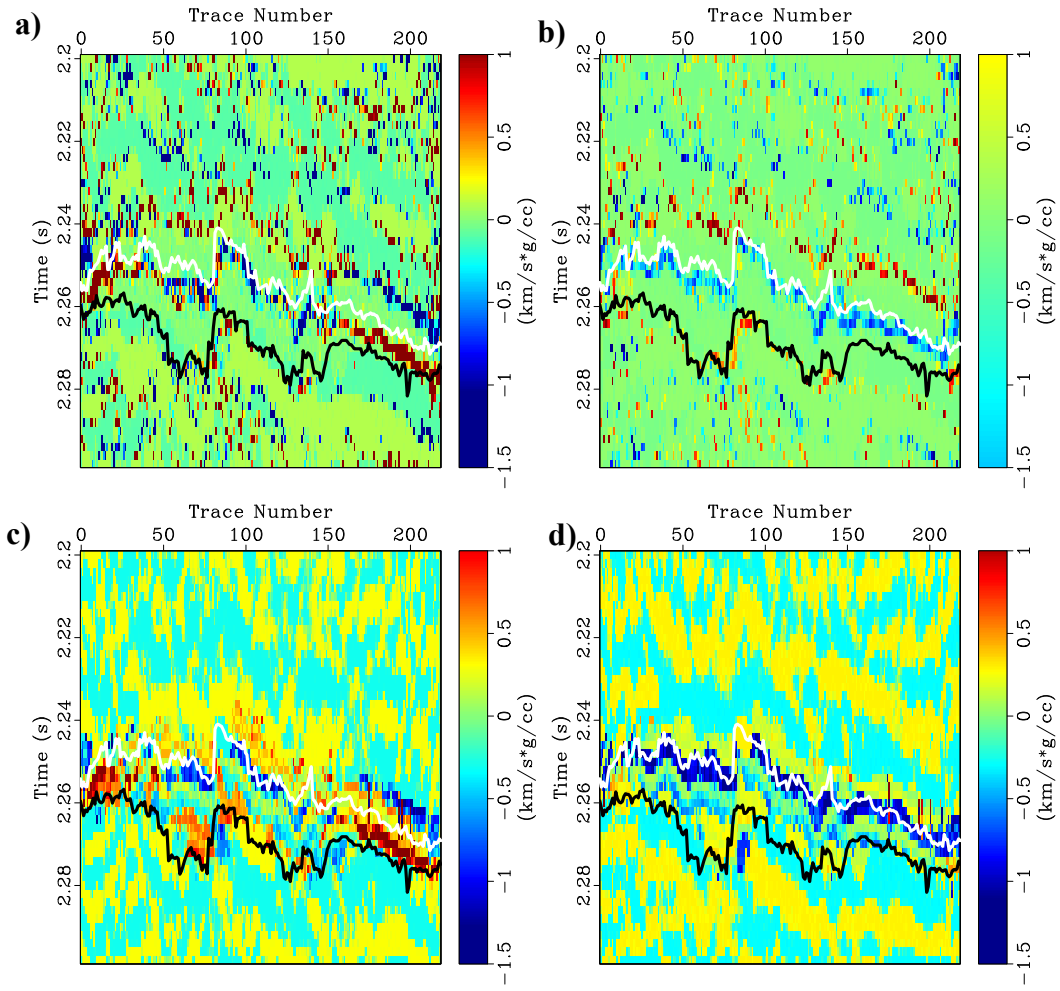
Figure 5.13: Comparison of inverted time-lapse differences for four different scenarios. (a) Conventional inversion of the data without warping; (b) double-difference inversion of the data without warping; (c) conventional inversion of the data with warping; (d) double-difference inversion of the data with warping.

Chapter 6: Conclusions and future directions

6.1 SUMMARY

I have presented novel applications of time reversal and plane-wave decompositions in seismic interferometry, full waveform inversion and reverse time migration. These three methods are currently the primary research focus areas in seismic exploration. The methods developed in this dissertation can have several applications, particularly for full wave applications in exploration geophysics. I have also shown an example of using prior information for robust time-lapse seismic inversion.

With the goal of utilizing the advantages of plane-wave transformation such as data reduction, directionality and focusing of direct wave, I developed a plane-wave based seismic interferometry method in chapter 2. I proved that the full redatumed seismic response can be retrieved in the plane-wave domain using a selected range of ray parameters. I tested the effectiveness of this method with transmission to reflection data retrieval, and VSP data for salt flank imaging. Results showed that with full ray parameters this approach is equivalent to conventional time-space domain approach but usually it can reduce the computational cost. The directionality with a choice of different ray parameters, which corresponds to the take-off angle at the surface, can also be used for better illumination of seismic interferometry. I also showed that this approach can be used to suppress artifacts of the CPG gather in super virtual interferometry and used this approach to a real OBS dataset. Far offset refractions have been significantly improved with this approach.

In chapter 3, I applied the plane-wave modeling approach to frequency domain full waveform inversion. This modeling strategy can be used both for the gradient and for the Hessian - the two key components needed for FWI. This is implemented as a multi-scale approach, which inverts discrete frequencies sequentially from low frequencies to

high frequencies. The inversion is carried out by a Gauss-Newton framework with a diagonal Hessian that is used to compensate for the source and receiver illuminations. Plane-wave encoding considerably reduces the computational burden and cross-talk artifacts are effectively suppressed by stacking over different ray parameters. Using a 2-D Overthrust example, I show that comparable results are obtained with my approach and with conventional shot-domain approach.

In chapter 4, I use the idea from the gradient calculation of FWI for time-domain reverse time migration. This is possible because numerical implementation of the gradient of the cost function in FWI is very similar as in RTM. L-p norms for FWI have been reported for better inversion result when the data contain outliers. RTM does not solve any data minimization problem, so "norm" in RTM does not make sense. However, these norms can be considered for scaling of the back-propagated receiver wavefield before applying the imaging condition. I tested this approach with different scaling strategies for non-Gaussian type noises. I found that in general scaling by its absolute norm (L-1 norm) produces better results than other approaches.

In chapter 5, I proposed a robust stochastic time-lapse seismic inversion strategy with an application of monitoring Cranfield CO_2 injection site. The key idea in this chapter is to use prior information to constrain the inversion process. This idea can also be used for other inversion algorithms, not only for the simultaneous inversion problem in this chapter. This inversion consists of two key steps with two types of prior information. The first step is the baseline inversion, which uses the fractal based hybrid prior from the well data to estimate the null space. The second step is to use the prior from the inverted parameters from the first step. The second step is known as double-difference inversion. With synthetic data, I showed that double-difference inversion obtains better result than a conventional two-pass approach. For the field data from

Cranfield site, the inverted results show time-lapse impedance changes that are consistent with CO_2 injection effects.

6.2 FUTURE RESEARCH

The interferometry relation used in this dissertation is based on the cross-correlation approach. Interferometry with a deconvolution approach and with plane-wave transformed data can be used to extend the scope of this dissertation. Other future research in plane-wave based seismic interferometry can be other applications where tau-p transform is applicable.

For plane-wave full waveform inversion and for reverse time migration, one direct application is to apply this approach to a real dataset and extend it to 3-D. For real data, we should use the "physical" time instead of the delay time to match the simulated plane-wave data. Incorporating elastic, attenuation and anisotropic effects can be interesting and challenging research topics as well. Various regularization, parameterization and optimization approaches can be incorporated for better performance of FWI for particular applications.

Appendix A: Proof of seismic interferometry using reflectivity method
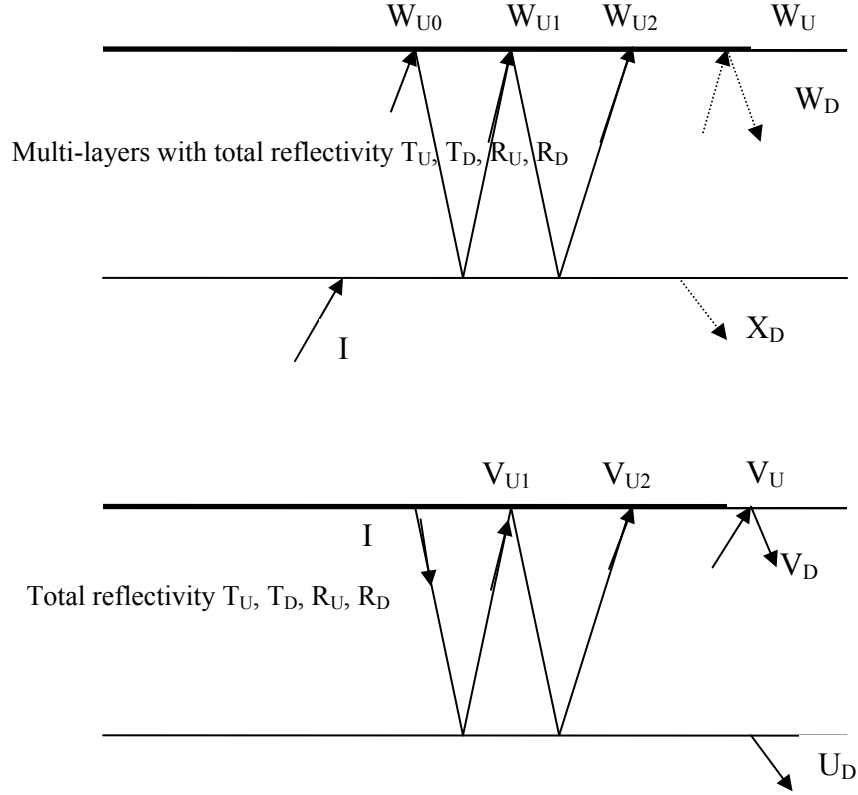
Figure A.1: The model to transform free surface recorded transmission responses to reflection responses.

Consider a model (Figure A.1) consisting of a free surface and a layer with incident plane waves P or SV from the deep. Received up-going wave $\mathbf{W}_U$ at the free surface consists of the $\mathbf{W}_{U0}$ (direct wave), $\mathbf{W}_{U1}$ (first reflection), and $\mathbf{W}_{U2}$ (second reflection). Kennett (1983) derived the following formula.

$$\mathbf{W}_U = \mathbf{W}_{U0} + \mathbf{W}_{U1} + \mathbf{W}_{U2} + \ldots = \mathbf{T}_U \mathbf{I} + \mathbf{R}_D \mathbf{R}_F \mathbf{T}_U + \mathbf{R}_D \mathbf{R}_F \mathbf{R}_D \mathbf{R}_F \mathbf{T}_U + \ldots$$
$$= \left[\mathbf{I} - \mathbf{R}_D \mathbf{R}_F \right]^{-1} \mathbf{T}_U$$

(A.1)

where $\mathbf{R}_F$ is the surface reflection matrix. $\mathbf{T}_U$, $\mathbf{T}_D$, $\mathbf{R}_U$ and $\mathbf{R}_D$ are the generalized transmission and reflection response from $z = 0$ to $z = z_N$. U means up going waves, D means down going waves. The term $[\mathbf{I} - \mathbf{R}_D \mathbf{R}_F]^{-1}$ includes all the multiples propagated in the region bounded by the free surface and the boundary of the last layer.

$$
\begin{aligned}
\mathbf{V}_U = \mathbf{V}_{U1} + \mathbf{V}_{U2} + \mathbf{V}_{U3} + ... &= \mathbf{R}_D + \mathbf{R}_D \mathbf{R}_F \mathbf{R}_D + \mathbf{R}_D \mathbf{R}_F \mathbf{R}_D \mathbf{R}_F \mathbf{R}_D + ... \\
&= (\mathbf{I} - \mathbf{R}_D \mathbf{R}_F)^{-1} \mathbf{R}_D
\end{aligned}
\tag{A.2}
$$

$$
\begin{aligned}
\mathbf{V}_D &= \mathbf{I} + \mathbf{R}_F \mathbf{R}_D + \mathbf{R}_F \mathbf{R}_D \mathbf{R}_F \mathbf{R}_D + \mathbf{R}_F \mathbf{R}_D \mathbf{R}_F \mathbf{R}_D \mathbf{R}_F \mathbf{R}_D + ... \\
&= \mathbf{I} + \mathbf{R}_F (\mathbf{R}_D + \mathbf{R}_D \mathbf{R}_F \mathbf{R}_D + \mathbf{R}_D \mathbf{R}_F \mathbf{R}_D \mathbf{R}_F \mathbf{R}_D + ...) \\
&= \mathbf{I} + \mathbf{R}_F (\mathbf{I} - \mathbf{R}_D \mathbf{R}_F)^{-1} \mathbf{R}_D
\end{aligned}
\tag{A.3}
$$

$$
\begin{aligned}
\mathbf{U}_D = \mathbf{T}_D \mathbf{V}_D &= \mathbf{T}_D (\mathbf{I} + \mathbf{R}_F (\mathbf{I} - \mathbf{R}_D \mathbf{R}_F)^{-1} \mathbf{R}_D) \\
&= \mathbf{T}_D (\mathbf{I} - \mathbf{R}_F \mathbf{R}_D)^{-1} = ((\mathbf{I} - \mathbf{R}_F{}^T \mathbf{R}_D{}^T)^{-1} \mathbf{T}_D{}^T)^T = \mathbf{W}_U{}^T
\end{aligned}
\tag{A.4}
$$

For the elastic case or P-SV wave modes case, the transmission matrix can be written as

$$
\mathbf{T}_D(k) =
\begin{bmatrix}
\overset{\downarrow\;\downarrow}{P\,P}\left(\dfrac{\rho_k V_P(k)\cos(\theta_k)}{\rho_{k-1} V_P(k-1)\cos(\theta_{k-1})}\right)^{1/2} & \overset{\downarrow\;\downarrow}{S\,P}\left(\dfrac{\rho_k V_P(k)\cos(\theta_k)}{\rho_{k-1} V_S(k-1)\cos(\phi_{k-1})}\right)^{1/2} \\[2em]
\overset{\downarrow\;\downarrow}{P\,S}\left(\dfrac{\rho_k V_S(k)\cos(\phi_k)}{\rho_{k-1} V_{P-1}(k)\cos(\theta_{k-1})}\right)^{1/2} & \overset{\downarrow\;\downarrow}{S\,S}\left(\dfrac{\rho_k V_S(k)\cos(\phi_k)}{\rho_{k-1} V_S(k-1)\cos(\phi_{k-1})}\right)^{1/2}
\end{bmatrix}
\tag{A.5}
$$

$$
\mathbf{T}_U(k) =
\begin{bmatrix}
\overset{\uparrow\;\uparrow}{P\,P}\left(\dfrac{\rho_{k-1} V_P(k-1)\cos(\theta_{k-1})}{\rho_k V_P(k)\cos(\theta_k)}\right)^{1/2} & \overset{\uparrow\;\uparrow}{S\,P}\left(\dfrac{\rho_{k-1} V_P(k-1)\cos(\theta_{k-1})}{\rho_k V_S(k)\cos(\phi_k)}\right)^{1/2} \\[2em]
\overset{\uparrow\;\uparrow}{P\,S}\left(\dfrac{\rho_{k-1} V_S(k-1)\cos(\phi_{k-1})}{\rho_k V_P(k)\cos(\theta_k)}\right)^{1/2} & \overset{\uparrow\;\uparrow}{S\,S}\left(\dfrac{\rho_{k-1} V_S(k-1)\cos(\phi_{k-1})}{\rho_k V_S(k)\cos(\phi_k)}\right)^{1/2}
\end{bmatrix}
\tag{A.6}
$$

where:
$$
\begin{pmatrix}
\overset{\downarrow\;\uparrow}{P\,P} & \overset{\downarrow\;\uparrow}{S\,P} & \overset{\uparrow\;\uparrow}{P\,P} & \overset{\uparrow\;\uparrow}{S\,P} \\
\overset{\downarrow\;\uparrow}{P\,S} & \overset{\downarrow\;\uparrow}{S\,S} & \overset{\uparrow\;\uparrow}{P\,S} & \overset{\uparrow\;\uparrow}{S\,S} \\
\overset{\downarrow\;\downarrow}{P\,P} & \overset{\downarrow\;\downarrow}{S\,P} & \overset{\uparrow\;\downarrow}{P\,P} & \overset{\uparrow\;\downarrow}{S\,P} \\
\overset{\downarrow\;\downarrow}{P\,S} & \overset{\downarrow\;\downarrow}{S\,S} & \overset{\uparrow\;\downarrow}{P\,S} & \overset{\uparrow\;\downarrow}{S\,S}
\end{pmatrix} = \mathbf{M}^{-1} \mathbf{N}
\tag{A.7}
$$

with

$$\overset{\downarrow}{P}\overset{\downarrow}{P} = \frac{2\rho_{k-1}\cos(\theta_{k-1})}{V_P(k)}\frac{F}{D} \tag{A.7.1}$$

$$\overset{\downarrow}{P}\overset{\downarrow}{S} = \frac{2\rho_{k-1}\cos(\theta_{k-1})\sin(\phi_k)}{V_S(k)V_S(k)}\frac{H}{D} \tag{A.7.2}$$

$$\overset{\downarrow}{S}\overset{\downarrow}{P} = -\frac{2\rho_{k-1}\cos(\phi_{k-1})\sin(\theta_k)}{V_P(k)V_P(k)}\frac{G}{D} \tag{A.7.3}$$

$$\overset{\downarrow}{S}\overset{\downarrow}{S} = \frac{2\rho_{k-1}\cos(\phi_{k-1})}{V_S(k)}\frac{E}{D} \tag{A.7.4}$$

$$\overset{\uparrow}{P}\overset{\uparrow}{P} = \frac{2\rho_k\cos(\theta_k)}{V_P(k-1)}\frac{F}{D} \tag{A.7.5}$$

$$\overset{\uparrow}{P}\overset{\uparrow}{S} = -\frac{2\rho_k\cos(\theta_k)\sin(\theta_{k-1})}{V_S(k-1)V_S(k-1)}\frac{G}{D} \tag{A.7.6}$$

$$\overset{\uparrow}{S}\overset{\uparrow}{P} = \frac{2\rho_k\cos(\phi_k)\sin(\theta_{k-1})}{V_P(k-1)V_P(k-1)}\frac{H}{D} \tag{A.7.8}$$

$$\overset{\uparrow}{S}\overset{\uparrow}{S} = \frac{2\rho_k\cos(\phi_k)}{V_S(k-1)}\frac{E}{D} \tag{A.7.9}$$

where D, E, F, G, H are cosine-dependent terms (Aki and Richards, 2002). Using the above equation, the transmission matrix can be explicitly written as

$$\mathbf{T}_D(k) = \begin{bmatrix} \frac{2F}{D}\sqrt{\frac{\rho_{k-1}\rho_k}{V_P(k-1)V_P(k)}}\cos(\theta_{k-1})\cos(\theta_k) & -2p\frac{G}{D}\sqrt{\frac{\rho_{k-1}\rho_k}{V_P(k)V_S(k-1)}}\cos(\phi_{k-1})\cos(\theta_k) \\ \frac{2pH}{D}\sqrt{\frac{\rho_{k-1}\rho_k}{V_{P-1}(k)V_S(k)}}\cos(\phi_k)\cos(\theta_{k-1}) & \frac{2E}{D}\sqrt{\frac{\rho_{k-1}\rho_k}{V_S(k-1)V_S(k)}}\cos(\phi_{k-1})\cos(\phi_k) \end{bmatrix} \tag{A.8}$$

$$\mathbf{T}_U(k) = \begin{bmatrix} \frac{2F}{D}\sqrt{\frac{\rho_{k-1}\rho_k}{V_P(k-1)V_P(k)}}\cos(\theta_{k-1})\cos(\theta_k) & \frac{2pH}{D}\sqrt{\frac{\rho_{k-1}\rho_k}{V_{P-1}(k)V_S(k)}}\cos(\phi_k)\cos(\theta_{k-1}) \\ -2p\frac{G}{D}\sqrt{\frac{\rho_{k-1}\rho_k}{V_P(k)V_S(k-1)}}\cos(\phi_{k-1})\cos(\theta_k) & \frac{2E}{D}\sqrt{\frac{\rho_{k-1}\rho_k}{V_S(k-1)V_S(k)}}\cos(\phi_{k-1})\cos(\phi_k) \end{bmatrix} \tag{A.9}$$

where p is the ray parameter. Equations A.8 and A.9 suggest that $\mathbf{T}_D{}^T = \mathbf{T}_U$.

Similarly, for the reflection matrix, it can be expressively written as

$$\mathbf{R}_F = \begin{bmatrix} \dfrac{\beta^2 \sin 2i_p \sin 2r_s - \alpha^2 \cos^2 2r_s}{\beta^2 \sin 2i_p \sin 2r_s + \alpha^2 \cos^2 2r_s} & \dfrac{-2\alpha\beta \sin 2i_p \cos 2r_s}{\beta^2 \sin 2i_p \sin 2r_s + \alpha^2 \cos^2 2r_s} \dfrac{\sqrt{\alpha\cos i_p}}{\sqrt{\beta\cos r_s}} \\[4mm] \dfrac{2\alpha\beta \sin 2i_p \cos 2r_s}{\beta^2 \sin 2i_p \sin 2r_s + \alpha^2 \cos^2 2r_s} \dfrac{\sqrt{\beta\cos r_s}}{\sqrt{\alpha\cos i_p}} & \dfrac{\beta^2 \sin 2i_p \sin 2r_s - \alpha^2 \cos^2 2r_s}{\beta^2 \sin 2i_p \sin 2r_s + \alpha^2 \cos^2 2r_s} \end{bmatrix} \tag{A.10}$$

After matrix multiplication, we could get

$$\mathbf{R}_F{}^* \mathbf{R}_F = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} \tag{A.11}$$

For lossless media case, the energy flux of the seismic wave before and after propagating in the region should be equal. Then we have the following identity:

$$\begin{aligned} \mathbf{U}_D^* \mathbf{U}_D &= \mathbf{V}_D^* \mathbf{V}_D - \mathbf{V}_U^* \mathbf{V}_U \\ &= (\mathbf{I} + \mathbf{R}_F \mathbf{V}_U + \mathbf{V}_U^* \mathbf{R}_F^* + \mathbf{V}_U^* \mathbf{R}_F^* \mathbf{R}_F \mathbf{V}_U) - \mathbf{V}_U^* \mathbf{V}_U = \mathbf{I} + \mathbf{R}_F \mathbf{V}_U + \mathbf{V}_U^* \mathbf{R}_F^* \end{aligned} \tag{A.12}$$

where $*$ is the conjugate transpose. We can also get this identity

$$\mathbf{W}_D^* \mathbf{W}_D = \mathbf{I} + \mathbf{R}_F \mathbf{V}_U + \mathbf{V}_U^* \mathbf{R}_F^* \tag{A.13}$$

Equation A.12 or A.13 is the interferometric relation, which suggests that the autocorrelation of the transmitted wave equals the waves received at the surface plus its time-reversed version plus a delta function.

Appendix B: Proof of the equivalence of shot-profile and plane-wave gradient

Incorporating the plane-wave encoding strategy, the forward simulation of the source wavefield is the solution to equations 3.4 and 3.5. It can be written as

$$\tilde{S}(\mathbf{x},\mathbf{p},\omega) = \begin{cases} \sum_{\mathbf{x_s}} S(\mathbf{x},\mathbf{x_s},\omega)\exp(i\omega\mathbf{p}(\mathbf{x_s}-\mathbf{x}_0)) & p_i \geq 0 \\ \sum_{\mathbf{x_s}} S(\mathbf{x},\mathbf{x_s},\omega)\exp(i\omega\mathbf{p}(\mathbf{x_s}-\mathbf{x}_{\max})) & p_i < 0, \end{cases} \tag{B-1}$$

where $S(\mathbf{x},\mathbf{x_s},\omega)$ is the source wavefield for the shot location at $(\mathbf{x_s},0)$ simulated with the shot-profile wave equation; $\tilde{S}(\mathbf{x},\mathbf{p},\omega)$ is the source wavefield with the plane-wave simulation.

Similarly, the back propagated data residual can be expressed as

$$\Delta\tilde{R}(\mathbf{x},\mathbf{p},\omega) = \begin{cases} \sum_{\mathbf{x_s}} \Delta R(\mathbf{x},\mathbf{x_s},\omega)\exp(i\omega\mathbf{p}(\mathbf{x_s}-\mathbf{x}_0)) & p_i \geq 0 \\ \sum_{\mathbf{x_s}} \Delta R(\mathbf{x},\mathbf{x_s},\omega)\exp(i\omega\mathbf{p}(\mathbf{x_s}-\mathbf{x}_{\max})) & p_i < 0. \end{cases} \tag{B-2}$$

After crosscorrelating the source wavefield and residual receiver wavefield, the gradient, with the ray parameter being $\mathbf{p}$, can be written as follows

$$\begin{aligned}
\tilde{g}(\mathbf{x},\mathbf{p_s}) &= \sum_{\omega}\mathrm{Re}\left\{\frac{-2\omega^2|\omega|f(\omega)}{v^3(\mathbf{x})}\tilde{S}(\mathbf{x},\mathbf{p_s},\omega)\Delta\tilde{R}^*(\mathbf{x},\mathbf{p_s},\omega)\right\} \\
&= \sum_{\omega}\sum_{\mathbf{x}_1}\sum_{\mathbf{x}_2}\frac{-2\omega^2|\omega|f(\omega)}{v^3(\mathbf{x})}S(\mathbf{x},\mathbf{x}_1,\omega)\Delta R^*(\mathbf{x},\mathbf{x}_2,\omega)\exp(i\omega\mathbf{p_s}(\mathbf{x}_1-\mathbf{x}_2)),
\end{aligned} \tag{B-3}$$

where $\mathbf{x}_1$ and $\mathbf{x}_2$ represent different shot locations for the source wavefiled and residual receiver wavefield.

Note that

$$\sum_{p=-\infty}^{\infty}\exp(i\omega p(\mathbf{x}_1-\mathbf{x}_2)) = \frac{1}{|\omega|}\delta(\mathbf{x}_1-\mathbf{x}_2). \tag{B-4}$$

110

Stacking equation A-3 over all ray parameters results in the plane-wave domain gradient

$$\tilde{g}(\mathbf{x}) = \sum_{\omega}\sum_{\mathbf{x_1}}\sum_{\mathbf{x_2}}\sum_{|\mathbf{p_s}|=-\infty}^{\infty} \frac{-2\omega^2|\omega|f(\omega)}{v^3(\mathbf{x})}S(\mathbf{x},\mathbf{x}_1,\omega)\Delta R^*(\mathbf{x},\mathbf{x}_2,\omega)\exp(i\omega\mathbf{p_s}(\mathbf{x}_1-\mathbf{x}_2))$$

$$= \sum_{\omega}\sum_{\mathbf{x_1}}\sum_{\mathbf{x_2}} \frac{-2\omega^2 f(\omega)}{v^3(\mathbf{x})}S(\mathbf{x},\mathbf{x}_1,\omega)\Delta R^*(\mathbf{x},\mathbf{x}_2,\omega)\delta(\mathbf{x}_1-\mathbf{x}_2) \qquad \text{(B-5)}$$

$$= \sum_{\omega}\sum_{\mathbf{x_s}} \frac{-2\omega^2 f(\omega)}{v^3(\mathbf{x})}S(\mathbf{x},\mathbf{x}_s,\omega)\Delta R^*(\mathbf{x},\mathbf{x}_s,\omega).$$

Thus, we have proved that

$$\tilde{g}(\mathbf{x}) = g(\mathbf{x}). \qquad \text{(B-6)}$$

However, real application of the gradient calculation is to simultaneously back-propagate the residual wavefield at all the receiver locations. This will result in some different filtering effects between plane-wave encoded gradient and shot-profile gradient.

Appendix C: Proof of the equivalence of shot-profile and plane-wave encoded diagonal Hessian

For writing simplicity, the proof shown below is based on the diagonal Hessian. However, it does not involve any modification to extend to the full Hessian matrix.

Similar to the relationship between the shot-profile and plane-wave extrapolated wavefield in the gradient calculation, the Green's function satisfies

$$\tilde{G}(\mathbf{x},\mathbf{p},\omega) = \begin{cases} \sum_{\mathbf{x_s}}G(\mathbf{x},\mathbf{x}_s,\omega)\exp(i\omega\mathbf{p}(\mathbf{x}_s-\mathbf{x}_0)) & p_i \geq 0 \\ \sum_{\mathbf{x_s}}G(\mathbf{x},\mathbf{x}_s,\omega)\exp(i\omega\mathbf{p}(\mathbf{x}_s-\mathbf{x}_{max})) & p_i < 0. \end{cases} \qquad \text{(C-1)}$$

The receiver-side plane-wave encoded Hessian is expressed as

$$\tilde{H}_0(\mathbf{x}, \mathbf{p_r}) = \sum_\omega \sum_{\mathbf{x_s}} \mathrm{Re}\left\{ \frac{4\omega^4 |\omega| |f(\omega)|^2}{v^6(\mathbf{x})} |G(\mathbf{x}, \mathbf{x_s}, \omega)|^2 |\tilde{G}(\mathbf{x}, \mathbf{p_r}, \omega)|^2 \right\}$$

$$= \sum_\omega \sum_{\mathbf{x_s}} \mathrm{Re}\left\{ \frac{4\omega^4 |\omega| |f(\omega)|^2}{v^6(\mathbf{x})} |G(\mathbf{x}, \mathbf{x_s}, \omega)|^2 \tilde{K}_\mathbf{r}(\mathbf{x}, \mathbf{p_r}, \omega) \right\}, \qquad \text{(C-2)}$$

with

$$\tilde{K}_\mathbf{r}(\mathbf{x}, \mathbf{p_r}, \omega) = \sum_{\mathbf{x_1}} \sum_{\mathbf{x_2}} G(\mathbf{x}, \mathbf{x_1}, \omega) G^*(\mathbf{x}, \mathbf{x_2}, \omega) \exp(i\omega \mathbf{p_r}(\mathbf{x_1} - \mathbf{x_2})). \qquad \text{(C-3)}$$

Integrating the equation C-3 over ray parameters, we get

$$\tilde{K}_\mathbf{r}(\mathbf{x}, \omega) = \sum_{\mathbf{x_1}} \sum_{\mathbf{x_2}} \sum_{\mathbf{p_r}} G(\mathbf{x}, \mathbf{x_1}, \omega) G^*(\mathbf{x}, \mathbf{x_2}, \omega) \exp(i\omega \mathbf{p_r}(\mathbf{x_1} - \mathbf{x_2}))$$

$$= \sum_{\mathbf{x_1}} \sum_{\mathbf{x_2}} G(\mathbf{x}, \mathbf{x_1}, \omega) G^*(\mathbf{x}, \mathbf{x_2}, \omega) \frac{1}{|\omega|} \delta(\mathbf{x_1} - \mathbf{x_2}) \qquad \text{(C-4)}$$

$$= \sum_{\mathbf{x_r}} \frac{1}{|\omega|} |G(\mathbf{x}, \mathbf{x_r}, \omega)|^2,$$

$$\tilde{H}_0(\mathbf{x}) = \sum_\omega \sum_{\mathbf{x_s}} \sum_{\mathbf{x_r}} \mathrm{Re}\left\{ \frac{4\omega^4 |f(\omega)|^2}{v^6(\mathbf{x})} |G(\mathbf{x}, \mathbf{x_s}, \omega)|^2 |\tilde{G}(\mathbf{x}, \mathbf{x_r}, \omega)|^2 \right\}$$

$$= H_0(\mathbf{x}). \qquad \text{(C-5)}$$

Equation C-4 shows the equivalence of the receiver-side plane-wave encoded Hessian and shot-profile Hessian.

A simultaneous plane-wave encoding uses both source and receiver ray parameters. It can be formulated as

$$\tilde{\tilde{H}}_0(\mathbf{x}, \mathbf{p_s}, \mathbf{p_r}) = \sum_\omega \mathrm{Re}\left\{ \frac{4\omega^4 |\omega f(\omega)|^2}{v^6(\mathbf{x})} |\tilde{G}(\mathbf{x}, \mathbf{p_s}, \omega)|^2 |\tilde{G}(\mathbf{x}, \mathbf{p_r}, \omega)|^2 \right\}$$

$$= \sum_\omega \mathrm{Re}\left\{ \frac{4\omega^4 |\omega f(\omega)|^2}{v^6(\mathbf{x})} |\tilde{K}_\mathbf{s}(\mathbf{x}, \mathbf{p_s}, \omega)|^2 |\tilde{K}_\mathbf{r}(\mathbf{x}, \mathbf{p_r}, \omega)|^2 \right\}, \qquad \text{(C-6)}$$

with $\tilde{K}_\mathbf{r}(\mathbf{x}, \mathbf{p_r}, \omega)$ given in equation C-3 and $\tilde{K}_\mathbf{s}(\mathbf{x}, \mathbf{p_s}, \omega)$ given as follows

$$\tilde{K}_{\mathrm{s}}(\mathbf{x},\mathbf{p}_{\mathrm{s}},\omega) = \sum_{\mathbf{x}_3}\sum_{\mathbf{x}_4} G(\mathbf{x},\mathbf{x}_3,\omega)G^*(\mathbf{x},\mathbf{x}_4,\omega)\exp(i\omega\mathbf{p}_{\mathrm{s}}(\mathbf{x}_3-\mathbf{x}_4)),\qquad\text{(C-7)}$$

The integration relationship in C-3 also holds for $\tilde{K}_{\mathrm{s}}(\mathbf{x},\mathbf{p}_{\mathrm{s}},\omega)$, i.e.

$$\tilde{K}_{\mathrm{s}}(\mathbf{x},\omega) = \sum_{\mathbf{x}_{\mathrm{s}}}\frac{1}{|\omega|}|G(\mathbf{x},\mathbf{x}_{\mathrm{s}},\omega)|^2 .\qquad\text{(C-8)}$$

Thus

$$\tilde{\tilde{H}}_0(\mathbf{x}) = \sum_{\omega}\sum_{\mathbf{x}_{\mathrm{s}}}\sum_{\mathbf{x}_{\mathrm{r}}}\mathrm{Re}\left\{\frac{4\omega^4|f(\omega)|^2}{v^6(\mathbf{x})}|G(\mathbf{x},\mathbf{x}_{\mathrm{s}},\omega)|^2|\tilde{G}(\mathbf{x},\mathbf{x}_{\mathrm{r}},\omega)|^2\right\}\qquad\text{(C-9)}$$
$$= H_0(\mathbf{x}).$$

Equation C-9 proves the equivalence of the source and receiver plane-wave encoded Hessian and shot-profile Hessian.

References

Alford R. M., K. R. Kelly, D. M. Boore, 1974, Accuracy of finite-difference modeling of the acoustic wave equation, Geophysics, 39, no. 6, 834-842.

Aki K. and P. G. Richards, 2002, Quantitative seismology, 2nd edition, University Science Books.

Amestoy, P. R., A. Guermouche, J. Y. L'Excellent, and S. Pralet, 2006, Hybrid scheduling for the parallel solution of linear systems: Parallel computing, 32, 136–156.

Aravkin, A., T. Leeuwen and F. Herrmann, 2011, Robust full-waveform inversion using the Student's t-distribution: 81st Annual International Meeting, SEG, Expanded Abstracts, 2669-2673.

Bakulin, A. and R. Calvert, 2006, The virtual source method: Theory and case study: Geophysics, 71, no. 4, SI139-SI150.

Balay, S., K. Buschelman, W. D. Gropp, D. Kaushik, L. C. McInnes, and B. F. Smith, , 2001, PETSc home page. http://www.mcs.anl.gov/petsc.

Boonyasiriwat, C., P. Valasek, P. Routh, W. Cao, G. T. Schuster, and B. Macy, 2009, An efficient multiscale method for time-domain waveform tomography: Geophysics, 74, no. 6, WCC59-WCC68.

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

Ben-Hadj-Ali, H., S. Operto and J. Virieux, 2011, An efficient frequency-domain full-waveform inversion method using simultaneous encoded sources: Geophysics, 76, no. 4, R109–R124.

Bharadwaj, P., G. T. Schuster, I. Mallison, and W. Dai, 2012, Theory of supervirtual refraction interferometry: Geophysical Journal International, 188, 263-273.

Browaeys T. J. and S. Fomel, 2009, Fractal heterogeneities in sonic logs and low-frequency scattering attenuation: Geophysics, 74, no. 2, WA77-WA92.

Brossier, R., Operto, S. and Virieux, J., 2009, Seismic imaging of complex onshore structures by 2D elastic frequency-domain full-waveform inversion, Geophysics, 74, no. 6, WCC105–WCC118.

Brossier, R., Operto, S. and Virieux, J., 2010, Which residual norm for robust elastic frequency-domain full waveform inversion?: Geophysics, 75, no. 3, R37–R46.

Bube, K. P., and R. T. Langan, 1997, Hybrid l1 / l2 minimization with applications to tomography: Geophysics, 62, 1183–1195.

Calvert, R., 2005, 4D technology: where are we, and where are we going?: Geophysical Prospecting, 53, 161–171.

Capdeville, Y., Y. Gung, and B. Romanowicz, 2005, Towards global earth tomography using the spectral element method: A technique based on source stacking: Geophysical Journal International, 162, 541–554.

Červený, V., 2001, Seismic ray theory: Cambridge University Press.

Claerbout, J. F., 1968, Synthesis of a layered medium from its acoustic transmission response, Geophysics, 33, 264-269.

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

Claerbout, J. F., 1985, Imaging the earth's interior: Blackwell Science Inc.

Claerbout, J., 1992, Earth soundings analysis: Processing vs inversion: Blackwell Scientific Publications.

Curtis A., P. Gerstoft, H. Sato, R. Snieder and K. Wapenaar, 2006, Seismic interferometry - turning noise into signal, The Leading Edge, 25, 1082-1092.

Curtis, A., 2009, Source-receiver seismic interferometry: 79th Annual International Meeting, SEG, Expanded Abstracts, 3655–3658.

De Basabe, J. D. and M. K. Sen, 2007, Grid dispersion and stability criteria of some common finite-element methods for acoustic and elastic wave equations, Geophysics, 72, no. 6, T81-T95.

De Basabe, J. D., M. K. Sen and M. F. Wheeler, 2008, The Interior Penalty Discontinuous Galerkin method for elastic wave propagation: grid dispersion, Geophysical Journal International, 175, 83-93.

Denli, H., and L. Huang, 2009, Double-difference elastic waveform tomography in the time domain: 79th Annual International Meeting, SEG, Expanded Abstracts, 2302–2306.

Diebold, J. B., and P. L. Stoffa, 1981, The traveltime equation, tau-p mapping and inversion of common midpoint data: Geophysics, 46, 238–254.

Druzhinin, A., and C. MacBeth, 2001, Robust cross-equalization of 4D-4C PZ migrated data at Teal South: 71st Annual International Meeting, SEG, Expanded Abstracts, 1670–1673.

Draganov, D., K. Wapenaar, and J. Thorbecke, 2006, Seismic interferometry: Reconstructing the Earth's reflection response: Geophysics, 71, no. 4, SI61– SI70.

Draganov, D., K.Wapenaar,W. Mulder, J. Singer, and A. Verdel, 2007, Retrieval of reflections from seismic background-noise measurements: Geophysical Research Letters, 34, L04305-1–L04305-4.

115

Dong, S., J. Sheng, and G. T. Schuster, 2006, Theory and practice of refraction interferometry: SEG Expanded Abstracts, 25, 3021–3025.

Dormy, E. and A. Tarantola, 1995, Numerical simulation of elastic wave propagation using a finite volume method, Journal of Geophysical Research, 100(B2), 2123-2134.

Fatti, J. L., G. C. Smith, P. J. Vail, P. J. Strauss, and P. R. Levitt, 1994, Detection of gas in sandstone reservoirs using AVO analysis: Geophysics, 59, 1362–1376.

Fink M., 1997, Time reversed acoustics: Physics Today, 50, 34-40.

Fletcher, R. F., P. Fowler, P. Kitchenside, and U. Albertin, 2005, Suppressing artifacts in prestack reverse-time migration: 75th Annual International Meeting, SEG, Expanded Abstracts, 2049–2051.

Fomel, S., 2007, Local seismic attributes: Geophysics, 72, no. 3, A29–A33.

Fomel, S., 2009, Velocity analysis using AB semblance: Geophysical Prospecting, 57, 311–321.

Fomel, S., and L. Jin, 2009, Time-lapse image registration using the local similarity attribute: Geophysics, 74, no. 2, A7–A11.

Foster, D. J., and C. C. Mosher, 1992, Suppression of multiple reflections using the Radon transform: Geophysics, 57, 386–395.

Frasier C. W., 1970, Discrete time solution of plane P-SV waves in a plane layered medium, Geophysics, 35, 197-219.

Grand S. P., 1987, Tomographic inversion for shear velocity beneath the North American plate, J. Geophys. Res., 92, pp. 14,065-14,090.

Guitton, A., and W.W. Symes, 2003, Robust inversion of seismic data using the Huber norm: Geophysics, 68, 1310–1319.

Guitton, A., B. Kaelin, and B. Biondi, 2007, Least-square attenuation of reverse-time migration artifacts: Geophysics, 72, no. 1, S19–S23.

Ha, T., Chung, W., Shin, C., 2009, Waveform inversion using a back-propagation algorithm and a Huber function norm: Geophysics, 74, no. 3, R15-R24.

Hall, S. A., C. MacBeth, J. Stammeijer, and M. Omerod, 2006, Time-lapse seismic analysis of pressure depletion in the southern gas basin: Geophysical Prospecting, 54, 63-73.

Hampson, D. P., B. H. Russell, and B. Bankhead, 2005, Simultaneous inversion of pre-stack seismic data: 75th Annual International Meeting, SEG, Expanded Abstracts, 1633-1636.

Herrmann, F. J., Y. Erlangga, and T. T. Y. Lin, 2009, Compressive sensing applied to full-wave form inversion: 71st Annual International Conference & Exhibition, EAGE, Extended Abstracts, S016.

Hewett, T. A., 1986, Fractal distribution of reservoir heterogeneity and their influence on fluid transport: 61st Annual Technical Conference, SPE, paper 15386.

Hong, T., and M. K. Sen, 2009, A new MCMC algorithm for seismic waveform inversion and corresponding uncertainty analysis: Geophysical Journal International, 177, no. 1, 14–32.

Hovorka, S. D., T. A. Meckel, R. H. Trevino, J. Lu, J.-P. Nicot, J.-W. Choi, D. Freeman, P. Cook, T. M. Daley, J. B. Ajo-Franklin, B. M. Freifeild, C. Doughty, C. R. Carrigan, D. L. Brecque, Y. K. Kharaka, J. J. Thordsen, T. J. Phelps, C. Yang, K. D. Romanak, T. Zhang, R. M. Holt, J. S. Lindler, and R. J. Butsch, 2011, Monitoring a large volume CO_2 injection: Year two results from secarb project at denburys Cranfield, Mississippi, USA: Energy Procedia, 4, 3478-3485. (10th International Conference on Greenhouse Gas Control Technologies).

Huber, P. J., 1973, Robust regression: Asymptotics, conjectures, and Monte Carlo: The Annals of Statistics, 1, 799–821.

Hurst, H. E., R. P. Black, and Y. M. Simaika, 1965, Long-term storage: an experimental study, Constable, London.

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

Ivansson, S., 1985, A study of methods for tomographic velocity estimation in the presence of low-velocity zone: Geophysics, 50, 969-988.

Johnson, T. C., R. J.Versteeg, H. Huang, and P. S. Routh, 2009, Data-domain correlation approach for joint hydrogeologic inversion of time-lapse hydrogeologic and geophysical data: Geophysics, 74, no. 6, F127–F140.

Kennett, B. L. N., 1983, Seismic Wave Propagation in Stratified Media, Cambridge University Press.

Komatitsch, D. and J. Vilotte, 1998, The spectral-element method: an efficient tool to simulate the seismic response of 2D and 3D geological structures, Bulletin of the Seismological Society of America, 88, 368-392.

Koster, K., Gabriels, P., Hartung, M., Verbeek, J., Deinum, G. and Staples, R., 2000, Time-lapse seismic surveys in the North Sea and their business impact: The Leading Edge, 19, no. 03, 286-293.

Krebs, J., J. Anderson, D. Hinkley, R. Neelamani, S. Lee, A. Baumstein, and M. D. Lacasse, 2009, Fast full-wavefield seismic inversion using encoded sources: Geophysics, 74, no. 6, WCC177–WCC188.

Lamb, H., 1904, On the propagation of tremors at the surface of an elastic solid: Trans Roy Soc. London, Ser A, 203, 1-42.

Landrø, M., Strønen, L.K. and Digranes, P., 2001, Mapping reservoir pressure and saturation changes using seismic methods--possibilities and limitations, First Break 19, 671-677.

Lee, J., Y. Kim, and C. Shin, 2011, Frequency-domain reverse time migration using the L_1-norm：81st Annual International Meeting, SEG, Expanded Abstracts, 3169-3173.

Levander, A. R., 1988, Fourth-order finite-difference P-SV seismograms: Geophysics, 53, 1425-1436.

Liu, F., D. W. Hanson, N. D. Whitmore, R. S. Day, and R. H. Stolt, 2006, Toward a unified analysis for source plane-wave migration: Geophysics, 71, no. 4, S129–S139.

Liu, F., G. Zhang, S. A. Morton, and J. P. Leveille, 2011, An effective imaging condition for reverse-time migration using wavefield decomposition: Geophysics, 76, no. 1, S29–S39.

Liu, F., M. K. Sen, and P. L. Stoffa, 2000, Dip selective 2-D multiple attenuation in the plane-wave domain: Geophysics, 65, 264–274.

Liu, F., N. D. Whitmore, D.W. Hanson, R. S. Day, and C. C.Mosher, 2004, The impact of reciprocity on prestack source plane wave migration: 74th Annual International Meeting, SEG, Expanded Abstracts, 1045–1048.

Lu, R., M. E. Willis, X. Campman, J. Ajo-Franklin, and M. N. Toksöz, 2008, Redatuming through a salt canopy and target-oriented salt-flank imaging: Geophysics,73, no. 3, S63–S71.

Lumley, D., 2001, Time-lapse seismic reservoir monitoring: Geophysics, 66, 50–53.

Mallinson, I., P. Bharadwaj, G. Schuster, and H. Jakubowicz, 2011, Enhanced refractor imaging by super-virtual interferometry: The Leading Edge, 30, 546–550.

Marfurt, K., 1984, Accuracy of finite-difference and finite-elements modeling of the scalar and elastic wave equation: Geophysics, 49, 533–549.

McIntosh, K. D., Y. Nakamura, T.-K. Wang, R.-C. Shih, A. T. Chen, and C.-S. Liu, 2005, Crustal-scale seismic profiles across Taiwan and the western Philippine Sea: Tectonophysics, 401, 23-54.

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

Mehta, K., Bakulin A., Sheiman J., Calvert R., and Snieder R., 2007, Improving the virtual source method by wavefield separation: Geophysics, 72, no. 4, V79-V86.

Mehta, K., R. Snieder, R. Calvert, and J. Sheiman, 2008, Acquisition geometry requirements for generating virtual-source data: The Leading Edge, 27, 620-629.

Metropolis, N., A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, 1953, Equation of state calculations by fast computing machines: Journal of Chemical Physics, 21, 1087–1092.

Menke, W., 1984, Geophysical data analysis: Discrete inverse theory: Academic Press, Inc.

Mora, P., 1989, Inversion=migration + tomography: Geophysics, 54, 1575–1586.

Mulder, W., and R.-E. Plessix, 2004, A comparison between one-way and two-way wave equation migration: Geophysics, 69, 1491–1504.

Mulder, W. A., 2005, Rigorous redatuming, Geophysical Journal International, 161, 401–415.

Oldham, R., 1906, The constitution of the earth: Quarterly Journal of the Geological Society of London, 62, 456–475.

Operto, S., J. Virieux, P. Amestoy, J. Y. L'Excellent, L. Giraud, and H. Ben- Hadj-Ali, 2007, 3D finite-difference frequency-domain modeling of viscoacoustic wave propagation using a massively parallel direct solver: A feasibility study: Geophysics, 72, no. 5, SM195–SM211.

Ostrander W. J., 1984, Plane-wave reflection coefficients for gas sands at nonnormal angles of incidence: Geophysics, 49, 1637-1648.

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.

Plessix, R.-E., and W. A. Mulder, 2004, Frequency-domain finite-difference amplitude-preserving migration: Geophysical Journal International, 157, 975–987.

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

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

Pyun, S., C. Shin, and W. Son, 2009, Frequency-domain waveform inversion using an L1-norm objective function: Exploration Geophysics, 40, 227-232.

Rickett, J. E., and D. E. Lumley, 2001, Cross-equalization data processing for time-lapse seismic reservoir monitoring: A case study from the Gulf of Mexico: Geophysics, 66, 1015-1025.

Rickett, J., and J. Claerbout, 1999, Acoustic daylight imaging via spectral factorization: Helioseismology and reservoir monitoring: The Leading Edge, 18, 957–960.

Rivière, B., M. Wheeler and V. Girault, 1999, Improved energy estimates for interior penalty, constrained and discontinuous Galerkin methods for elliptic problems. Part I, Computational Geosciences, 3, 337-360.

Romero, L. A., D. C. Ghiglia, C. C. Ober, and S. A. Morton, 2000, Phase encoding of shot records in prestack migration: Geophysics, 65, 426–436.

Roth, M., and M. Korn,1993, Single scattering theory versus numerical modelling in 2-D random media: Geophysical Journal International, 112, 124–140.

Rowbotham, P. S., D. Marion, P. Lamy, E. Insalaco, P.A. Swaby, andY. Boisseau, 2003, Multidisciplinary stochastic impedance inversion: Integrating geological understanding and capturing reservoir uncertainty: Petroleum Geoscience, 9, 287–294.

Sabra, K., Gerstoft, P., Roux, P., Kuperman, W.A., Fehler, M.C., 2005, Extracting time-domain Green's function estimates from ambient seismic noise: Geophysical Research Letters, 32, 1–5.

Sava, P., and S. Fomel, 2006, Time-shift imaging condition in seismic migration: Geophysics, 71, no. 6, S209–217.

Schuster, G. T., 1985, A hybrid BIE+Born series modeling scheme: Generalized Born series: J. acoust. Soc. Am., 77, 865-879.

Schuster, G. T., 2001, Theory of daylight/interferometric imaging: Tutorial: 63rd Annual Conference and Exhibition, EAGE, Extended Abstracts, A32-1-A32-4.

Schuster, G. T., 2009, Seismic Interferometry: Cambridge University Press.

Schuster, G. T., J. Yu, J. Sheng, and J. Rickett, 2004, Interferometric/daylight seismic imaging. Geophysical Journal International, 157, 838-852.

Schuster, G. T. and M. Zhou, 2006, A theoretical overview of model-based and correlation-based redatuming methods: Geophysics, 71, no. 4, SI103-SI110.

Sen, M. K., 2006, Seismic inversion: 120 pp, Socitey of Petroleum Engieers Publications, Richardson, Texas.

Sen, M. K., and P. L. Stoffa, 1991, Nonlinear one-dimensional seismic waveform inversion using simulated annealing: Geophysics, 56, 1624–1638.

Sen, M. K., and P. L. Stoffa, 1995, Global optimization methods in geophysical inversion: Elsevier Science Publ. Co., Inc.

Sen, M. K., and I. G. Roy, 2003, Computation of differential seismograms and iteration adaptive regularization in prestack waveform inversion: Geophysics, 68, 2026–2039.

Sen, M. K., and A. Mukherjee, 2003, t-p analysis in transversely isotropic media, Geophysical Journal International: 154, 647-658.

Shin, C., and D. Min, 2006, Waveform inversion using a logarithmic wavefield: Geophysics, 71, no. 3, R31–R42.

Shin, C., and Y. H. Cha, 2008, Waveform inversion in the Laplace domain: Geophysical Journal International, 173, 922–931.

Shin, C., Jang, S., Min, D.-J., 2001, Improved amplitude preservation for prestack depth migration by inverse scattering theory: Geophysical Prospecting, 49, 592-606.

Sil, S. and M. K. Sen, 2009, Seismic critical-angle anisotropy analysis in the tau-p domain: Geophysics, 74, no. 4, A53–A57.

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

Snieder, R., A. Grêt, H. Douma, and J. Scales, 2002, Coda wave interferometry for estimating nonlinear behavior in seismic velocity, Science, 295, 2253-2255.

Snieder, R., K. Wapenaar, and K. Larner, 2006, Spurious multiples in seismic interferometry of primaries, Geophysics, 71, no. 4, SI111-SI124.

Srivastava, R. P., and M. K. Sen, 2009, Fractal-based stochastic inversion of poststack seismic data using very fast simulated annealing: Journal of Geophysics and Engineering, 2009, 6, 412-425.

Srivastava, R. P., and M. K. Sen, 2010, Stochastic inversion of pre-stack seismic data using fractal_based initial models: Geophysics, 75, no. 3, R47-R59.

Stefani, J., and S. D. Gopa, 2001, On the power-law behavior of subsurface heterogeneity: 61st Annual International Meeting, SEG, Expanded Abstracts, 2033–2036.

Stoffa, P. L., ed., 1989, Tau-p: A plane wave approach to the analysis of seismic data: Kluwer Academic Publishers.

Stoffa, P. L., M. K. Sen, R. Seifoullaev, R. Pestana, and J. Fokemma, 2006, Plane-wave Depth migration: Geophysics,71, no. 6, S261-S272.

Symes, W. W., 2007, Reverse-time migration with optimal checkpointing: Geophysics, 72, no. 5, SM213–222.

Tang, Y., 2009, Target-oriented wave-equation least-squares migration/inversion with phase-encoded Hessian: Geophysics, 74, no. 6, WCA95–WCA107.

Tang, Y. and S. Lee, 2010, Preconditioning full waveform inversion with phase-encoded Hessian, 80th Annual International Meeting, SEG, Expanded Abstracts, 1034–1038.

Tatham, R. H., 1989, Tau-p filtering, in P. L. Stoffa, ed., Tau-p: A plane wave approach to the analysis of seismic data: Kluwer Academic Publishers, 35–70.

Tao Y., K. Spikes and M. K. Sen, 2011, Stochastic seismic inversion using both fractal and low-frequency priors: 81st Annual International Meeting, SEG, Expanded Abstracts, 2732-2736.

Tao Y. and M. K. Sen, 2012, Shallow splay fault properties of the Nankai Trough accretionary wedge inferred from seismic inversion: J. Geophys. Eng., 9, 1-11.

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

Tarantola, A., 1987, Inverse problem theory: Methods for data fitting and model parameter estimation: Elsevier Scientific Publ. Co., Inc.

Tarantola, A., 1988, Theoretical background for inversion of seismic reflection data: Pure and Applied Geophysics, 128, 365–399.

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

Thomson, W. T., 1950, Transmission of elastic waves through a stratified solid medium, Journal of Applied Physics, 21, 89-93.

Trifunac, M. D., 1971, Surface motion of a semi-cylindrical alluvial valley for incident plane SH waves: Bull. seism. Soc. Am., 61, 1755-1770.

Tromp, J., C. Tape, and Q. Liu, 2005, Seismic tomography, adjoint methods, time reversal, and banana-doughnut kernels: Geophysical Journal International, 160, 195–216.

Vasconcelos, I. and R. Snieder, 2008a, Interferometry by deconvolution, Part 1—Theory for acoustic waves and numerical examples, Geophysics, 73, no. 3, S115-S128.

Vasconcelos, I., R Snieder, and B. Hornby, 2008b, Imaging internal multiples from subsalt VSP data — Examples of target-oriented interferometry: Geophysics, 73, no. 4, S157-S168.

Vedanti, N., and M. K. Sen, 2009, Seismic inversion tracks in situ combustion: A case study from Balol oil field, India: Geophysics, 74, no. 4, B103–B112.

Vidal, A. C., J. van der Neut, D. Draganov, G. G. Drijkoningen, and K. Wapenaar, 2011, Retrieval of reflections from ambient noise using the incident fields (point-spread function) as a diagnostic tool: 73th Conference and Exhibition, EAGE.

Vigh, D., and E. Starr, 2008, 3D prestack plane-wave, full waveform inversion: Geophysics, 73, no. 5, VE135–VE144.

Virieux, J., and S. Operto, 2009, An overview of full waveform inversion in exploration geophysics: Geophysics, 74, no. 6, WCC1–WCC26.

Waldhauser, F., and W. L. Ellsworth, 2000, A Double-Difference Earthquake Location Algorithm: Method and Application to the Northern Hayward Fault, California: Bulletin of the Seismological Society of America, 90, 1353–1368.

Wapenaar K., 2003, Synthesis of an inhomogeneous medium from its acoustic transmission response, Geophysics , 68, 1756-1759.

Wapenaar, K., 2004, Retrieving the elastodynamic Green's function of an arbitrary inhomogeneous medium by cross correlation: Physical Review Letters, 93, 254301.1-254301.4.

Wapenaar K., 2006, Green's function retrieval by cross-correlation in case of one-sided illumination: Geophysical Research Letters, 33, L19304-1–L19304-6.

Wapenaar K., D. Draganov, R. Snieder, X. Campman, and A.Verdel, 2010a, Tutorial on seismic interferometry: Part 1 - Basic principles and applications: Geophysics, 75, no.5, 195-209.

Wapenaar, K., E. Slob, R. Snieder, and A. Curtis, 2010b, Tutorial on seismic interferometry: Part 2 - Underlying theory and new advances: Geophysics, 75, no.5, 211-227.

Watanabe, T., S. Shimizu, E. Asakawa, and T. Matsuoka, 2004, Differential waveform tomography for time-lapse crosswell seismic data with application to gas hydrate production monitoring: 74th Annual International Meeting, SEG, Expanded Abstracts, 2323–2326.

Whitmore, N. D., 1983, Iterative depth migration by backward time propagation: 53rd Annual International Meeting, SEG, Expanded Abstracts, 382–385.

Whitmore, N. D., 1995, An imaging hierarchy for common angle plane wave seismograms: PhD thesis, University of Tulsa.

Xu, S., Y. Zhang, and B. Tang, 2011, 3D angle gathers from reverse time migration: Geophysics, 76, no. 2, S77–S92.

Yan, R., and X. B. Xie, 2009, A new angle -domain imaging condition for prestack reverse-time migration: 79th Annual International Meeting, SEG, Expanded Abstracts, 2784-2787.

Yoon, K., and K. J. Marfurt, 2006, Reverse-time migration using the pointing vector: Exploration Geophysics, 37, 102–107.

Zhang, R., and J. Castagna, 2011, Seismic sparse-layer reflectivity inversion using basis pursuit decomposition: Geophysics, 76, no. 6, R147–R158.

Zhang, Y., Sun, J., C. Notfors, Gray, S. H., Chernis, L., and Young, J., 2005, Delayed-shot 3D depth migration: Geophysics, 70, no. 5, E21–E28.

Zhang, Y. and D. Wang, 2009, Traveltime information-based wave-equation inversion, Geophysics, 74, no. 6, WCC27–WCC36.

Zhang, Y., and G. Zhang, 2009, One-step extrapolation method for reverse time migration: Geophysics, 74, no. 4, A29–A33.

Zhang, Y., and J. Sun, 2009, Practical issues in reverse-time migration: True amplitude gathers, noise removal and harmonic-source encoding: First Break, 26, 19–25.

Zheng Y., Barton P., and Singh S., 2011, Strategies for elastic full waveform inversion of time-lapse Ocean Bottom Cable (OBC) seismic data, 81st Annual Meeting, SEG, Expanded Abstracts, 4195-4200.

Vita

Yi Tao was born on February 6, 1983 (or December 24 in lunar calendar), in Chongqing, to Hongxiang Tao and Cuilan Chen. He entered University of Science and Technology of China in Hefei, Anhui, China in 2001 and received a BS degree in Geophysics in 2005. He then joined Institute of Geology and Geophysics, Chinese Academy of Sciences in Beijing, China and received M.S. in Geophysics in 2008. In fall, 2008, he entered the University of Texas at Austin to start his PhD study in Jackson School of Geosciences, the University of Texas at Austin. During the summer of 2010 and 2011, he worked twice as a geoscience intern for ConocoPhillips in Houston, Texas. After graduation, he will join ConocoPhillips.

Permanent address (or email): 7117 Wood Hollow Dr., Apt 921, Austin, TX, 78731

This dissertation was typed by Yi Tao.