RICE UNIVERSITY

The Nonlinear Differential Semblance Algorithm for Plane Waves in Layered Media

by

Dong Sun

A Thesis Submitted
in Partial Fulfillment of the
Requirements for the Degree

Master of Arts

Approved, Thesis Committee:

William W. Symes, Noah G. Harding Professor, Chair
Computational and Applied Mathematics

Matthias Heinkenschloss, Professor
Computational and Applied Mathematics

Yin Zhang, Professor
Computational and Applied Mathematics

Houston, Texas

April 2008

Abstract

The Nonlinear Differential Semblance Algorithm for Plane Waves in Layered Media

by

Dong Sun

This thesis proposes an alternative approach to the output least-squares (OLS) seismic inversion for layered-media. The latter cannot guarantee a reliable solution for either synthetic or field data, because of the existence of many spurious local minima of the objective function for typical data, which lack low-frequency energy. To recover the low-frequency lacuna of typical data, I formulate waveform inversion into a differential semblance optimization (DSO) problem with artificial low-frequency data as control variables. To my knowledge, this approach is the first version of differential semblance with non-linear modeling that may properly accounts for nonlinear effects of wave propagation, such as multiple reflections. Numerical experiments with synthetic data indicate the smoothness and convexity of the proposed objective function. These results suggest that gradient-related algorithms may successfully approximate a global minimizer from a crude initial guess for typical band-limited data.

Acknowledgements

It is a pleasure to thank the many people who made this thesis possible.

I would like to express my deep and sincere gratitude to my thesis advisor Dr. William Symes for his consistent support and encouragement throughout my research work, for his sound advice, good guidance, and lots of great ideas, for his precious time on proof-reading my manuscripts word by word, and invaluable help on many other things.

I would like to express my sincere gratitude to my committee, Dr. Matthias Heinkenschloss and Dr. Yin Zhang, for their extensive edification, advice and help.

I also owe my sincere gratitude to ExxonMobil Upstream Research Company for giving me the great opportunity to work there during the summer of 2008. I extended my work and learned a lot from many of the experts there. Especially, David Hinkley and Jerry Krebs helped me a lot and provided me with lots of good advice on this project.

My sincere appreciations go to Dr. Janice Hewitt and Dr. Danny Sorensen for their great teaching and advice on improving my thesis writing and oral presentation skills. And great thanks go to Yang He, Rami Nammour, Chao Wang, Igor Terentyev, Klaus Wiegand, Saifon Chaturantabut, Davood Shamsi, and Letty for their help and advice during my writing this thesis.

This research was supported by The Rice Inversion Project (TRIP). Current TRIP Sponsors are Amerada Hess, BP, Chevron, ConocoPhillips, ExxonMobil, Landmark Graphics, Shell Research, Total E&P USA, and WesternGeco.

Last but not the least, I would like to express my most gratitude to my beloved wife and families for all their love, support and great encouragement.

Contents

List of Figures

Chapter 1

Introduction

In this thesis, I propose a nonlinear differential semblance algorithm to solve the inverse problem of reflection seismology for a layered constant-density acoustic media. I describe an implementation and provides some evidence that this approach may avoid the severe convergence difficulties associated with the classical output least squares (OLS) seismic inversion, and accounts in a natural way for nonlinear effects (such as multiple reflection) frequently encountered in actual data.

The introduction chapter is intended to provide a historical and scientific base for the work presented and put the thesis into context. More specifically, this chapter starts with an overview of the underling inverse problem of reflection seismology (Section 1.1). Section 1.2 then describes a common approach to this inverse problem — waveform inversion. Particularly, I discuss some important references about the OLS inversion and its intrinsic difficulties, which motivate this work. Next, Section 1.3 renders the idea, goal, and contribution of this thesis. Finally, this chapter ends with an agenda of this thesis.

1.1 Overview of the Inverse Problem

This section briefly reviews the geophysical experiment leading to this inverse problem.

A common objective of reflection seismology is to make inferences about physical features (*model*) of subsurface (e.g., velocity distribution, impedance profile, etc.) from data (*seismogram*) recorded on or near the surface. In general, with reasonably idealized setting, the laws of physics provide the governing equations for computing the data values given a model. This is called the *forward problem*. A common idealized setting in reflection seismology is based on the assumption that the earth is a linearly acoustic isotropic body supporting wave propagation governed by acoustic wave equations.

In the inverse problem, the reflection seismic experiment introduces a mechanical vibration at a point on or near the surface of the earth; the mechanical response of the earth to the excitation is measured on or near the surface; the aim is to reconstruct the physical properties (model) from a set of measurements (data). Usually, this inverse problem does not have unique solutions, because: (1) the amount of data is finite and cannot carry sufficient information to determine the model uniquely (underdetermination), or, (2) the data has more degrees of freedom than those of the desired model (overdetermination) and are inconsistent (because of measurement errors).

Fortunately, it is possible to construct a type of inverse through minimization of an objective function that measures the difference between two points in the data space. Thus, the inversion becomes a model-based data-fitting process that provides a "best fit" solution to the inverse problem. This inversion is called *waveform inversion*,

and has been studied extensively over 30 years. A common objective function is the "least squares" objective function, which yields the maximum likelihood criterion if experimental errors have Gaussian distributions Tarantola and Vallette (1982). However, this approach (*output least squares inversion*) is impractical in exploration seismology because of some intrinsic difficulties, particularly the existence of many spurious local minima of the objective function for typical seismic data, which lacks low-frequency energy.

This thesis proposes an alternative approach to waveform inversion for layered acoustic media to avoid the severe convergence difficulties associated with the output least squares (OLS) inversion. To recover the low-frequency lacuna of typical data, I formulate waveform inversion into a differential semblance optimization (DSO) problem with artificial low-frequency data as control variables. To my knowledge, this approach is the first version of differential semblance algorithm with non-linear modeling that properly accounts for nonlinear effects (such as multiple reflections) of wave propagation. Numerical experiments with synthetic data indicate the smoothness and convexity of the proposed objective function so that gradient-related algorithms can approximate the global minimum from a coarse initial guess for typical band-limited data.

The following paragraphs present a literature review on waveform inversion, and describe the alternative developed in the rest of the thesis.

1.2 Waveform Inversion

Waveform inversion is an important model-based data-fitting approach to reflection seismology, which aims to determine the features of subsurface structure from seismic

reflection data collected by groups of receivers (hydrophones or geophones) located on or near surface. The most familiar objective function for waveform inversion is the "least squares" objective function. It is popular because: (1) it is very simple and yields the maximum likelihood criterion if experimental errors have Gaussian distributions; (2) it does not require picked travel time and can take into account essentially any physics of seismic wave propagation and reconstruct detailed features of subsurface structure.

Tarantola and Valette (1982) stated a general definition of the nonlinear least squares inversion, which is valid for various kinds of problems (including discrete and continuous, overdetermined and underdetermined, linear and nonlinear problems). Here comes an abstract setting for the least-squares inverse problem over a constant density acoustics media: The model space $\mathcal{M}$ is a set of possible velocity distributions v, and usually of large degrees of freedom especially for three-dimensional problems (e.g., 50^{50}); the data space $\mathcal{D}$ consists of samples d of reflection response (data) on or near the surface over a time interval. $\mathcal{D}$ is regarded as a Hilbert space with norm $\|.\|$. The forward map $S : \mathcal{M} \to \mathcal{D}$ is a function of the input velocity model v, denoted by $S[v]$, which builds a nonlinear relation between $\mathcal{M}$ and $\mathcal{D}$. The simplest version of data fitting inversion is an Output Least Squares problem:

$$\min_{v \in \mathcal{M}} J_{OLS} := \frac{1}{2}\|S[v] - d_{obs}\|^2.$$

Most attempts to minimize J_{OLS} are to compute the gradient of J_{OLS} with respect to v and search in the direction related to this gradient for an update. The gradient vanishes at a stationary point, which could be a minimum of J_{OLS}. Gauss-Newton and nonlinear conjugate gradient are examples of these kinds of methods. With some

version of the L^2 norm in M, the gradient can be computed through standard adjoint state method and written as $\nabla_v J_{OLS} = DS[v]^* (S[v] - d_{obs})$, where $DS[v]^*$ is the adjoint of the linearized forward map $DS[v]$ of $S[v]$ at the point v. In Gauss-Newton algorithm, the searching direction can be expressed as $\left(-(DS^* DS)^\dagger \nabla_v J\right)$, which is the solution of the linearized least squares problem

$$\min_{\delta v} \frac{1}{2} \left\| DS[v]\delta v - (S_{data} - S[v]) \right\|^2 .$$

Lailly (1983) applied the adjoint state method to seismic inverse problem and found that DS^* is equivalent to a migration operator. The linearized inversion can be computed through conjugate gradient method. Tarantola (1984a) discussed solving the linearized problem using iterative algorithms, and showed that the rigorous solution of the linearized seismic inversion can be achieved using the classical methods of migration. As a generalization, Tarantola (1984b) developed a gradient-related iterative approach to solve the nonlinear least-squares inverse problem in the acoustic approximation for seismic reflection data with nonlinear effects (such as multiple reflection). No numerical examples were provided in Tarantola (1984b).

These kinds of methods are called Newton-like iterative approaches, which only use local information of a current iterate v and yield local convergence. It is possible to use global optimization methods to minimize J_{OLS} such as genetic (Sen and Stoffa, 1991b) and simulated annealing (Sen and Stoffa, 1991a) methods. These methods use some random search strategies to traverse the model space in order to find the global minimum which corresponds to the smallest objective value. Though global methods don't need a good start model and gradient, they require a great many of evaluations of the objective function (forward problem) before they converge. Considering that

a model space in reflection seismology usually has millions or even billions degrees of freedom, global methods are currently infeasible. So far, only Newton-like methods have been computationally feasible for such a large scale problem.

In this work, I use conjugate gradient method to obtain 1D linearized inversion as an updating direction for Gauss-Newton algorithm which is applied to solve the nonlinear least squares problem. More discussion about the least squares inversion comes in Chapter 2.

Modifications of the OLS Inversion

Though the OLS inversion with Newton-like approaches is conceptually attractive, its applications in reflection seismology have been strictly restricted by two major obstacles (Symes, 2007). The first is the computational intensity of wave field modeling and various computation required by the OLS inversion, especially in 3D. This computational obstacle is weakening with continuous advances in computer hardware and simulation techniques.

The second obstacle is more fundamental. OLS objective function is very ill-conditioned and has many spurious local minima which will trap any Newton-like iteration. Therefore this inversion doesn't work with any Newton-like optimization method unless the starting velocity model is so accurate that it has the same velocity trend (long scale structure) as the true velocity model. This fact is well observed and discussed in literature. Gauthier et al. (1986) implemented the first realistic example of multidimensional, nonlinear inversion of multioffset seismic reflection data and proved the feasibility of the nonlinear inversion method proposed by Tarantola (1986). This paper showed that the OLS problem is strongly nonlinear and has secondary

minima. Also, it demonstrated that the OLS inversion is good for estimating short scale structure but cannot recover the long scale structure. Santosa and Symes (1989) explored in detail the success and limitations of the OLS inversion in the context of the layered velocity model. They partly released the obstacle by redefining the least-squares problem to match only the precritical part of the data. But their approach still essentially suffered from the same impedient discussed above. Symes and Carazzone (1992) illustrated the high non-convexity of the OLS objective function clearly via a plot of the mean square error over a line segment connecting constant back ground velocity with the reference velocity. I present similar plots in Chapter 3 for both the OLS inversion and the proposed method. These plots demonstrate the proposed method is superior to the OLS inversion for layered media.

The main factor appears to drive the above behavior of output least squares inversion is the band-limitation of typical field data, especially the lack of low frequencies, which leads to the reconstruction ambiguous (Santosa and Symes, 1989). Lots of work has shown that the impedance as a function of vertical travel time in a layered acoustic medium could be reconstructed from the impulse response, which contains all frequency components down to 0 Hz, (Bamberger et al., 1979; Symes, 1981, 1986b; Sacks and Santosa, 1987). For several dimensional problem, numerical examples indicate that impulse responses may determine constant-density acoustic models via the OLS inversion (**?**Shin and Min, 2006). Low-frequency data appear to contain information about the trend of the true model. The OLS inversion cannot infer the velocity trend from bandlimited reflection data.

Many attempts have been tried to deal with the local minima issue associated with the OLS inversion.

A number of papers tried to diminish the problem of local minima by a decomposi-

tion of the seismic inversion problem by scale. Kolb et al. (1986) suggested a pre-stack continuum inversion algorithm for 1D acoustic medium. This algorithm first recovers the low-frequency trend of the velocity model via inversion of the low-frequency part of the data. Next, a progressive downward determination process is employed to infer the velocity distribution layer by layer. The numerical results demonstrate the efficiency of this continuum inversion process only for data with the very low-frequency components. This approach inspires me with the continuum low-frequency inversion strategy for data with low-frequency components down to 0 Hz, which leads to a much more efficient approach than the conventional inversion approach. I use this strategy to solve the least-squares subproblem embedded in the proposed algorithm. For 2D pre-stack seismic inversion, **?** showed that a multiscale approach is effective in releasing the difficulty of local minima only for data with much lower frequencies than is normally available in realistic seismic data sets.

Shin and Min (2006) introduced a new objective function to overcome the non-convexity associated with the classical objective function. The new objective can take into account phase and amplitude separately or simultaneously, and then yield three different inversions. Some tests showed that this approach could lead to a better result than the conventional least-squares inversion for some synthetic data with very low-frequencies down to 0.3121 Hz. While the inversion results were not good for data without frequencies below 5 Hz.

All the above approaches adopt special strategies to solve the OLS inversion. But none essentially release the problem of local minima that has been the main impediment to full waveform inversion.

Differential Semblance

In contrast, the differential semblance approach is based on a modified least-squares principle which essentially avoids the non-convexity of the OLS inversion, hence leads to a well-behaved inversion (Symes and Carazzone, 1991; Symes, 1991a, 1993, 1999). Many versions of the differential semblance algorithm have been implemented all based on linearized scattering theory, e.g., (Symes and Versteeg, 1993; Chauris and Noble, 2001; Verm and Symes, 2006; Li and Symes, 2007). Some theoretical evidence exists that a similar algorithm based on (nonlinear) scattering might be feasible, and account in a natural way for nonlinear effects (such as multiple reflection) frequently encountered in actual data (Symes, 1991b). This thesis aims to develop and clarify such a nonlinear differential semblance algorithm in a relatively simple wave propagation framework, that of plane waves propagating and scattering in a layered acoustic medium.

1.3 Idea, Goal, and Claim of the Proposed Method

The task of this thesis is to introduce an alternative approach to the OLS inversion for layered acoustic media.

Through Radon transform, I decompose the original wave equation into a series of 1D plane-wave equations characterized by vertical velocities $v[c,p] = \frac{c}{\sqrt{1-c^2p^2}}$, where c is the true depth dependent velocity model, p is the plane-wave slowness. Each of these 1D problems corresponds to a nonlinear least squares inversion. It has been shown that 1D models could be reconstructed from their impulse responses which are the data with all frequencies, especially the very low frequencies (Bube and Burridge, 1983; Symes, 1986b). I demonstrate this key fact by some numerical experiments

in Chapter 2. This fact means that with the very low-frequency data, each 1D least squares inversion yields a unique "best fit" solution. Therefore, I can view the velocity model $v[c, p]$ as a function of the data, via the solvability of the 1D impulsive inverse problem.

However, field experiments don't have the very low frequency data. The most important goal of this thesis is trying to get around the limitation of not having the lowest frequencies available. The central idea of this work is to add in artificial low frequency data as controls in order to make the 1D nonlinear inversions solvable. From the previous discussion, each $v[c, p]$ can be regarded as a function of the low frequency controls. Through those 1D inversions, I get a set of vertical velocities $v[c, p]$, from which the target velocity models $c[p](z)$ are computed. If the low frequency data added at the beginning was correct, all the vertical velocities v should lead to the same $c[p](z)$, i.e. the velocity model $c[p](z)$ should not depend on slowness p, because the earth is unique.

Hence, this inverse problem is posed as:

$$\min_{d_l} \frac{1}{2} \|Q[v]\|^2 \quad s.t. \quad \|S_l[v] - (d_{obv} + d_l)\| \simeq 0,$$

where $Q[v]$ satisfies that $Q[v] = 0 \Rightarrow \frac{\partial c}{\partial p} = 0$, d_{obv} are the observed band-limited data, and d_l are the low frequency controls.

This objective function aims to quantify the *coherency condition* ($\frac{\partial c}{\partial p} = 0$) and is parameterized over a space of low-frequency data. Chapter 3 demonstrates the convexity of the objective function over line segments that connect the true data with the data which have the true high frequency data components but some specified or randomly perturbed low frequency components. These "scan" experiments demonstrate

the smoothness and convexity of the proposed DS objective function, and thus in some extent verify the feasibility and efficiency of Newton-like optimization method. I derives the theoretical computation of the gradient of the objective function with respect to low-frequency controls (Appendix A). Thereby, a standard Newton-like optimization approach can be used to solve this inverse problem. Implementation of the proposed algorithm will be a future work.

The next chapter presents the method in detail. Chapter 4 shows some numerical results that demonstrate the feasibility of the proposed method. In Chapter 5, the thesis ends with some conclusions and discussions of prospects for further development.

Chapter 2

Theory and Method

This chapter elaborates the construction of the proposed differential semblance approach via three sections. First, I set up the layered constant-density acoustic model, introduce the plane-wave decomposition and present the original settings via Extended Modeling concept, which renders a general form of the inverse problem of reflection seismology. Second, as the original problem is reduced to a set of one-dimensional plane-wave problems, I review some results about one-dimensional inverse problem of wave propagation, and the important role played by the very low-frequency information. And I try to see the relation between velocity models and low-frequency components of data. Third, based on Extended Modeling concept and the relation between velocity models and low-frequency information, I propose the differential semblance approach with nonlinear modeling.

2.1 The Plane Wave Model and its Extension

As the first step of developing an algorithm, this thesis concerns a simple model for reflection seismology — the layered constant-density acoustic model — to simplify the derivation and focus on the main tasks: avoiding the intrinsic difficulties of the OLS inversion, recovering the low-frequency lacuna, and accounting for nonlinear effects of wave propagation. The layered media assumption is a reasonable approximation as sediments retain a large degree of lateral homogeneity in many areas. With this assumption, the ideas of the proposed approach could be most clearly expressed and illustrated numerically, and some rigorous mathematical backup is available. Extensions of the proposed method to more complicated media go beyond the intention of the present paper.

In the layered constant-density acoustic model, the wave field potential $u(x, z, t)$ $(x, z \in I\!R)$ is governed by the wave equation

$$\begin{aligned} \left(\frac{1}{c^2(z)}\frac{\partial^2}{\partial t^2} - \nabla^2\right) u(x, z, t) &= \omega(t)\delta(x, z), \\ u(x, z, t) = u_t(x, z, t) &\equiv 0, \quad t < 0, \end{aligned} \tag{2.1}$$

where $c(z)$ is the sound velocity, and the right-hand side is an isotropic point energy source with the source wavelet $\omega(t)$. Notice that $\omega(t)$ is usually chosen to be band-limited, as is required by observations of the spectra of seismograms: for various physical limitations, real reflection seismograms don't have fourier components at very low ($< \xi_l$ Hz) and very high ($> \xi_h$ Hz) temporal frequencies [1]. Assume that $c(z)$ is a function only of the depth variable z and subject to some constraints and regularity condition, such as $0 < c_{min} \leq c(z) \leq c_{max}$, $c(z) = c_0$ for $z < 0$, $c(z) = c_b$

[1] The positive numbers ξ_l and ξ_h depend on specific physical settings of real experiments. For example, $\xi_l = 5, \xi_h = 60$.

for $z > z_{max}$, and $c \in H^1_{loc}(I\!R)$.

The seismogram is a sampling of the pressure field $\frac{\partial u}{\partial t}(x, 0, t)$ at a number of "receiver" points over a time interval $0 \leq t \leq t_{max}$. I adopt the idealization that the "receiver" points form the continuum $z = 0$ and that the measurement of $\frac{\partial u}{\partial t}$ is continuous in time. Given the source time function $\omega(t)$, the seismogram becomes a function of the sound velocity:

$$s[c](x,t) := \frac{\partial u}{\partial t}(x,0,t), \quad 0 \leq t \leq t_{max}.$$

The goal is to find $c(z)$ for $0 \leq z \leq z_{max}$ from the observed seismogram s_{data} such that $s[c] \simeq s_{data}$.

Now, we introduce the Radon transformed field

$$U(p,z,t) = \int dx\, u(x,z,t+px), \;\; p \in I\!R.$$

Given the layered medium assumption and $|p|c(z) < 1$, I decompose the original problem into a set of 1-D plane-wave problems

$$\left(\frac{1}{v^2(p,z)}\frac{\partial^2}{\partial t^2} - \frac{\partial^2}{\partial z^2}\right) U(p,z,t) = \omega(t)\delta(z), \qquad (2.2)$$
$$U(p,z,t) \equiv 0, \quad t < 0,$$

where the *vertical velocity* $v(p,z) = c(z)/\sqrt{1 - c^2(z)p^2}$, and p denotes the ray parameter (slowness). The *vertical travel-time* is

$$T(z,p) = \int_0^z d\zeta \; \frac{1}{v(p,\zeta)}.$$

For any α with $0 < \alpha < 1$ (for example, $\alpha = 0.1$), let $p_{max} = \frac{1}{c_{max}}\sqrt{1-\alpha^2}$, and $t_{max}(p) = 2\,T(z_{max}, p)$ for $p \in [-p_{max}, p_{max}]$.
Then, the plane-wave seismogram is defined by

$$S_\omega[c](p,t) := \frac{\partial U}{\partial t}(p,0,t) \qquad \text{for all } (p,t) \in \mathbf{P}, \tag{2.3}$$

where

$$\mathbf{P} = \{(p,t) : |p| \le p_{max},\ 0 \le t \le t_{max}(p)\}\,.$$

Remark. In fact, one can define (α) precritical depth function

$$Z_\alpha(p) = \max\left\{z : 0 \le \zeta \le z, c(\zeta)|p| \le \sqrt{1-\alpha^2}\right\}.$$

Let $p^*_{max} = \max\left\{p : c(z)|p| \le \sqrt{1-\zeta^2},\ 0 \le z \le z_{max}\right\}$. Then the plane-wave seismogram can be defined on the (α) precritical region $\mathbf{P}^*$ defined by

$$\mathbf{P}^* = \{(p,t) : |p| \le p^*_{max},\ 0 \le t \le t_{max}(p)\}\,.$$

(Santosa and Symes, 1989, Chapter 2)
During the inversion, the region $\mathbf{P}^*$ is unknown. One needs some sophisticated strategy to update the computing region to make it as close to $\mathbf{P}^*$ as possible. For the simplicity of implementation, this thesis concerns a fixed domain $\mathbf{P}$, which is only a subset of the precritical region $\mathbf{P}^*$.

Notice that Equation (2.2) becomes a one-dimensional equation for each slowness p, which governs the propagation of a plane wave.

Given the plane-wave seismogram $d(p,t)$ (i.e., $d = S_{data}$ [2]), this thesis focuses on

[2] S_{data} can be computed from s_{data} by Radon transform. To focus on the principal algorithm, I

the inverse problem:

$$\begin{aligned} &\text{Find} \quad c(z) \in \mathbf{M} \\ &s.t. \quad S_\omega[c] \simeq d, \end{aligned} \tag{2.4}$$

where the *model space* $\mathbf{M}$ denotes a set of possible velocity models, incorporating bounds on values (e.g., $0 \leq c_{min} \leq c(z) \leq c_{max}$, $c(z) = c_0$ for $z < 0$, $c = c_b$ for $z > z_{max}$) and other regularity constraints (e.g., $c(z) \in H^1_{loc}(I\!R)$).

This inversion is often constructed through minimization of an objective function to obtain a "best fit" solution. The most common objective function is the least-squares objective function, which yields the OLS inversion. This thesis aims to construct a new objective function that leads to the proposed differential semblance (DS) approach to avoid the intrinsic difficulties associated with the OLS inversion discussed in Chapter 1.

Before elaborating the construction of the proposed DS approach, I would like to present the inverse problem (2.4) via a unifying concept — Extended Modeling, discussed in (Symes, 2007), which sets up a general framework for the inverse problem of reflection seismology.

Extended Modeling

Recall that regarding the source time function $\omega(t)$ is known, the forward map $S_\omega : \mathbf{M} \longrightarrow \mathbf{D}$ is defined by (2.3), where $\mathbf{D}$ is the *data space*.

---

leave out this computation and assume that S_{data} is known.

The *extended* model space $\overline{\mathbf{M}}$ is defined as

$$\overline{\mathbf{M}} := \{c(p,z) \quad : \quad 0 \le c_{min} \le c(p,z) \le c_{max} \ \texttt{for} \ 0 \le z \le z_{max} \ \texttt{and} \ |p| \le p_{max},$$
$$c(p,z) = c_b \ \texttt{for} \ z > z_{max}, c \in H^1\} .$$

Then, the corresponding vertical velocity to $c(p,z)$ is

$$v(p,z) := \frac{c(p,z)}{\sqrt{1 - c^2(p,z)\, p^2}},$$

and the extended forward map $\overline{S}_\omega : \overline{\mathbf{M}} \longrightarrow \mathbf{D}$ is defined as

$$\overline{S}_\omega[c(p,z)](p,t) := \frac{\partial U}{\partial t}(p,0,t) \qquad \text{for all } (p,t) \in \mathbf{P}, \tag{2.5}$$

where $U(p,z,t)$ satisfies (2.2).
The extension operator $E : \mathbf{M} \longrightarrow \overline{\mathbf{M}}$ is defined as

$$E[c(z)] := c(p,z),$$

where $c(p,z) = c(z)$ for all $z \in I\!R$.

Then, the *extension* of model $S_\omega : \mathbf{M} \longrightarrow \mathbf{D}$ consists of

- the extended model space $\overline{\mathbf{M}}$;
- the extension operator $E : \mathbf{M} \longrightarrow \overline{\mathbf{M}}$;
- the extended modeling operator $\overline{S}_\omega : \overline{\mathbf{M}} \longrightarrow \mathbf{D}$ satisfying $S_\omega[c] = \overline{S}_\omega[E[c]]$ for any $c \in \mathbf{M}$.

For this extended model, the extended inverse problem is: given $d \in \mathbf{D}$, find $\bar{c}(p,z) \in$

$\overline{\mathbf{M}}$ so that $\overline{S}_\omega[\bar{c}] \simeq d$. Notice that $E[\mathbf{M}] \subset \overline{\mathbf{M}}$. $E[\mathbf{M}]$ corresponds to the "physical models". While the extended models could be "unphysical" in the sense that $\bar{c}(p,z) \in \overline{\mathbf{M}}$ could vary with respect to p. A solution $\bar{c}$ is physically meaningful only if $\bar{c} = E[c]$ for some $c \in \mathbf{M}$, and then c is a solution of the original inverse problem, i.e., $S_\omega[c] = \overline{S}_\omega[\bar{c}] \simeq d$. That is, to find a solution to the extended inverse problem that belongs to $E[\mathbf{M}]$ to solve the original inverse problem. Generally, "$\simeq$" is in the least-squares sense.

To turn the inverse problem into an optimization problem, we need to figure out an operator to measure the extent to which a solution to the extended inverse problem is physically meaningful. Since $E[\mathbf{M}]$ is a linear subspace of $\overline{\mathbf{M}}$, any linear operator vanishing on this subspace gives rise to a quadratic form which can serve as such an objective. An *annihilator* is a map $A[\bar{c}]$ from $\overline{\mathbf{M}}$ to some other Hilbert space $\mathbf{H}$ so that

$$\bar{c} \in E[\mathbf{M}] \Longleftrightarrow A[\bar{c}] = 0.$$

With the above notations, a general form of the inverse problem could be stated as

$$\begin{aligned} \min_{\bar{c}\in\overline{\mathbf{M}}} \quad & J_A[\bar{c}, d] := \frac{1}{2}\|A[\bar{c}]\|^2_{\mathbf{H}} \\ s.t. \quad & \|\overline{S}_\omega - d\|^2_{\mathbf{D}} \simeq 0, \end{aligned} \tag{2.6}$$

where $\|\cdot\|_{\mathbf{H}}$ and $\|\cdot\|_{\mathbf{D}}$ respectively stand for Hilbert norms in the space $\mathbf{H}$ and $\mathbf{D}$.

Recall that problem (2.1) is reduced to a set of 1D plane-wave problems (2.2) via Radon Transform. Before building the proposed DS approach, I would like to review some results about 1D inverse problem of wave propagation and extract some relation between low-frequency components of data and extended velocity models. Based on this relation, in section 2.3, I will propose the differential semblance approach via using

some type of differential semblance as an annihilator. In Chapter 3, I will illustrate the behavior of the corresponding objective J_{DS} through numerical experiments (scan tests).

2.2 1D Plane-Wave Problem and Low-frequencies' Influence

In this section, I will recall some results about the one-dimensional inverse problem of reflection seismology and try to build some relation between low-frequency information and velocity models.

Recall the notations given in section 2.1. Equation (2.2) is a one-dimensional form for each fixed p, which governs the propagation of a plane wave.

For fixed $p \in \{p : |p| \leq p_{max}\}$, let's define the map $F_{p,\omega} : \mathbf{H_1}[0, z_{max}] \longrightarrow \mathbf{H_2}[0, t_{max}(p)]$ as

$$F_{p,\omega}[y] := \frac{\partial U}{\partial t}(p, 0, t) \qquad \text{for all } 0 \leq t \leq t_{max}(p),$$

where U satisfies (2.2) with source wavelet $\omega(t)$ and $v(p, z) = y(z)$ for $z \in [0, z_{max}]$. For the choice of $\omega = \delta$, the corresponding $\frac{\partial U}{\partial t}(p, 0, t)$ is called an impulsive response. Both this source and its response have all frequency components. In this case, given the members of $\mathbf{H_1}$ and $\mathbf{H_2}$ are sufficiently differentiable, $F_{p,-\delta}$ is one-to-one. That is, given $d(p, t) \in \texttt{Range}(F)$ $(0 \leq t \leq t_{max}(p))$, there is uniquely determined $v(p, z)$ on $[0, z_{max}]$ so that $F_{p,-\delta}[v(p, z)] = d(p, t)$. See (Bamberger et al., 1979; Symes, 1981, 1986b; Sacks and Santosa, 1987) for details.

If we change the depth scale by replacing depth z by its corresponding travel time $\tau(z) = \int_0^z \frac{dz}{v(p,z)}$, then similarly we can define a map $\widetilde{F}_{p,\omega} : \mathbf{H_1}[0, t_{max}(p)/2] \longrightarrow$

$\mathbf{H_2}[0, t_{max}(p)]$ as

$$\widetilde{F}_{p,\omega}[\tilde{y}] := \frac{\partial \tilde{U}}{\partial t}(p, 0, t) \qquad \text{for all } 0 \leq t \leq t_{max}(p),$$

where $\tilde{U}$ defined by

$$\tilde{U}(p, \tau(z), t) := U(p, z, t),$$

satisfies

$$\begin{aligned} \frac{1}{\tilde{v}(p,\tau)} \frac{\partial^2 \tilde{U}}{\partial t^2} - \frac{\partial}{\partial \tau} \frac{1}{\tilde{v}(p,\tau)} \frac{\partial \tilde{U}}{\partial \tau} &= 0 \qquad (2.7) \\ \frac{\partial \tilde{U}}{\partial \tau} &= v(p,0)\omega(t) \\ \tilde{U}(p,\tau,t) = \tilde{U}_t(p,\tau,t) &\equiv 0, \quad t << 0. \end{aligned}$$

Let $r = \frac{\partial}{\partial \tau}(\log \tilde{v})$ (i.e., $\tilde{v} = \exp(\int_0^\tau r)$), then equation (2.7) becomes

$$\frac{\partial^2 \tilde{U}}{\partial t^2} - \frac{\partial^2 \tilde{U}}{\partial \tau^2} + r \frac{\partial \tilde{U}}{\partial \tau} = 0, \qquad (2.8)$$

which defines a map $\widehat{F}_{p,\omega}[r] = \frac{\partial \tilde{U}}{\partial t}(p, 0, t)$, $0 < t < t_{max}(p)$.

Note that

$$\widehat{F}_{p,\omega}[r] = \omega * \widehat{F}_{p,\delta}[r]. \qquad (2.9)$$

Given $\omega = \delta$, we have the following theorem (Symes, 1986b,Theorem 0.3):

Theorem 1: $\widehat{F}_{p,\omega}$ is a C^1 diffeomorphism: $L^2[0, \frac{t_{max}(p)}{2}] \longrightarrow L^2[0, t_{max}(p)]$.

So $\widetilde{F}_{p,-\delta}$ is a C^1-diffeomorphism of $H^1[0, t_{max}(p)/2]$ into $L^2[0, t_{max}(p)]$.

Hence, given impulsive response $d(p,t)$, $v(p,z)$ and $\tilde{v}(p,\tau)$ are uniquely determined, and $\left(\widetilde{F}_{p,-\delta}\right)^{-1}$ is continuously differentiable.

Remark. Lots of work has shown that inversion for the impedance as a function of vertical travel-time in a layered acoustic medium is well-posed provided that all frequency components down to 0 Hz are available in the data. Here I could review these results about 1-D inversions for the velocity, since the density is assumed to be constant.

Let's write $d(p,t)$ as

$$d(p,t) = d_l(p,t) + d_{l^\perp}(p,t), \tag{2.10}$$

where

$$d_l(p,t) = \int_{|\xi|\le\xi_l} d\xi\, e^{2\pi i\xi t}\eta(p,\xi)$$

and

$$d_{l^\perp}(p,t) = \int_{|\xi|>\xi_l} d\xi\, e^{2\pi i\xi t}\eta(p,\xi),$$

in which

$$\eta(p,\xi) = \int dt\, e^{-2\pi i\xi t} d(p,t),$$

and ξ_l is a positive number (e.g., $\xi_l = 5$).

Then, given $d_{l^\perp}(p,t)$ (or $\eta(p,\xi)$ for $|\xi| > \xi_l$), $v(p,z)$ and $\tilde{v}(p,\tau)$ can be regarded as functions of $d_l(p,t)$ (or $\eta(p,\xi)$ for $|\xi| \le \xi_l$). And $\tilde{v}[d_l]$ (or $\tilde{v}[\eta]$) is continuously differentiable respect to d_l (or $\eta(p,\xi)$ for $|\xi| \le \xi_l$). Generally, the 1D inverse problem is posed as a least squares problem. In practice, due to various physical limitations, $\omega(t)$ and the corresponding reflection seismogram is band-limited. The instability resulting from the absence of high frequency information may be ameliorated by regularization. In contrast, the lack of low-frequency data poses a very serious obstacle to successful inversion. Here I demonstrate this key fact by a numerical experiment. Figure 2.1 shows the amplitude spectra of the two source time functions, of which the

one on the left doesn't have the very low frequency energy and the other one does. Figure 2.2 presents the target velocity profile v (solid line), the starting model v_0 (dot line), and the results v_{inv} got from the OLS inversions corresponding to the two source time functions. The data fitting performance for the two inversions is showed in Figure 2.3. Apparently both estimated models fit the data very well. However, only the inversion with the presence of the very low frequency energy leads to the estimated model with the same velocity trend as the target. The other one is far from the target velocity, i.e., it is a spurious local minimum.

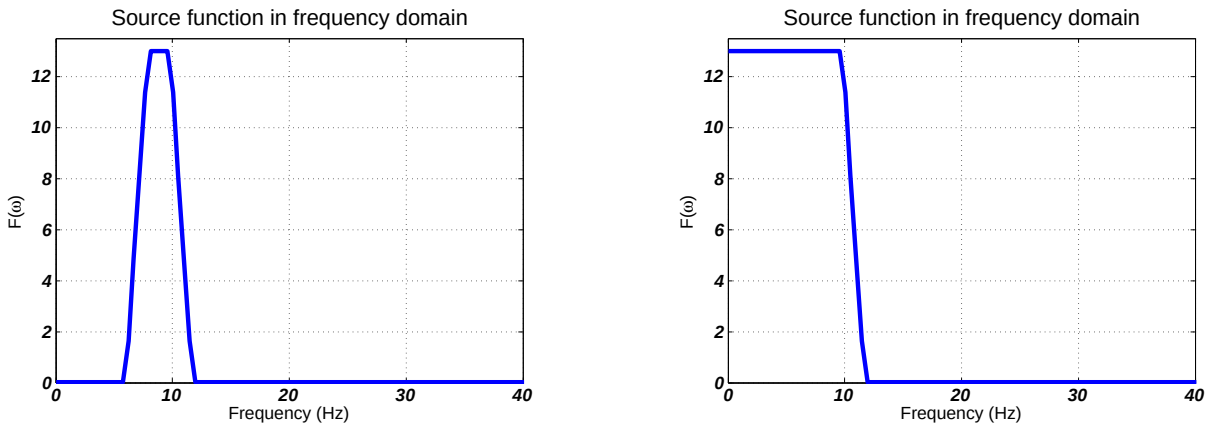

Figure 2.1: Source time functions (in frequency domain): the upper one with the very low-frequency components, the other without those components.

Thus, with the very-low frequency information, a 1D least-squares problem is solvable in the sense that the inversion could recover the long-scale structure of some velocity model. Thus, if given a source time function $\omega(t)$ with the low-frequency components down to 0 Hz, the least-squares inversion does associate a velocity model $v(p,t)$ with a data point $d(p,t)$ which has the very low-frequency components.

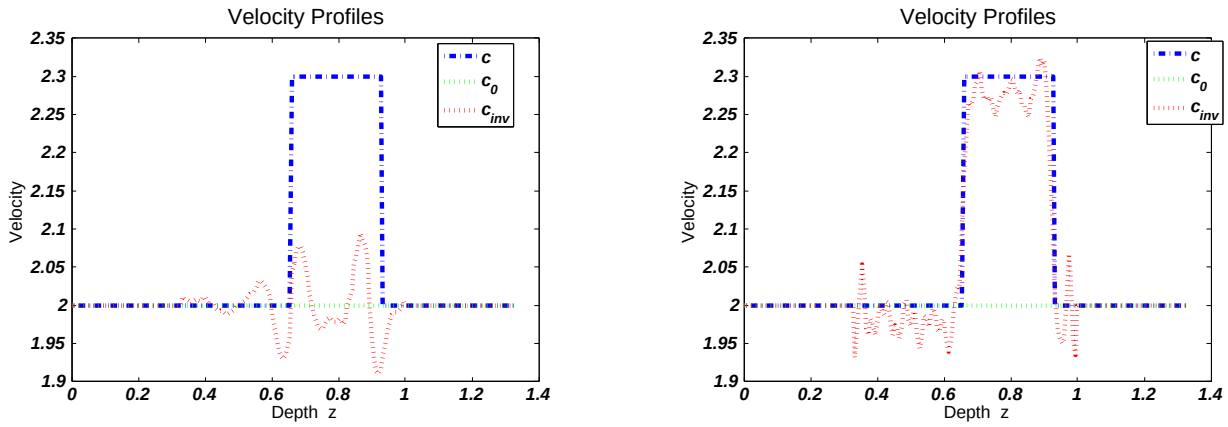

Figure 2.2: Comparison of two velocity profiles (dot line): the one on the left derived from OLS inversion using the source without the very-low frequency components, the other one derived from OLS inversion using the source with low frequency data. (the target velocity profile (dash-dot line), the starting velocity $v_0(z) = 2$)

Separating $d(p,t)$ similarly as (2.10) does,

$$d(p,t) = d_l(p,t) + d_{l^\perp}(p,t),$$

where

$$d_l(p,t) = \int_{|\xi|\le\xi_l} d\xi\, e^{2\pi i\xi t}\eta(p,\xi)$$

and

$$d_{l^\perp}(p,t) = \int_{\xi_l<|\xi|\le\xi_h} d\xi\, e^{2\pi i\xi t}\eta(p,\xi),$$

in which

$$\eta(p,\xi) = \int dt_{|\xi|\le\xi_h}\, e^{-2\pi i\xi t}d(p,t),$$

and $0 < \xi_l < \xi_h$ (e.g., $\xi_l = 5, \xi_h = 60$).

Then, given $d_{l^\perp}(p,t)$ (or $\eta(p,\xi)$ for $_{\xi_l<|\xi|\le\xi_h}$), $v(p,z)$ can be regarded as functions of

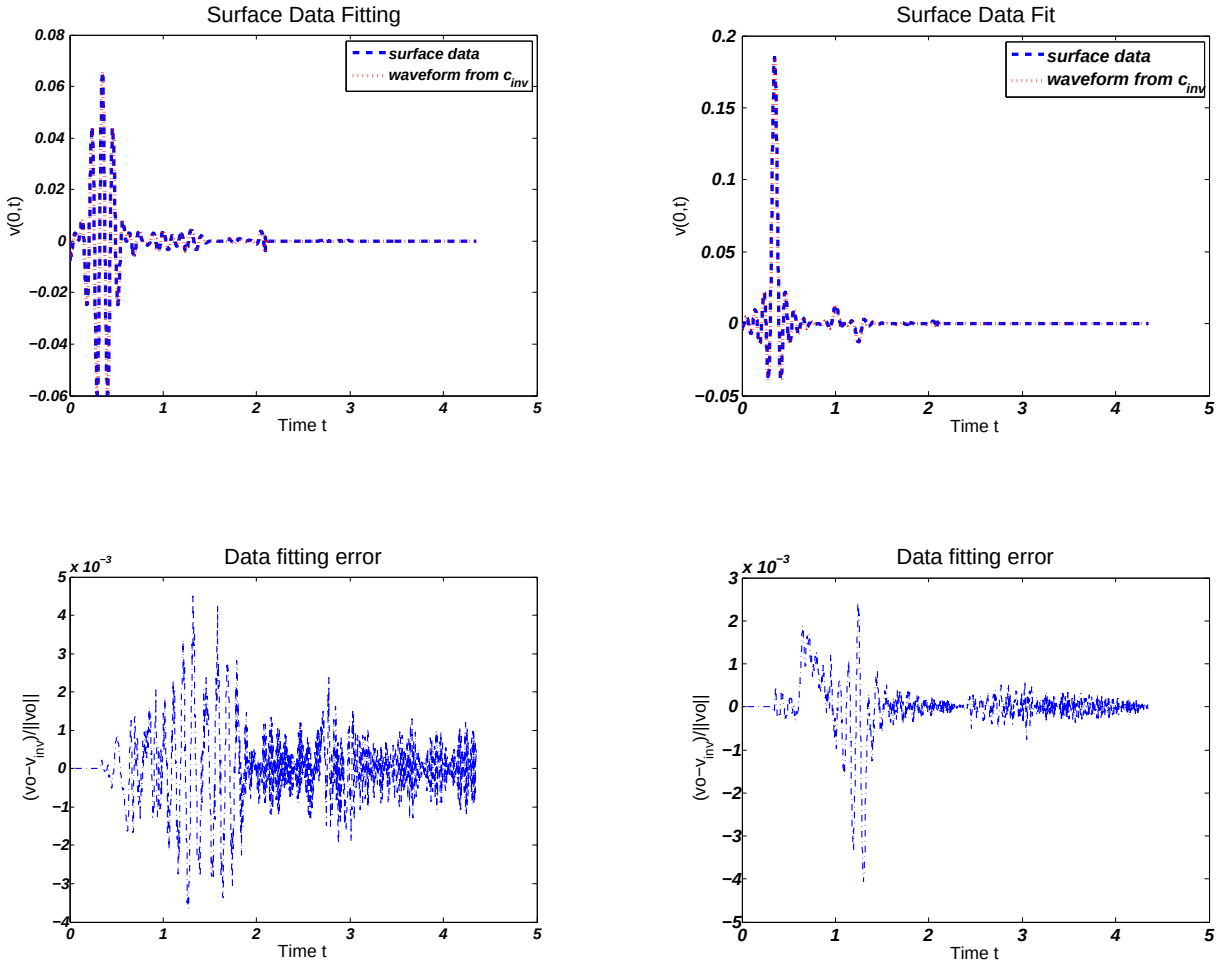

Figure 2.3: Data fitting performance for the two experiments: the left two plots are for the experiment without the very low-frequency energy; the right two is for the experiment with the very low-frequency energy. The upper two plots compare the observed and predicted seismograms. The lower two render the relative data fitting error.

$d_l(p,t)$ (or $\eta(p,\xi)$ for $|\xi| \leq \xi_l$).

Based on the above observation, I will propose a Differential Semblance approach to recover the low-frequency lacuna and release the difficulties associated with the OLS inversion.

2.3 Differential Semblance Approach

In this section, I develop a new approach to the inverse problem (2.4), which falls into the category of differential semblance methods. Differential semblance criterion has been discussed in detail in (Symes and Carazzone, 1991; Symes, 1991a, 1993, 1999). The underlying idea is the concept of semblance of redundant images, i.e., due to the high redundancy of a typical survey, predictions of some model parameters are redundant and unlikely to be consistent (flat in common image panels) unless the velocity model is correct.

Recall Extended Modeling concept in section 2.1 via a simple diagram

$$\begin{array}{ccc} \mathbf{M} & \xrightarrow{S_\omega} & D\,. \\ {\scriptstyle E}\downarrow & \nearrow_{\overline{S}_\omega} & \\ \overline{\mathbf{M}} & & \end{array}$$

A general form of the target inverse problem could be stated as (2.6), i.e.,

$$\begin{aligned} \min_{\bar{c}\in\overline{\mathbf{M}}} \quad & J_A[\bar{c},d] := \frac{1}{2}\|A[\bar{c}]\|^2_{\mathbf{H}} \\ s.t. \quad & \|\overline{S}_\omega - d\|^2_{\mathbf{D}} \simeq 0, \end{aligned}$$

where A is an annihilator.

From the discussion about the solvability of the 1D impulsive inverse problem (see

section 2.2), we have

$$\bar{c}(p,z) = \left(\overline{S}_\delta\right)^{-1}[d_l + d_{l^\perp}].$$

For each fixed p, given the source time function

$$\omega_b(t) = \int_{\xi_l<|\xi|\le\xi_h} d\xi\, e^{2\pi i\xi t} g(\xi)$$

and the corresponding reflection response

$$d_b(p,t) = \int_{\xi_l<|\xi|\le\xi_h} d\xi\, e^{2\pi i\xi t}\eta(p,\xi),$$

take

$$\omega(t) = \omega_l(t) + \omega_b(t),$$

where

$$\omega_l(t) = \int_{|\xi|\le\xi_l} d\xi\, e^{2\pi i\xi t} g(\xi).$$

Then, regarding the source time wavelet $\omega(t)$ with the low-frequency components down to 0 Hz and $d_b(p,t)$ fixed, $v(p,z)$ can be regarded as a function of $d_l(p,t)$ (or $\eta(p,\xi)$ for $|\xi| \le \xi_l$), where

$$d_l(p,t) = \int_{|\xi|\le\xi_l} d\xi\, e^{2\pi i\xi t}\eta(p,\xi). \tag{2.11}$$

Through 1-D OLS inversion, adding $d_l(p,t)$, we could estimate $v(p,z)$ for each p, and then get $\bar{c}(p,z) = v(p,z)\Big/\sqrt{1+v^2(p,z)\,p^2}$. Generally, $\bar{c}(p,z)$ computed through this process will depend on p. But a physically meaningful $\bar{c}(p,z)$ shouldn't depend on p, because of the uniqueness of the subsurface structure! This proposes somewhat a *coherency condition*, that is $\partial\bar{c}\,/\partial p = 0$. (Symes, 1991b)

Hence, in this case, the annihilator is chosen as

$$A[\bar{c}] := \frac{\partial \bar{c}}{\partial p}.$$

I can now state a version of the inverse problem as:

$$\begin{aligned} \min_{\eta(p,\xi),\,(p,\xi)\in\Omega} \quad & J_{DS} := \frac{1}{2}\left\|A[\bar{c}]\right\|^2 \\ s.t. \quad & \|\overline{S}_\omega[\bar{c}](p,t) - d_b(p,t) - d_l[\eta](p,t)\|_{\mathbf{D}} \simeq 0,\ p \in [0, p_{max}] \end{aligned} \tag{2.12}$$

where $A[\bar{c}] := \frac{\partial \bar{c}}{\partial p}$, $\Omega =: \{(p,\xi) : 0 \le p \le p_{max}, |\xi| \le \xi_l\}$.

Remark: To make the computation more efficient and stable, one can choose another $A[\bar{c}]$ with the property that $A[\bar{c}] = 0 \Rightarrow \frac{\partial \bar{c}}{\partial p} = 0$. This issue will be discussed in next chapter.

As for all versions of Newton-related method, smoothness of the objective J_{DS} is essential. Now let's adopt the travel-time trick used by Symes (1991b) to show that J_{DS} is continuously differentiable with respect to the low-frequency components η. Let

$$\tau(z, \bar{c}) := \int_0^z \frac{d\sigma}{\bar{c}(\sigma, p)},$$

and

$$\tilde{\bar{c}}(\tau(z, \bar{c})) := \bar{c}(z).$$

Then

$$\frac{dz}{d\tau} = \tilde{\bar{c}}(\tau(z, \bar{c})),$$

and

$$\begin{aligned} J_{DS} &= \int dp\, dz \left(\frac{\partial \bar{c}}{\partial p}\right)^2 \\ &= \int dp\, d\tau\, \tilde{\bar{c}} \left(\frac{\partial \tilde{\bar{c}}}{\partial p}\right)^2. \end{aligned}$$

The map

$$Q\left[\tilde{\bar{c}}\right] := (\tilde{\bar{c}})^{1/2} \left(\frac{\partial \tilde{\bar{c}}}{\partial p}\right)$$

is continuously differentiable with respect to $\tilde{\bar{c}}$ over $\overline{M}$. Theorem 1 tells us that for each p, $\tilde{\bar{c}}\left[d_l(p,t)\right]$ is continuously differentiable with respect to $d_l(p,t)$ over $Range(\widetilde{F}_{p,\omega})$. And, $d_l(p,t)[\eta(p,\xi)]$ is continuously differentiable with respect to $\eta(p,\xi)$.(see (2.11)) Thus, J_{DS} is continuously differentiable with respect to η.

Note that $\eta(p,\xi)$ must be positive to ensure $d_l(p,t) \in Range(\widetilde{F}_{p,\omega})$. See equation (2.9).

I will sketch the gradient computation for the new objective function in Appendix A. The proposed inverse procedure is summarized in Figure 2.4. This thesis aims to clarify the ideas of this new approach and verify the convexity of the proposed optimization problem through numerical experiments. In next chapter, I will demonstrate via some numerical experiments that this problem is smooth and convex so that one may apply a Newton-like method to solve it.

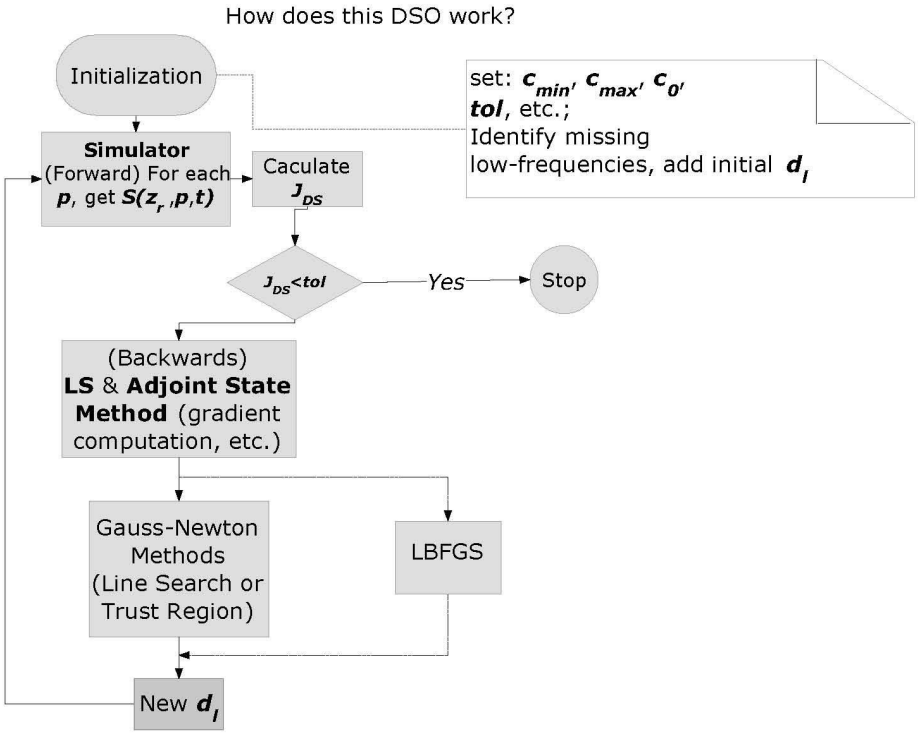

Figure 2.4: Flow chart of the proposed differential semblance approach.

Chapter 3

Numerical Experiments

In this chapter I will conduct some primary numerical experiments ("scan" tests), in which the proposed differential semblance (DS) objective function is evaluated over line segments in the space of low-frequency controls, to display the smoothness and convexity of the proposed objective function. More specifically, I evaluate the DS objective function at some data points D_μ(for some $\mu \in [0, 1.5]$) defined by

$$D_\mu = \{(1-\mu)D_{lpert}(p_i) + \mu D_{obv}(p_i)\}_{i=1}^{N_p}, \tag{3.1}$$

where data $D_{lpert}(p_i)$ at slowness p_i $(i = 1, 2, \cdots, N_p)$ differ from the observed data $D_{obv}(p_i)$ only by their low-frequency components.

3.1 Two-Layer Media

In this section I will present a set of numerical experiments using a two-layer velocity model $c^*(z)$ exhibited in Figure 3.1 to generate the plane wave seismograms of

Figure 3.3 for different slowness

$$\begin{aligned} p \;=\; & (0, 0.0149, 0.0299, 0.0448, 0.0597, 0.0747, 0.0896, 0.1045, 0.1195, 0.1344, 0.1493, \\ & 0.1643, 0.1792, 0.1941, 0.2091, 0.2240, 0.2389, 0.2539, 0.2688, 0.2837, 0.2987, \\ & 0.3136, 0.3285, 0.3435, 0.3584, 0.3733, 0.3883, 0.4032, 0.4181, 0.4330) \end{aligned} \tag{3.2}$$

by numerically solving Equation (2.2) with the absorbing boundary conditions on the surface and bottom, and with the time source function $w(t)$ plotted in Figure 3.2. I use this two-layer model because this simple model presents the most fundamental block embedded in many complicated models.

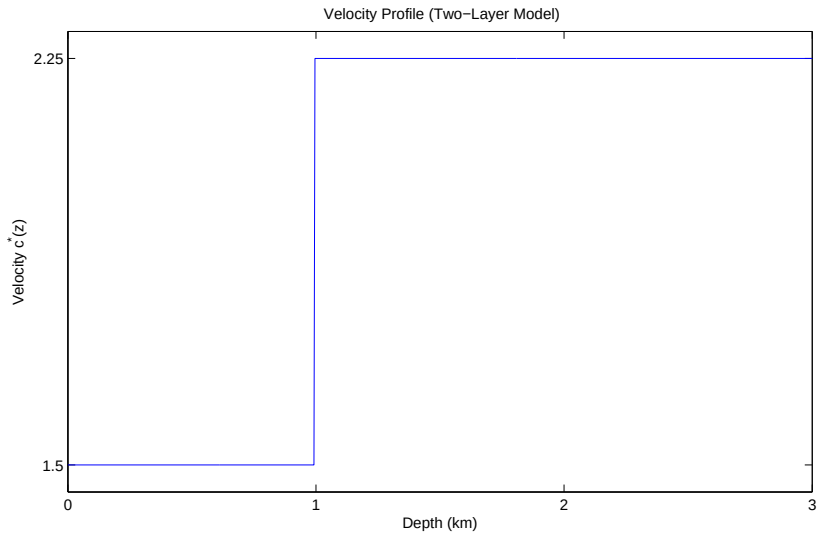

Figure 3.1: Velocity profile (a two-layer model).

For each slowness, I solve the corresponding 1-D least squares problem (**??**) for $v(p, z)$ and then compute $c[p](z)$ from $c[p](z) = \frac{v(p,z)}{\sqrt{1+v^2p^2}}$. More specifically, in each 1-D inversion, the iteration stops when $fobj_{new} \leq 1.e \times 10^{-4} * fobj^0$ or $\|gradient_{new}\| \leq 1.e \times 10^{-4}$ or $fobj_{new} \geq fobj_{prev}$. The resulting velocities $c_{inv}[p](z)$

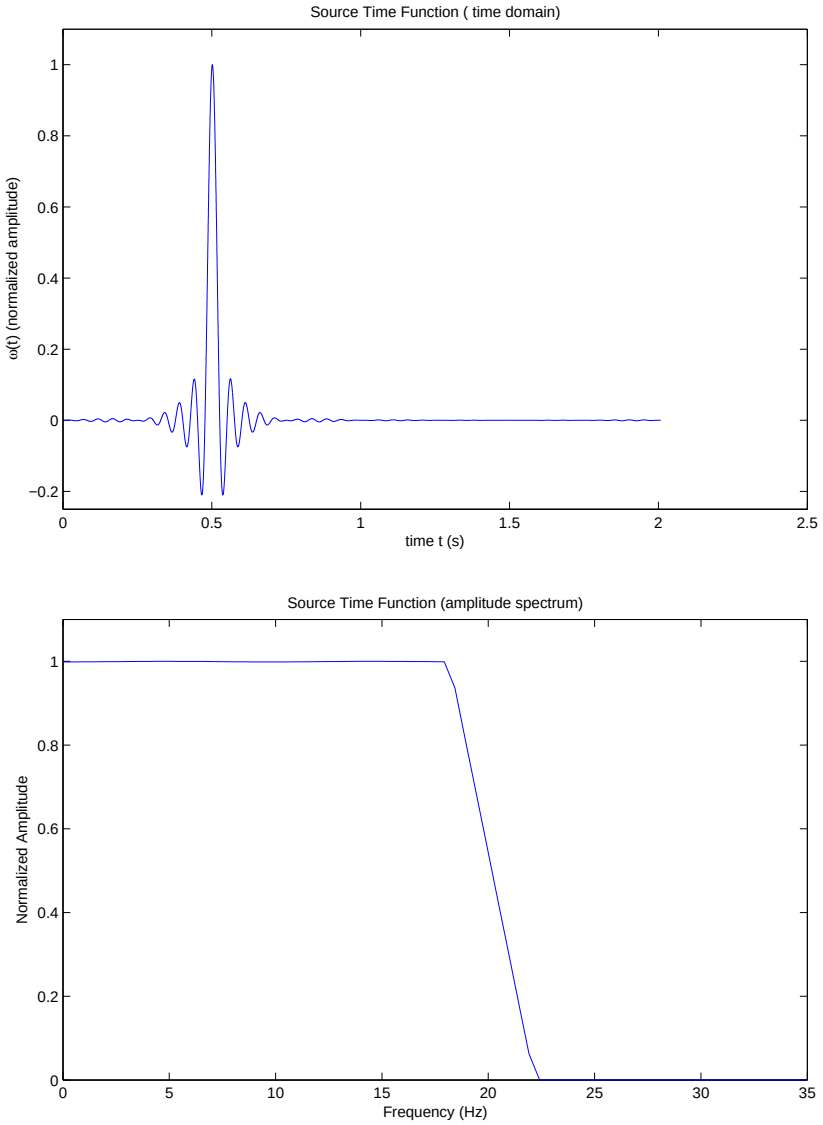

Figure 3.2: Normalized source time function in time domain (upper figure) and in frequency domain (lower figure).

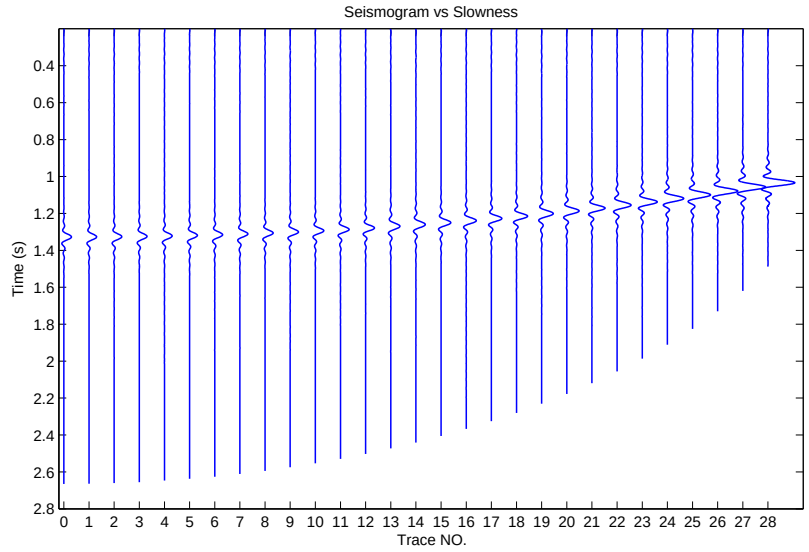

Figure 3.3: Scaled seismograms vs slowness p. The i^{th} trace stands for the seismogram for the i^th slowness (i.e., $p(i+1)$). And $p = (0, 0.0149, \cdots, 0.4330)$ (see Form (3.2)). ($p_{max} < \frac{1}{max(c)} = 0.4444$)

for the synthetic seismogram with true low-frequency components are plotted in Figure 3.4. The Figure 3.5 shows the corresponding data fitting performance for these 1-D OLS inversions.

Figure 3.8 renders a 1-D scan of OLS objective function, which evaluates this objective function at velocities c_μ defined by

$$c_\mu(z) = (1-\mu)c_{hom} + \mu c^*(z)$$

with $\mu = 0.0, 0.1, \cdots, 1.2$. This scan demonstrates once again the multimodality of this objective function, which severely jeopardize the application of Newton-like methods.

As a contrast, the curve in Figure 3.7 interpolates samples of the DS objective

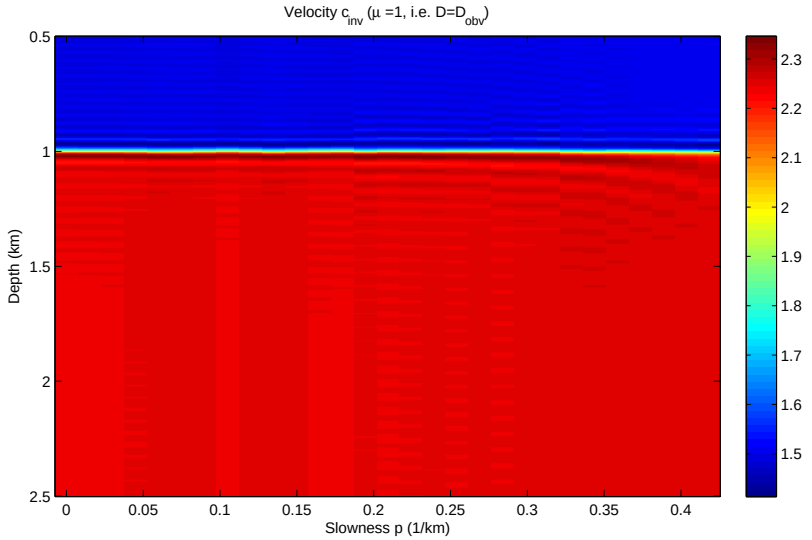

Figure 3.4: Velocities $c_{inv}[p](z)$ derived from 1-D OLS inversions (Four-layer model). And $p = (0, 0.0149, \cdots, 0.4330)$ (see Form (3.2)). $(p_{max} < \frac{1}{max(c)} = 0.4444)$

function at data point D_μ defined by Equation (3.1) with $\mu = 0.0, 0.1, \cdots, 1.2$, p specified in Form (3.2), and the perturbed seismogram $D_{lpert}(p_i)$ at slowness p_i $(i = 1, 2, \cdots, N_p = 29)$, of which the low-frequency components (0 to $5Hz$) are the corresponding low-frequency components of the seismogram derived from the homogeneous velocity model $c_{hom}(z) = 1.84$. Figure 3.6 plots the velocities derived from the 1-D OLS inversions when $\mu = 0$ (i.e., $D_\mu = D_{lpert}$), which clearly presents the bending of the velocity $c_{inv}(p, z)$ with respect to slowness p. This 1-D scan of the DS objective function appears to exhibit the smoothness (at least at the sample scale) and convexity near the minimum. Also, the minimum is achieved at the data point with correct low-frequency components ($\mu = 1$). Therefore, at least if restricted to this 1-D slice, the proposed DSO overcomes the severe convergence difficulties of the OLS inversion, and would recover the velocity model in a few steps of Newton-like methods.

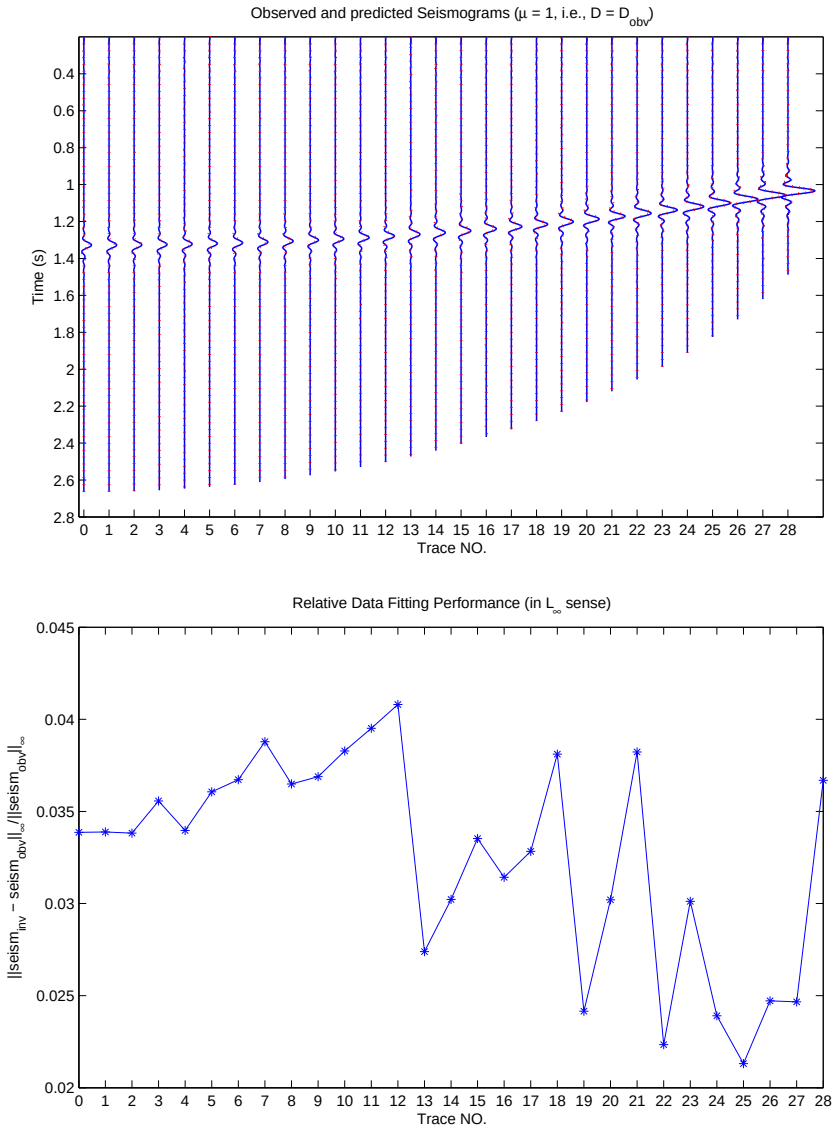

Figure 3.5: Data fitting performance. Upper figure: the observed seismograms (blue solid curves) vs the seismograms based on the velocity models derived from the OLS inversions (red dot curves); Lower figure: relative data fitting errors in l_∞ sense.

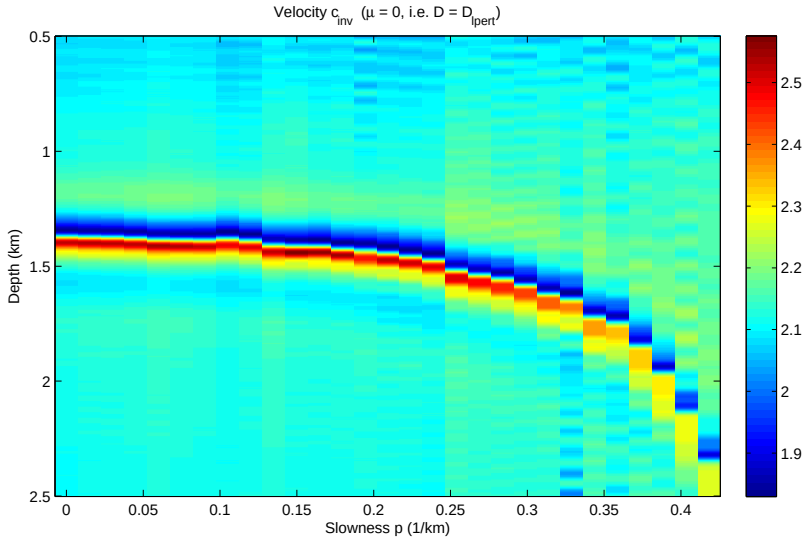

Figure 3.6: Velocity $c_{inv}(p, z)$ at $\mu = 0$ (i.e., $D_\mu = D_{lpert}$). Here $p = (0, 0.0149, \cdots, 0.4330)$ (see Form (3.2)), $D_{lpert}(p)$ has the same low-frequency components (0 to $5Hz$) of the seismogram derived from the homogeneous velocity model $c_{hom}(z) = 1.84$.

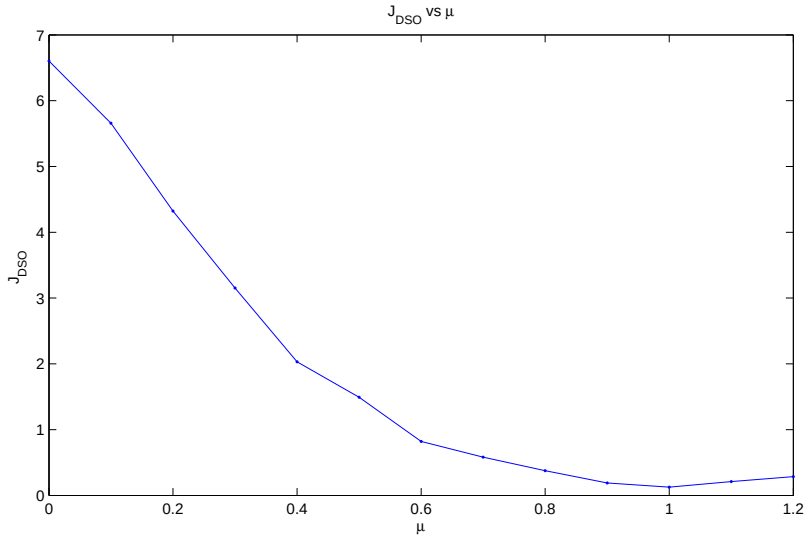

Figure 3.7: The value $J_{DSO}[D_\mu]$ plotted against μ. Here $\mu = 0.0, 0.1, 0.2, \cdots, 1.2$, $p = (0, 0.0149, \cdots, 0.4330)$ (see Form (3.2)), and $D_{lpert}(p)$ has the same low-frequency components (0 to $5Hz$) of the seismogram derived from the homogeneous velocity model $c_{hom}(z) = 1.84$.

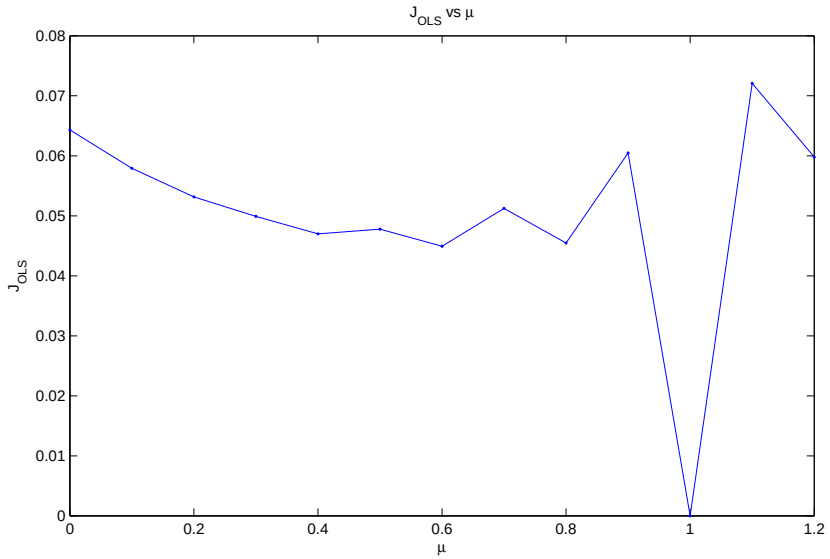

Figure 3.8: The value $J_{OLS}[c_\mu]$ plotted against μ. Here $\mu = 0.0, 0.1, \cdots, 1.2$, $p = (0, 0.0149, \cdots, 0.4330)$ (see Form (3.2)), and $c_\mu(z) = (1-\mu)c_{hom} + \mu c^*(z)$, where $c_{hom}(z) = 1.84$ and $c^*(z)$ is the target velocity model.

3.2 Four-Layer Medium

In this section I will present two groups of "scan" tests using a four-layer velocity model $c^*(z)$ exhibited in Figure 3.9 with different boundary conditions on the surface, i.e., the absorbing boundary and free surface boundary conditions. The following tests use the same time source function $w(t)$ as the one employed in the previous section. The synthetic seismogram D_{obv} with the very low frequency components are generated via numerically solving Equation (2.2) for each slowness with corresponding boundary conditions. In this section, instead of directly partitioning the slowness p evenly, I discretize the slowness p such that p^2 is separated evenly. In such a way, more traces are assigned for larger slowness to achieve better efficiency and performance, because the incorrectness of c_{inv} increases dramatically as slowness becomes larger and closer to the critical value, as can been seen in Figure 3.6.

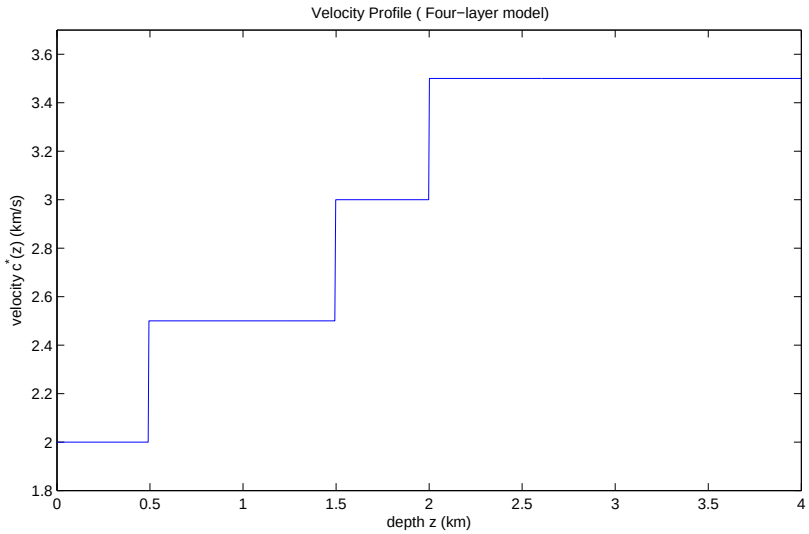

Figure 3.9: Velocity profile (a four-layer model).

Experiments with absorbing boundary on the surface

As in the previous test for two-layer model, the absorbing boundary conditions are employed on the surface and bottom of the four-layer media. To evaluate the differential semblance objective at some data point, I solve for each slowness the corresponding 1-D least squares problem (**??**) to get $v(p,z)$, and then compute $c[p](z)$ from $c[p](z) = \frac{v(p,z)}{\sqrt{1+v^2p^2}}$, and finally compute the differential semblance.

The curve in Figure 3.11 interpolates samples of the DS objective function at data point D_μ defined by Equation (3.1) with $\mu = 0.0, 0.1, 0.2, \cdots, 1.5$,

$$\begin{aligned} p \;=\; & (0, 0.0497, 0.0703, 0.0861, 0.0994, 0.1111, 0.1217, 0.1315, 0.1406, 0.1491, 0.1572, \\ & 0.1648, 0.1722, 0.1792, 0.1860, 0.1925, 0.1988, 0.2049, 0.2109, 0.2166, 0.2223, \\ & 0.2278, 0.2331, 0.2384, 0.2435, 0.2485, 0.2534), \end{aligned}$$

and the perturbed seismogram $D_{lpert}(p_i)$ at slowness p_i $(i = 1, 2, \cdots, 27)$ differs from the observed seismogram $D_{obv}(p_i)$ only by the very low-frequency components (0 to $5Hz$). More specifically, the low-frequency components (0 to $5Hz$) of D_{lpert} are the corresponding low-frequency components of the seismogram derived from the homogeneous velocity model $c_{hom}(z) = 2$. This 1-D slice through the DS objective function appears to exhibit the smoothness (at least at the sample scale) and convexity near the minimum. Also, the minimum is achieved at the data point with correct low-frequency components ($\mu = 1$). Figure 3.10 shows the $c_{inv}(p, z)$ driven from selected data points D_μ $(\mu = 0, 0.4, 0.7, 1.0, 1.2, 1.5)$. As a contrast, Figure 3.12 presents a similar "scan" experiment, which evaluates the OLS objective function at velocities c_μ defined by

$$c_\mu(z) = (1 - \mu)c_{hom} + \mu c^*(z)$$

with $\mu = 0.0, 0.1, 0.2, \cdots, 1.5$. This 1-D scan shows that the minimum is achieved at the desired targets ($\mu = 1$). But it (Figure 3.12) clearly demonstrates the multimodality, that is, this objective function has many spurious local minima, which badly jeopardizes the application of Newton-like methods.

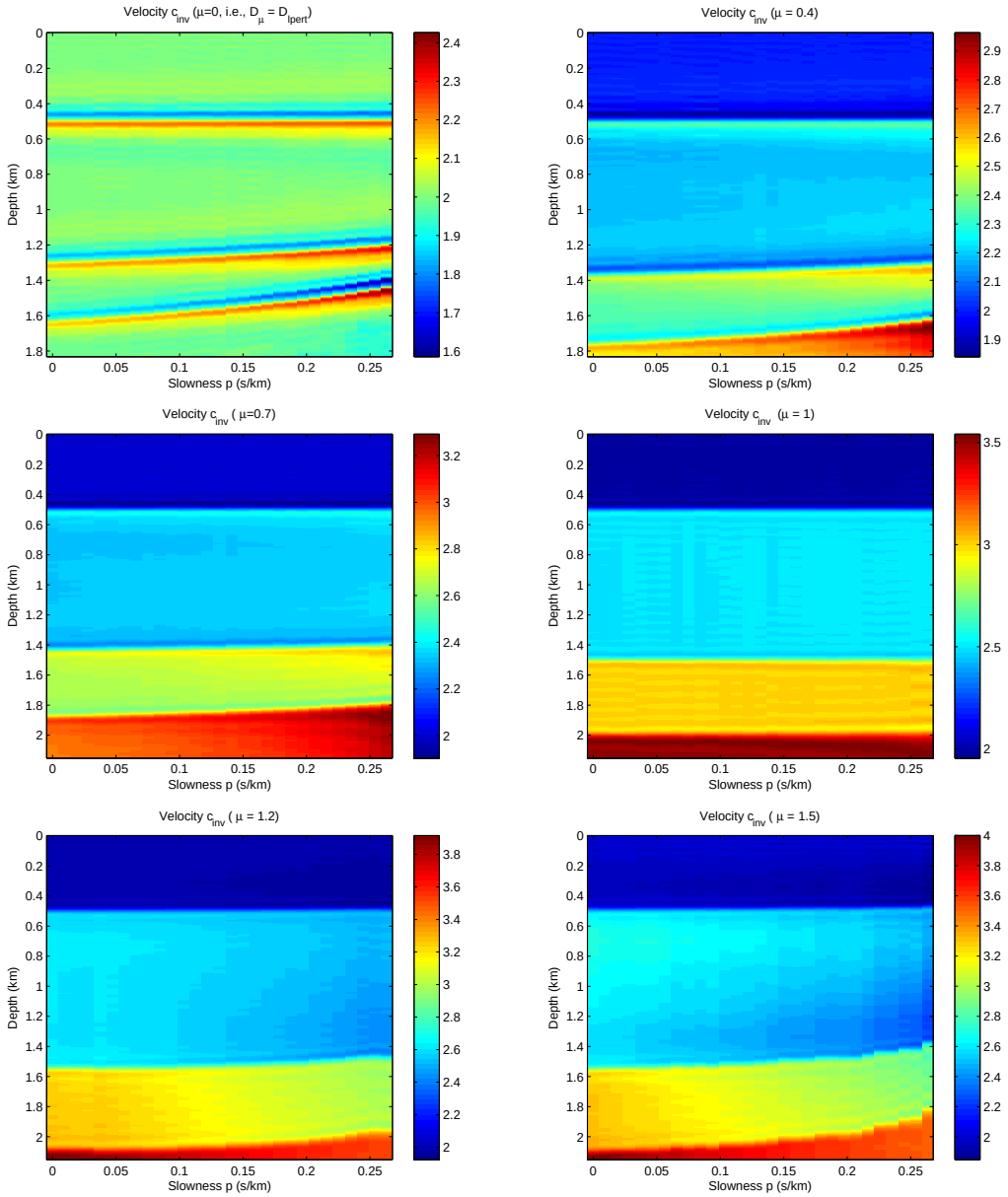

Figure 3.10: Velocities $c_{inv}[p](z)$ derived from 1-D OLS inversions for different data points, i.e., $\mu = 0, 0.4, 0.7, 1.0, 1.2, 1.5$.
And $p = (0, 0.0497, \cdots, 0.2485, 0.2534)$. $(p_{max} < \frac{1}{max(c)} = 0.2857)$

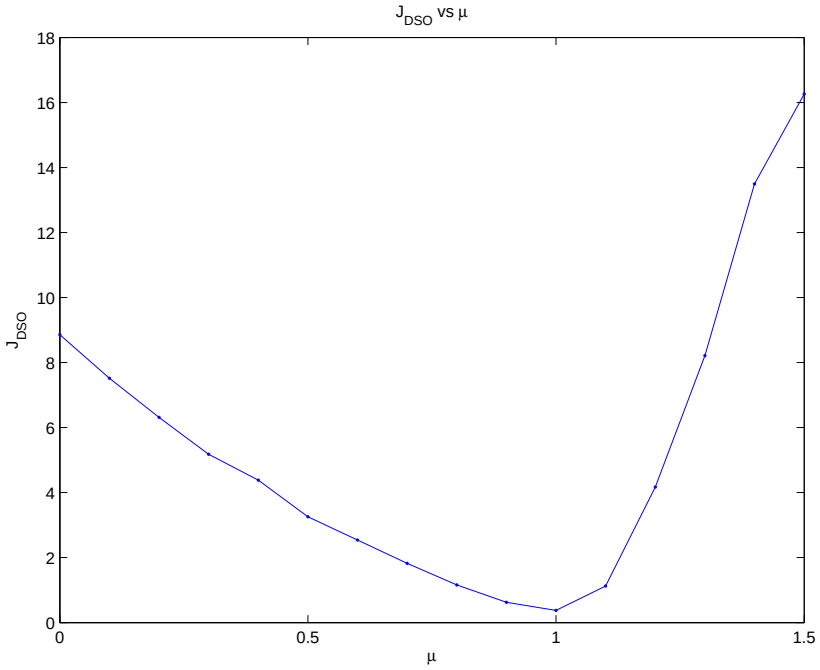

Figure 3.11: The value $J_{OLS}[c_\mu]$ plotted against μ. Here $\mu = 0.0, 0.1, \cdots, 1.5$, $p = (0, 0.0497, \cdots, 0.2485, 0.2534)$, and $D_\mu = (1-\mu)D_{lpert} + \mu D_{obv}$, where D_{lpert} has the same low-frequency components (0 to $5Hz$) of the seismogram derived from the homogeneous velocity model $c_{hom}(z) = 2$.

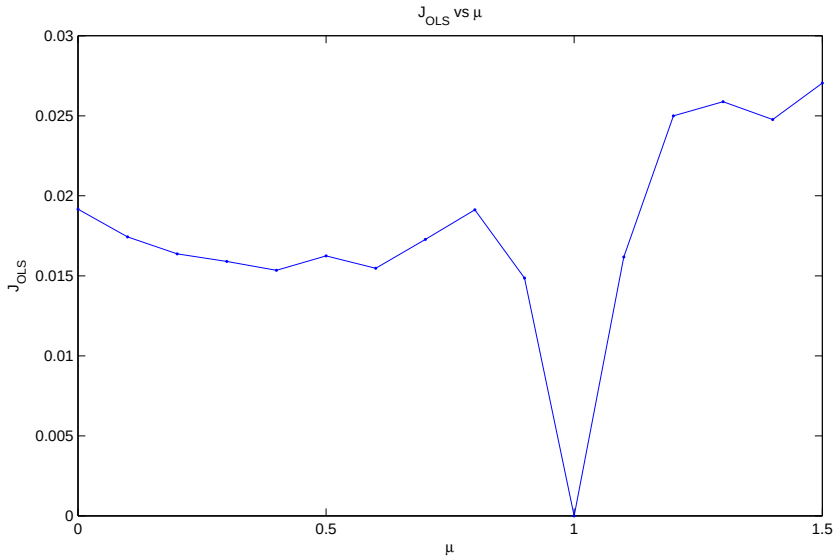

Figure 3.12: The value $J_{OLS}[c_\mu]$ plotted against μ. Here $\mu = 0.0, 0.1, 0.2, \cdots, 1.5$, $p = (0.0, 0.0165, \cdots, 0.1320)$, and $c_\mu(z) = (1-\mu)c_{hom} + \mu c^*(z)$, where $c_{hom}(z) = 2$ and $c^*(z)$ is the target velocity model.

Experiments with free surface boundary condition

An important objective of the proposed algorithm is to account for the nonlinear effects of wave propagation such as multiple reflections. Hence, it is desired to know how this new differential semblance objective behave for problems with free surface, which is an important cause of multiple reflections. In the following tests, the free surface boundary condition is employed.

The curves in Figure 3.14 evaluate the DS objective function at data point D_μ defined by Equation (3.1) with $\mu = 0.0, 0.1, 0.2, \cdots, 1.2$ and $\mu = 0.0, 0.05, 0.1, \cdots, 1.2$

respectively, and with

$$
\begin{aligned}
p \;=\; & (0, 0.0382, 0.0541, 0.0662, 0.0765, 0.0855, 0.0937, 0.1012, 0.1081, 0.1147, 0.1209, \\
& 0.1268, 0.1324, 0.1379, 0.1431, 0.1481, 0.1529, 0.1576, 0.1622, 0.1667, 0.1710, \\
& 0.1752, 0.1793, 0.1834, 0.1873, 0.1912, 0.1950, 0.1987, 0.2023, 0.2059, 0.2094, \\
& 0.2129, 0.2163, 0.2196, 0.2229, 0.2262, 0.2294, 0.2326, 0.2357, 0.2388, 0.2418, \\
& 0.2448, 0.2478, 0.2507, 0.2536, 0.2565, 0.2593, 0.2621, 0.2649, 0.2676),
\end{aligned}
$$

and the perturbed seismogram $D_{lpert}(p_i)$ at slowness p_i $(i = 1, 2, \cdots, 50)$ differs from the observed seismogram $D_{obv}(p_i)$ only by the very low-frequency components (0 to $5Hz$). More specifically, the low-frequency components (0 to $5Hz$) of D_{lpert} are the corresponding low-frequency components of the seismogram derived from the homogeneous velocity model $c_{hom}(z) = 2$. Figure 3.13 shows the $c_{inv}(p, z)$ driven from selected data points D_μ $(\mu = 0, 0.2, 0.4, 0.7, 1, 1.2)$. The two 1-D slices through the DS objective function exhibit the convexity near the minimum. Also, the minimum is achieved at the data point with correct low-frequency components $(\mu = 1)$. But for the finer grids of μ, the lower figure of Figure 3.14 appears to be flat near $\mu = 1$ and possess some bumps. These features could badly jeopardize the application of Newton-like methods. Those small bumps may come from the numerical errors accumulated during all the approximating computations. To improve the behavior of DS objective function, one can apply a number of strategies to decrease numerical errors, such as using smaller tolerance for 1-D inversions, employing finer slowness grids, choosing different expressions of differential semblance objective and adopting some regularization techniques to smooth $c_{inv}(p, z)$ with respect to p, etc.. I consider some of these improving strategies in the rest of this section.

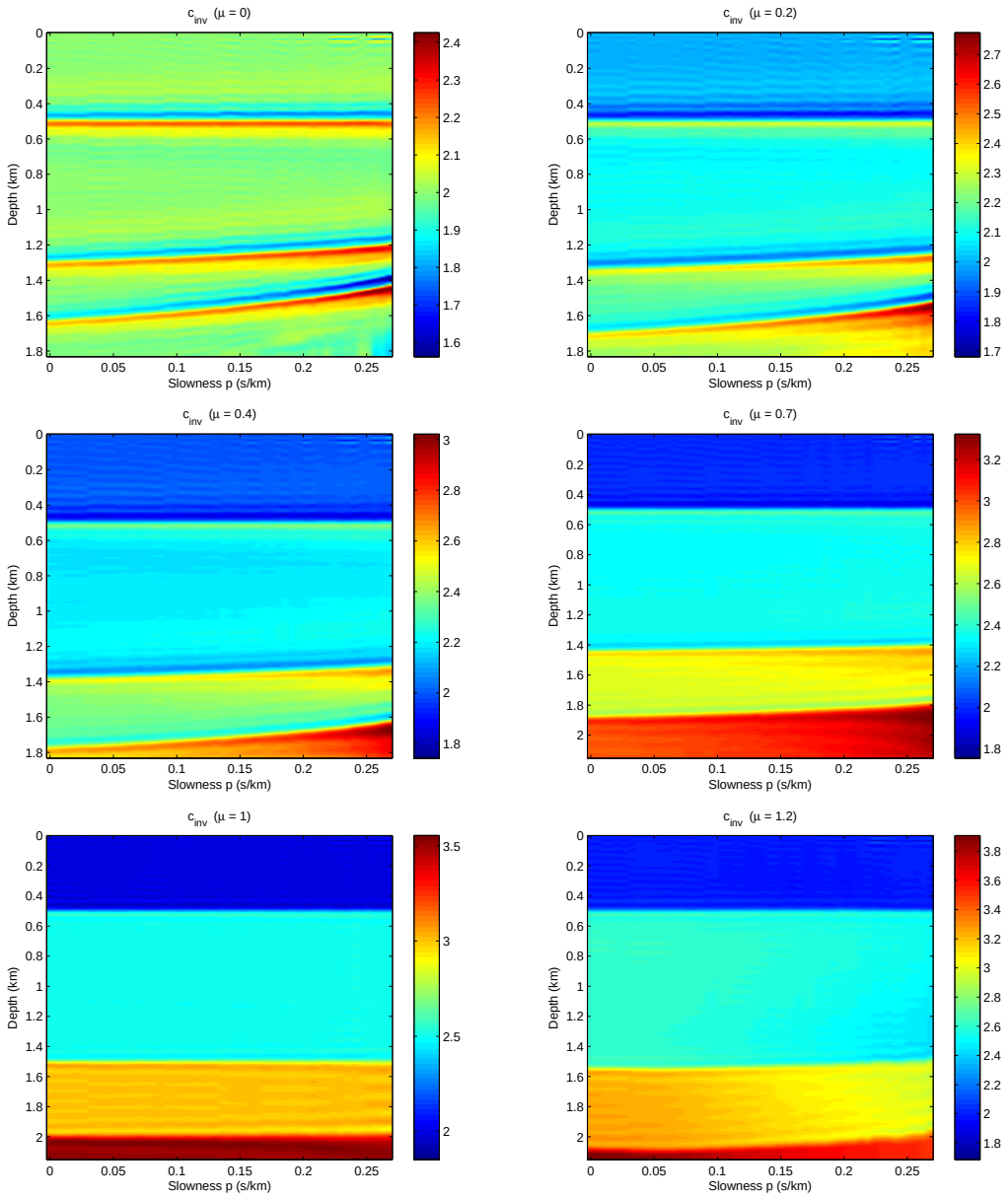

Figure 3.13: Velocities $c_{inv}[p](z)$ derived from 1-D OLS inversions for different data points, i.e., $\mu = 0, 0.2, 0.4, 0.7, 1.0, 1.2$.
And $p = (0, 0.0382, \cdots, 0.2649, 0.2676)$. $(p_{max} < \frac{1}{max(c)} = 0.2857)$

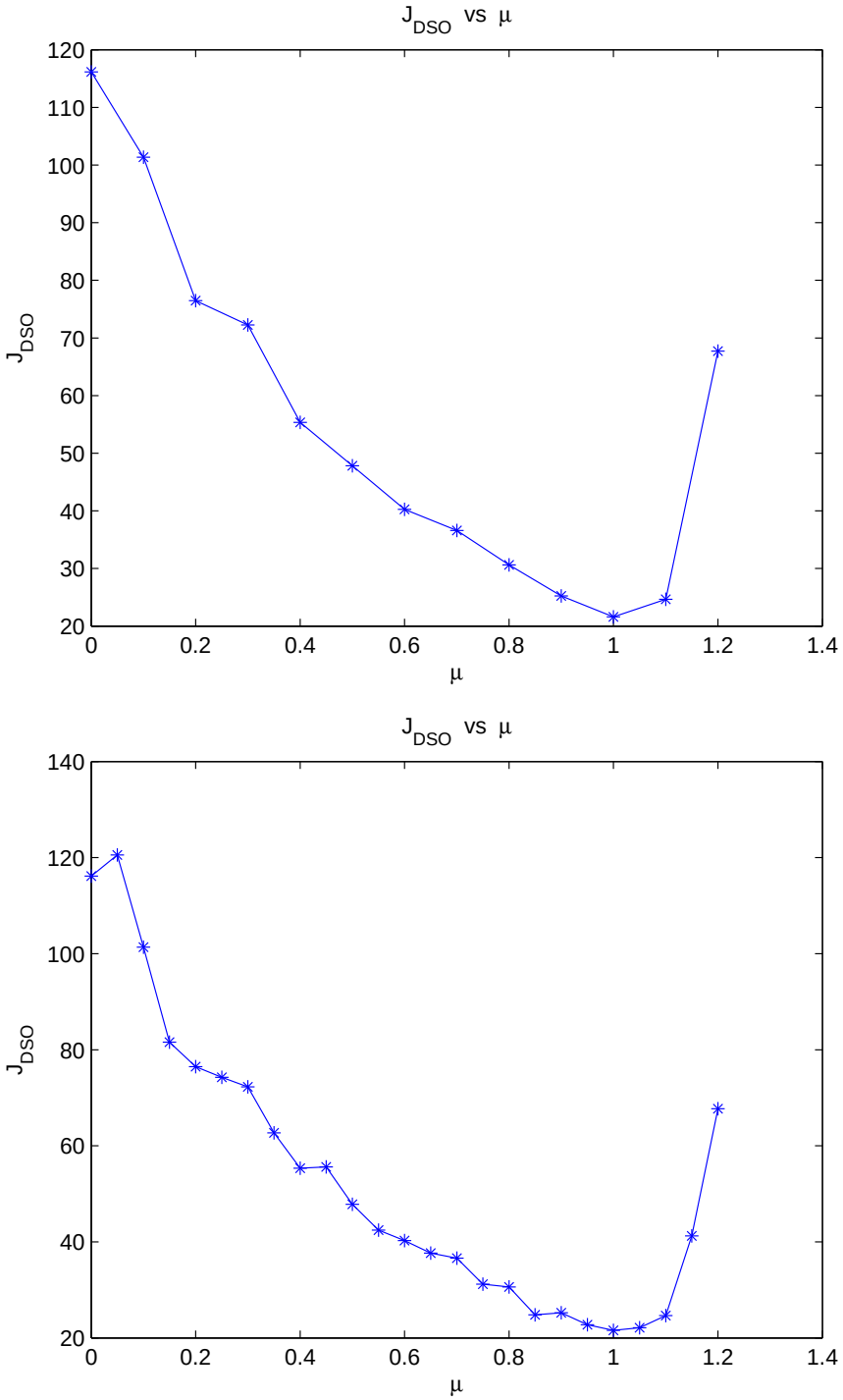

Figure 3.14: The value $J_{OLS}[c_\mu]$ plotted against μ. Here $\mu = 0.0, 0.1, \cdots, 1.2$ for the upper figure, $\mu = 0.0, 0.05, 0.1 \cdots, 1.2$ for the lower figure, and $p = (0, 0.0382, \cdots, 0.2649, 0.2676)$, and $D_\mu = (1-\mu)D_{lpert} + \mu D_{obv}$, where D_{lpert} has the same low-frequency components (0 to $5Hz$) of the seismogram derived from the homogeneous velocity model $c_{hom}(z) = 2$.

Smoothing $c_{inv}(p,z)$ with respect to p

Figure 3.16 presents a scan of the DS objective function through the data points D_μ defined by Equation (3.1) with $\mu = 0.0, 0.1, 0.2, \cdots, 1.2$ and $\mu = 0.0, 0.05, 0.1, \cdots, 1.2$ respectively. All the settings are the same as the previous scan test for the four-layer media with free surface except that $c_{inv}(p,z)$ is smoothed with respect to p via minimizing the Total Variation of $c_{inv}(p,z)$ with respect to p for each z. Figure 3.17 shows the $c_{inv}(p,z)$ driven from selected data points D_μ ($\mu = 0, 0.2, 0.4, 0.7, 1, 1.2$). These two 1-D slices through the DS objective function exhibit the smoothness (at least at the sample scale) and convexity. Also, the minimum is achieved at the data point with correct low-frequency components ($\mu = 1$).

Take DS objective as $J_{DSO} := \frac{1}{2}\|\frac{\partial(1/c)}{\partial p}\|^2$

Given the same settings as the previous scan test for the four-layer free surface media with smoothing c_{inv} with respect to p, take $J_{DSO} = \frac{1}{2}\|\frac{\partial(1/c)}{\partial p}\|^2$. Figure 3.17 shows the $\frac{1}{c_{inv}(p,z)}$ for selected data points D_μ ($\mu = 0, 0.2, 0.4, 0.7, 1, 1.2$). Figure 3.18 presents the corresponding scans of this objective function, which exhibit convexity and better smoothness than the previous test, though the bottoms of them are nearly flat. Also, the minimum is achieved $\mu = 1$.

Remark I confess that the examples so far do not show clearly any influence of nonlinearity, and that it is needed to create some other examples with more reflectors in them.

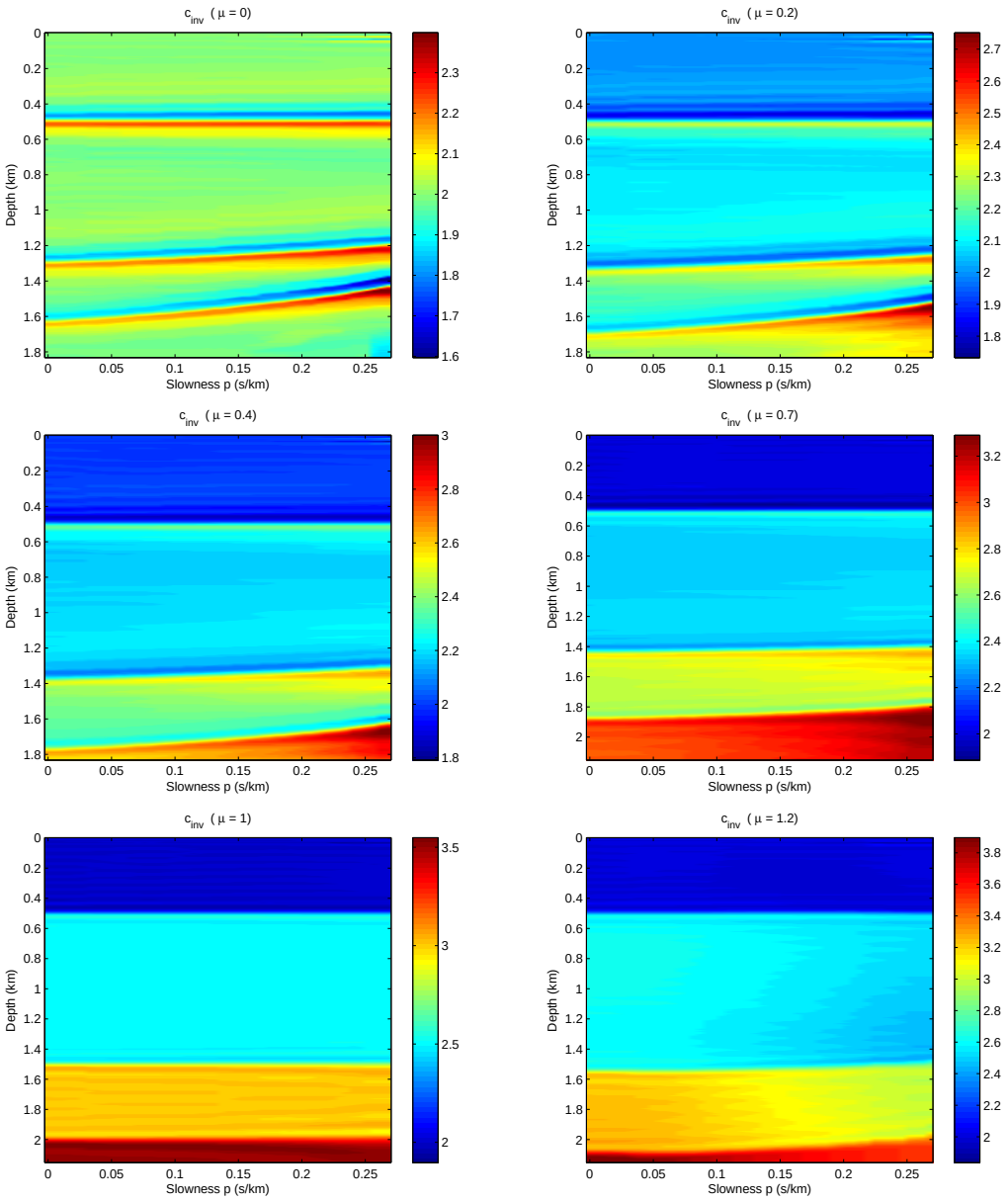

Figure 3.15: Velocities $c_{inv}[p](z)$ derived from 1-D OLS inversions for different data points after smoothing in p. Here $\mu = 0, 0.2, 0.4, 0.7, 1.0, 1.2$.
And $p = (0, 0.0382, \cdots, 0.2649, 0.2676)$. $(p_{max} < \frac{1}{max(c)} = 0.2857)$

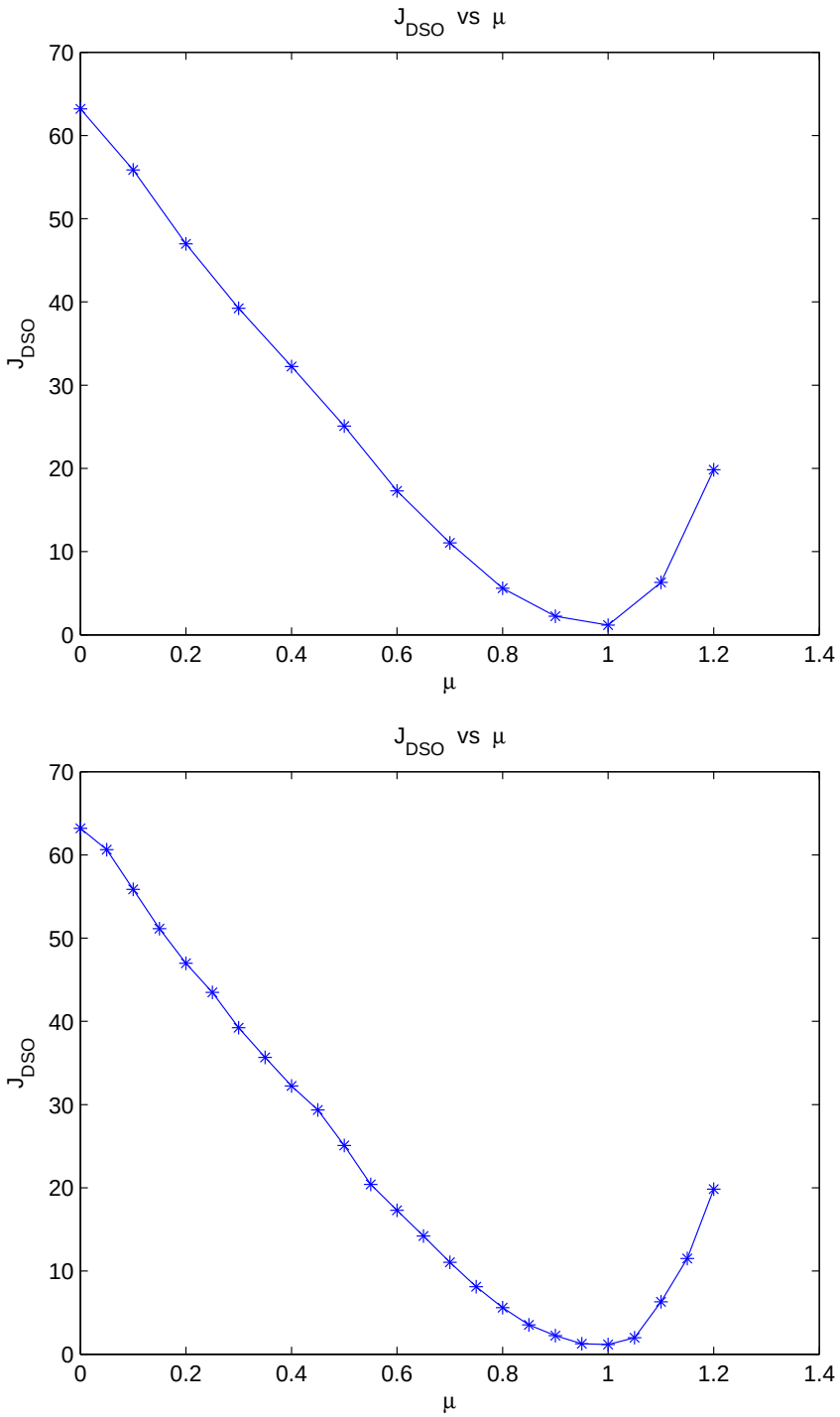

Figure 3.16: The value $J_{OLS}[c_\mu]$ plotted against μ. Here $\mu = 0.0, 0.1, \cdots, 1.2$ for the upper figure, $\mu = 0.0, 0.05, 0.1 \cdots, 1.2$ for the lower figure, and $p = (0, 0.0382, \cdots, 0.2649, 0.2676)$, and $D_\mu = (1-\mu)D_{lpert} + \mu D_{obv}$, where D_{lpert} has the same low-frequency components (0 to $5Hz$) of the seismogram derived from the homogeneous velocity model $c_{hom}(z) = 2$.

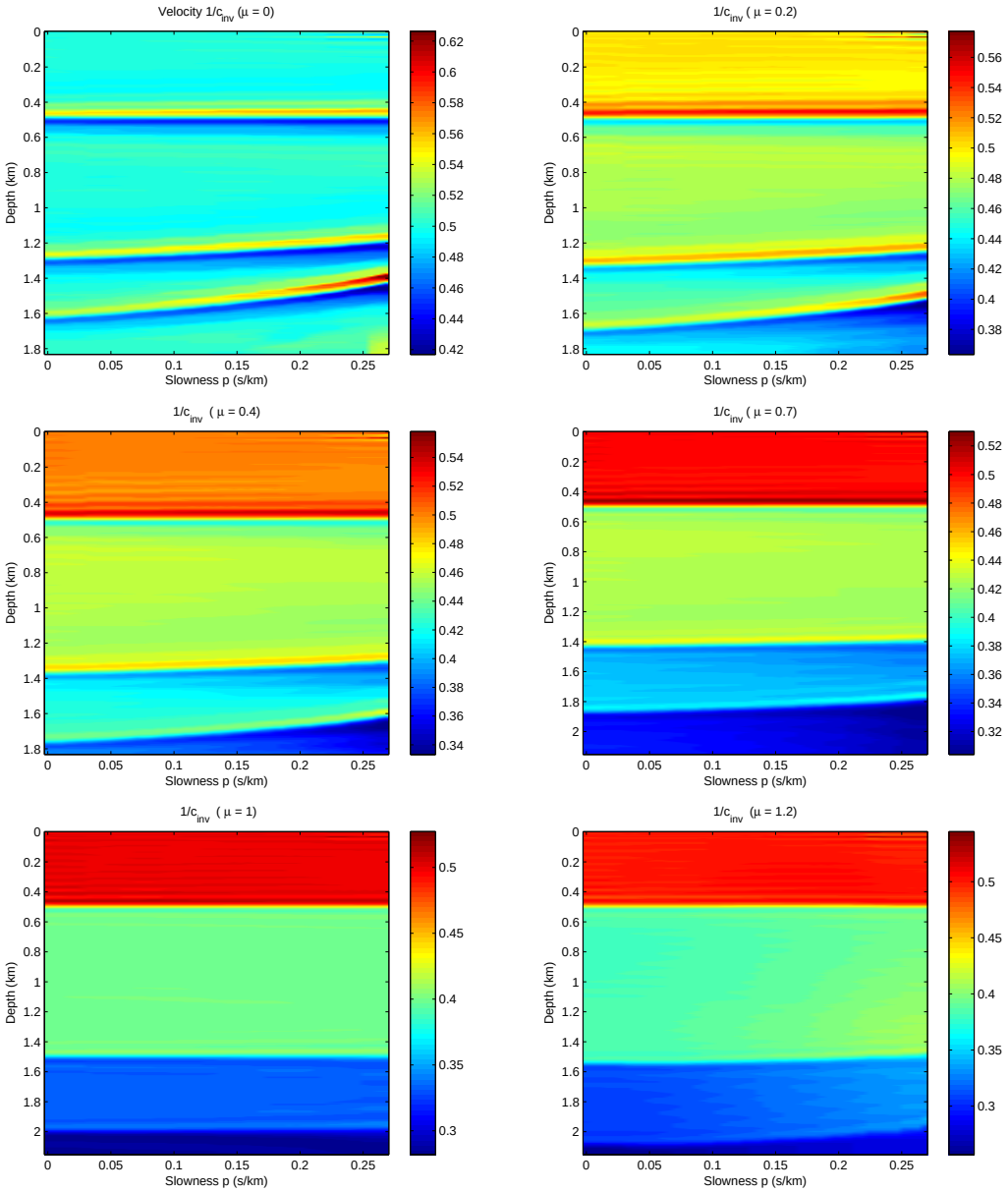

Figure 3.17: $\frac{1}{c_{inv}}$ derived from 1-D OLS inversions for different data points after smoothing in p. Here $\mu = 0, 0.2, 0.4, 0.7, 1.0, 1.2$.
And $p = (0, 0.0382, \cdots, 0.2649, 0.2676)$. ($p_{max} < \frac{1}{max(c)} = 0.2857$)

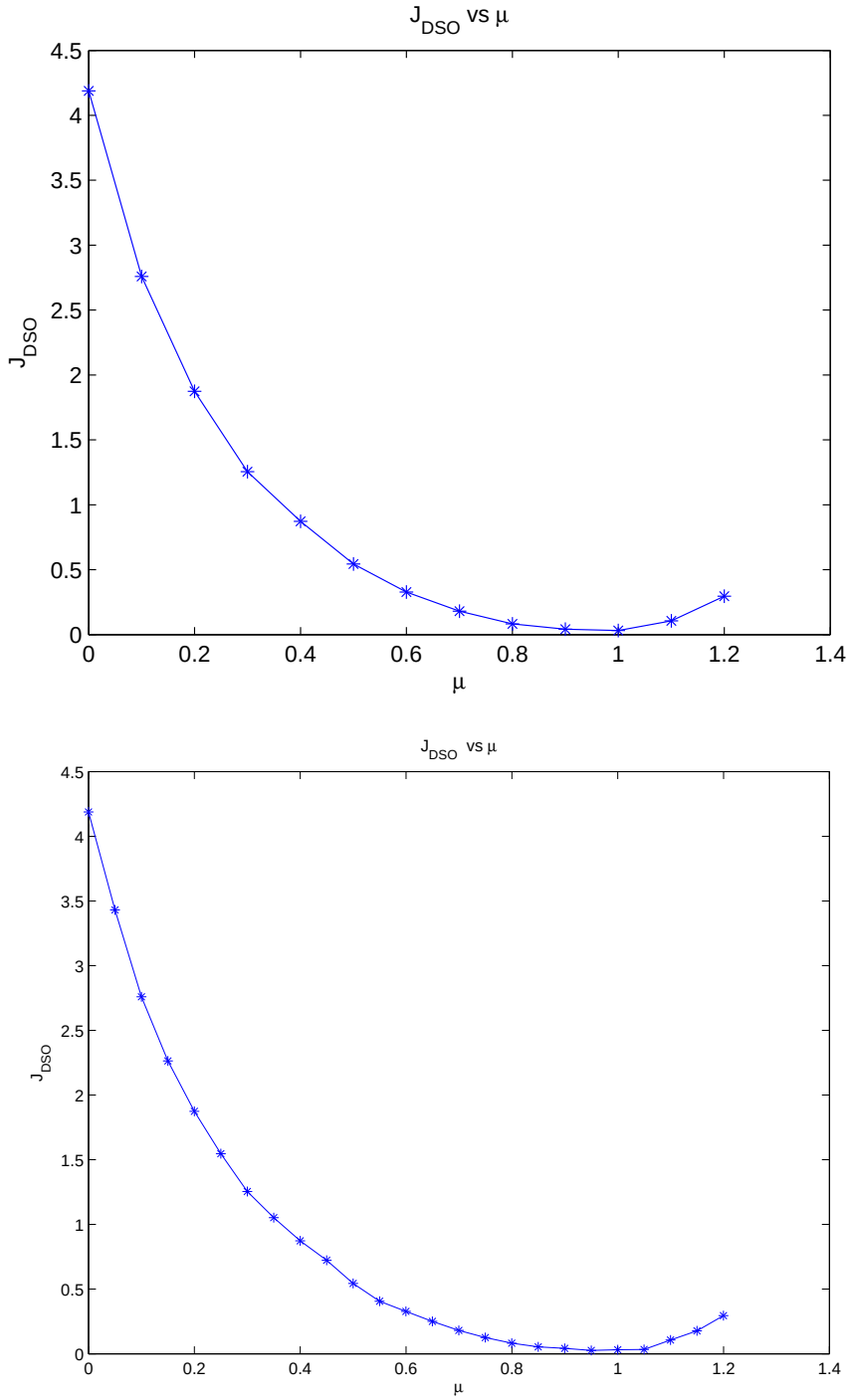

Figure 3.18: The value $J_{OLS}[c_\mu] = \frac{1}{2}\|\frac{\partial(1/c)}{\partial p}\|^2$ plotted against μ. Here $\mu = 0.0, 0.1, \cdots, 1.2$ for the upper figure, $\mu = 0.0, 0.05, 0.1 \cdots, 1.2$ for the lower figure, and $p = (0, 0.0382, \cdots, 0.2649, 0.2676)$, and $D_\mu = (1-\mu)D_{lpert} + \mu D_{obv}$, where D_{lpert} has the same low-frequency components (0 to $5Hz$) of the seismogram derived from the homogeneous velocity model $c_{hom}(z) = 2$.

Chapter 4

Discussion and Conclusion

From the numerical experiments discussed in Chapter 3, I draw the following conclusions:

First, for synthetic data, the target model is amongst the minima of J_{DSO}. Thus, the velocity distribution obtained by minimization of J_{DSO} is consistent with the target model.

Second, 1-D slices of the DS objective function are sort of smooth and convex, provided large enough ranges of slowness; while the corresponding scans of the OLS objective function demonstrate strong multi-modality.

Third, Newton-like methods seem to provide a promising approach to solve the proposed DSO problem, because of the smoothness and convexity demonstrated by numerical experiments. Implementation of such a Newton-like method will be the next work.

The above three points make the proposed method a promising alternative to wave-

form inversion. There are still some important concerns about the proposed approach:

First, DS is based on precritical slowness, i.e., $c|p| < 1$. For better performance, the construction of DS objective needs as large range of slowness as possible. However, the very depth of "precritical" depends on the model, i.e., on c, which is to be sought. Thus, a prediction about the large enough slowness, needed to construct the proposed DSO problem, is very important. I discuss such a prediction for two-layer models in Appendix B. What's more, to balance the efficiency and reliability of this DSO approach, a strategy to adaptively determine the range of slowness appears desirable, which will be a future work.

Second, one should adopt some regularization strategy in order to get a more reliable and smooth DS objective function. Additionally, lots of adjustments could help decrease numerical errors further, such as adjusting the stopping criteria, optimizing regularization parameters, refining the slowness grid (e.g., evenly partition the squares of slowness), etc.

As a summary, I have described a new DSO approach for velocity estimation. I have demonstrated by some numerical experiments how this new objective function avoids the non-convexity of least-squares forms. Also, the calculation of gradient and the prediction of large enough slowness (for two-layer media) are addressed in the appendix. This promising approach avoids the severe convergence difficulties associated with the classical output least squares (OLS) seismic inversion, and would lead to reasonably accurate result from a coarse initial guess for typical band-limited data.

Appendix A

Calculation of Gradient via Adjoint State Method

Recall that the proposed DS optimization problem is

$$\begin{aligned} \min_{\eta(p,\xi),\,(p,\xi)\in\Omega} \quad & J_{DS} := \frac{1}{2}\left\|A[\bar{c}]\right\|^2 \\ s.t. \quad & \bar{c} = \text{Argmin}\,\frac{1}{2}\left\|\overline{S}_\omega[\bar{c}](p,t) - d_b(p,t) - d_l[\eta](p,t)\right\|_{\mathbf{D}}, \text{ for } p\in[0,p_{max}] \end{aligned} \tag{A-1}$$

where $A[\bar{c}] := \frac{\partial \bar{c}}{\partial p}$, $\Omega =: \{(p,\xi) : 0 \le p \le p_{max}, |\xi| \le \xi_l\}$, p is slowness, d_l are the artificial low-frequency controls that make up the missing very low-frequency data.

Now let's compute the gradient $\nabla_\eta J_{DS}$. Assume all the derivatives in the following computation exist. The computation consists of three steps.

Step 1. Compute $\delta_\eta d_l$

Recall that

$$d_l(p,t) = \int_{|\xi|\le\xi_l} d\xi\, e^{2\pi i\xi t}\eta(p,\xi).$$

Applying regular perturbation to the above equation, we have

$$\delta d_l(p,t) \approx \int_{|\xi|\le\xi_l} d\xi\, e^{2\pi i\xi t}\delta\eta(p,\xi) = Y\delta\eta(p,\xi) \tag{A-2}$$

Step 2. Compute $\delta_{d_l}\bar{c}$

Let

$$E[\bar{c},d_l] := \frac{1}{2}\left\|\overline{S}_\omega[\bar{c}](p,t) - d_b(p,t) - d_l[\eta](p,t)\right\|_{\mathbf{D}}.$$

The first order development of E gives:

$$\begin{aligned}\delta E &= \frac{1}{2}\left\{\left\|\overline{S}_\omega[\bar{c}+\delta\bar{c}]-d_b-d_l\right\|_{\mathbf{D}}^2-\left\|\overline{S}_\omega[\bar{c}]-d_b-d_l\right\|_{\mathbf{D}}^2\right\}\\ &\approx \left\langle\overline{S}_\omega[\bar{c}]-d_b-d_l, D\overline{S}_\omega[\bar{c}]\delta\bar{c}\right\rangle_{\mathbf{D}}\\ &= \left\langle D\overline{S}_\omega[\bar{c}]^T\left(\overline{S}_\omega[\bar{c}]-d_b-d_l\right), \delta\bar{c}\right\rangle_{\mathbf{M}}.\end{aligned}$$

The first order necessity of the least-squares subproblem gives:

$$D\overline{S}_\omega[\bar{c}]^T\left(\overline{S}_\omega[\bar{c}]-d_b-d_l\right)=0.$$

Applying regular perturbation to the above equation, we have

$$D\overline{S}_\omega[\bar{c}]^T\left(D\overline{S}_\omega[\bar{c}]\delta\bar{c}--\delta d_l\right)\approx 0,$$

i.e.,

$$D\overline{S}_\omega[\bar{c}]^T D\overline{S}_\omega[\bar{c}]\delta\bar{c}\approx D\overline{S}_\omega[\bar{c}]^T\delta d_l.$$

Thus,

$$D_{d_l}\bar{c}=\left(D\overline{S}_\omega[\bar{c}]^T D\overline{S}_\omega[\bar{c}]\right)^\dagger D\overline{S}_\omega[\bar{c}]^T \tag{A-3}$$

$$\delta_{d_l}\bar{c}=D_{d_l}\bar{c}\delta d_l \tag{A-4}$$

Step 3. Compute $\delta_{\bar{c}}J_{DS}$

The first order development of J_{DS} gives:

$$\begin{aligned}
\delta_{\bar{c}} J_{DS} &= \frac{1}{2}\left\{\left\|\frac{\partial \bar{c}+\delta\bar{c}}{\partial p}\right\|_{\mathbf{M}}^2 - \left\|\frac{\partial \bar{c}}{\partial p}\right\|_{\mathbf{M}}^2\right\} \\
&\approx \left\langle \frac{\partial \bar{c}}{\partial p}, \frac{\partial \delta\bar{c}}{\partial p}\right\rangle_{\mathbf{M}} \\
&= -\left\langle \frac{\partial^2 \bar{c}}{\partial p^2}, \delta\bar{c}\right\rangle_{\mathbf{M}} + \int dz \left(\frac{\partial \bar{c}}{\partial p}\delta\bar{c}\right)\Big|_0^{p_{max}} \\
&= -\left\langle \frac{\partial^2 \bar{c}}{\partial p^2}, \delta\bar{c}\right\rangle_{\mathbf{M}}.
\end{aligned}$$

Thus,

$$\begin{aligned}
\delta_\eta J_{DS} &= -\left\langle \frac{\partial^2 \bar{c}}{\partial p^2}, \delta\bar{c}\right\rangle_{\mathbf{M}} \\
&= -\left\langle \frac{\partial^2 \bar{c}}{\partial p^2}, D_{d_l}\bar{c}\delta d_l\right\rangle_{\mathbf{M}} \\
&= -\left\langle \left(D_{d_l}\bar{c}\right)^T \frac{\partial^2 \bar{c}}{\partial p^2}, \delta d_l\right\rangle_{\mathbf{D}} \\
&= -\left\langle \left(D_{d_l}\bar{c}\right)^T \frac{\partial^2 \bar{c}}{\partial p^2}, Y\delta\eta\right\rangle_{\mathbf{D}} \\
&= -\int_{|\xi|\le\xi_l} d\xi\, Y^T \left(D_{d_l}\bar{c}\right)^T \frac{\partial^2 \bar{c}}{\partial p^2}\delta\eta
\end{aligned}$$

Since

$$\delta_\eta J_{DS} = \int_{|\xi|\le\xi_l} d\xi\, \nabla_\eta J_{DS}\delta\eta,$$

$$\nabla_\eta J_{DS} = Y^T \left(D_{d_l}\bar{c}\right)^T \frac{\partial^2 \bar{c}}{\partial p^2}. \tag{A-5}$$

Appendix B

Prediction of the large enough slowness for two-layer model

A two layer model is:

$$c_r(z) = \begin{cases} c_* & \text{for} \quad 0 \le z < z_*, \\ c_2 & \text{for} \quad z \ge z_*. \end{cases}$$

In the inverse problem, we try to approximate the reference velocity c_* by a trial velocity c_t which is updated in every step. For the reference velocity c_*, the vertical velocity is

$$v_*(p) = \frac{c_*}{\sqrt{1-(p\,c_*)^2}} \ ,$$

where slowness $|p| \le p^*_{max}$ ($|p^*_{max}\, c_r(z)| < 1$). Introducing a mechanical vibration at the point $z = 0$, one expects to receive a response corresponding to the velocity discontinuity located at z_*. The corresponding two way travel time $\tau(p)$ of this response is:

$$\tau(p) = \frac{2z_*}{v_*(p)} \ .$$

In the inverse process, the vibration is considered to travel at the speed $v_t(p)$, and then the discontinuity seems to be located at the depth $z_t(p)$ computed by:

$$z_t(p) = \frac{1}{2} v_t(p)\, \tau(p) \ ,$$

where $v_t(p) = \frac{c_t}{\sqrt{1-(p\,c_t)^2}}$, in which $|p| \le p^t_{max}$ ($|p^t_{max}\, c_t(z)| < 1$).
Thus,

$$\frac{z_t(p)}{z_*(p)} = \frac{v_t(p)}{v_*(p)} = \frac{c_t\sqrt{1-(p\,c_*)^2}}{c_*\sqrt{1-(p\,c_t)^2}}.$$

Let $a = \frac{c_t}{c_*}$, then

$$z_t(p) - z_*(p) = \left(a \frac{\sqrt{1-(p\,c_*)^2}}{\sqrt{1-(a\,p\,c_*)^2}} - 1 \right) z_* \qquad \text{for all } p \in P, \tag{B-1}$$

where $P := \{p : |p| \le p_{max}\}$, and $p_{max} := \min\{p^*_{max}, p^t_{max}\}$.

To differentiate $v_t(p)$ from $v_*(p)$, the difference between $z_t(p)$ and $z_*(p)$ must be no less than some scalar λ which is related to the wavelength determined by the problem settings, such as the frequency band of the source wavelet, the range of possible velocity distribution, etc.. The aim of this appendix is to predict the large enough p (say p_{lowmax}) that could make true $|z_t(p) - z_*(p)| \ge \lambda$. The following paragraphs discuss how to achieve such a prediction of p_{low} in three cases.

Case 1 : $a = 1$, i.e., $c_t = c_*$.

For all $p \in P$, $z_t = z_*$. No need to do inversion.

Case 2 : $a > 1$, i.e., $c_t > c_*$.

Then, $|z_t(p) - z_*(p)| = z_t(p) - z_*(p)$. And

$$\begin{aligned}
|z_t(p) - z_*(p)| \ge \lambda \iff & a \frac{\sqrt{1-(p\,c_*)^2}}{\sqrt{1-(a\,p\,c_*)^2}} - 1 \ge \frac{\lambda}{c_*} \\
\iff & \frac{a^2\,(1-(p\,c_*)^2)}{1-(a\,p\,c_*)^2} \ge (1 + \frac{\lambda}{c_*})^2 \\
\iff & \frac{a^2-1}{1-(a\,p\,c_*)^2} \ge (1 + \frac{\lambda}{c_*})^2 - 1 \\
\iff & p^2 \ge \frac{1}{c_t^2} \left(1 - \frac{a^2-1}{\left(\frac{\lambda}{z_*}+1\right)^2 - 1} \right)
\end{aligned} \tag{B-2}$$

Let $b = 1 - \frac{a^2-1}{\left(\frac{\lambda}{z_*}+1\right)^2-1}$, then:

- if $0 < b < 1$, i.e.,

$$\begin{aligned} 0 < \frac{a^2-1}{\left(\frac{\lambda}{z_*}+1\right)^2-1} < 1 &\iff a^2 < \left(\frac{\lambda}{z_*}+1\right)^2 \\ &\iff a < \frac{\lambda}{z_*}+1 \\ &\iff \frac{c_t-c_*}{c_*} < \frac{\lambda}{z_*} \\ &\iff c_t - c_* < \frac{2\,\lambda}{\tau(0)}\,, \end{aligned} \tag{B-3}$$

  then

$$\text{Form (B-2)} \Longleftrightarrow |p| \geq \frac{1}{c_t}\sqrt{\left(1 - \frac{a^2-1}{\left(\frac{\lambda}{z_*}+1\right)^2-1}\right)}\,; \tag{B-4}$$

- if $b \leq 0$, i.e., $c_t - c_* \geq \frac{2\,\lambda}{\tau(0)}$, then all $p \in P$ works.

- $b \ngeq 1$, because $\frac{a^2-1}{\left(\frac{\lambda}{z_*}+1\right)^2-1} \geq 0$.

The above discussion concludes that if the difference between c_t and c_* is larger than $\frac{2\,\lambda}{\tau(0)}$, then for any possible p, $|z_t(p) - z_*(p)| \geq \lambda$, i.e., one could differentiate $v_t(p)$ from $v_*(p)$; if $0 < c_t - c_* < \frac{2\,\lambda}{\tau(0)}$, then for

$$p \in P \bigcap \left\{ p : |p| \geq \frac{1}{c_t}\sqrt{\left(1 - \frac{a^2-1}{\left(\frac{\lambda}{z_*}+1\right)^2-1}\right)} \right\},$$

$|z_t(p) - z_*(p)| \geq \lambda$.

Hence, in this case,

$$p_{lowmax} = \frac{1}{c_t}\sqrt{\left(1 - \frac{a^2-1}{\left(\frac{\lambda}{z_*}+1\right)^2 - 1}\right)}. \tag{B-5}$$

Case 3 : $a > 1$, i.e., $c_t < c_*$.

Then, $|z_t(p) - z_*(p)| = z_*(p) - z_t(p)$. And

$$\begin{aligned} |z_t(p) - z_*(p)| \geq \lambda \iff & 1 - a\frac{\sqrt{1-(p\,c_*)^2}}{\sqrt{1-(a\,p\,c_*)^2}} \geq \frac{\lambda}{c_*} \\ \iff & \frac{\lambda}{z_*} < 1 \quad \text{and} \quad \frac{a^2\,(1-(p\,c_*)^2)}{1-(a\,p\,c_*)^2} \leq (1 - \frac{\lambda}{c_*})^2 \\ \iff & \frac{\lambda}{z_*} < 1 \quad \text{and} \quad p^2 \geq \frac{1}{c_t^2}\left(1 - \frac{a^2-1}{\left(1-\frac{\lambda}{z_*}\right)^2 - 1}\right) \end{aligned} \tag{B-6}$$

Let $q = 1 - \frac{1-a^2}{1-\left(1-\frac{\lambda}{z_*}\right)^2}$, then:

- if $0 < q < 1$, i.e.,

$$\begin{aligned} 1 - a^2 < 1 - \left(1 - \frac{\lambda}{z_*}\right)^2 \iff & a^2 > \left(1 - \frac{\lambda}{z_*}\right)^2 \\ \iff & a > 1 - \frac{\lambda}{z_*} \\ \iff & \frac{c_* - c_t}{c_*} < \frac{\lambda}{z_*} \\ \iff & c_* - c_t < \frac{2\,\lambda}{\tau(0)}\,, \end{aligned} \tag{B-7}$$

then

$$\text{Form (B-6)} \iff \frac{\lambda}{z_*} < 1 \quad \text{and} \quad |p| \geq \frac{1}{c_t}\sqrt{\left(1 - \frac{1-a^2}{1-\left(\frac{1-\lambda}{z_*}\right)^2}\right)}\,; \tag{B-8}$$

- if $q \leq 0$, i.e., $c_* - c_t \geq \frac{2\lambda}{\tau(0)}$, then all $p \in P$ works, provided $\frac{\lambda}{z_*} < 1$.
- $q \ngeq 1$, because $\frac{1-a^2}{1-\left(1-\frac{\lambda}{z_*}\right)^2} \geq 0$.

The above discussion concludes that: given $\frac{\lambda}{z_*} < 1$, if the difference between c_t and c_* is larger than $\frac{2\lambda}{\tau(0)}$, then for any possible p, $|z_t(p) - z_*(p)| \geq \lambda$, i.e., one could differentiate $v_t(p)$ from $v_*(p)$; if $0 < c_* - c_t < \frac{2\lambda}{\tau(0)}$, then for

$$p \in P \bigcap \left\{ p : |p| \geq \frac{1}{c_t} \sqrt{\left(1 - \frac{1-a^2}{1-\left(1-\frac{\lambda}{z_*}\right)^2}\right)} \right\},$$

$|z_t(p) - z_*(p)| \geq \lambda$.

Hence, in this case,

$$p_{lowmax} = \frac{1}{c_t} \sqrt{\left(1 - \frac{1-a^2}{1-\left(1-\frac{\lambda}{z_*}\right)^2}\right)}. \tag{B-9}$$

As a summary, I have got the predictions (Forms (B-5) and (B-9)) on the smallest slowness required to differentiate a trial velocity c_t from the reference velocity c_*. The above discussion tells us that: with z_* decreasing (to 0) or c_t approaches to c_*, p_{lowmax} increases (to p_{max}), and thus the difficulty increases for differentiating c_t from c_*.

REFERENCES

Bamberger, A., G. Chavent, and P. Lailly, 1979, About the stability of the inverse problem in 1-d wave equation — application to the interpretation of seismic profiles: Applied Mathematics and Optimization, **5**, 1–47.

Bube, K. and R. Burridge, 1983, The one dimensional inverse problem of reflection seismology: SIAM Review, **25**, 497–559.

Burstedde, C. and O. Ghattas, 2007, Algorithmic strategies for full waveform inversion: 1d experiments: SEG Annual meeting.

Chauris, H. and M. Noble, 2001, Two-dimensional velocity macro model estimation from seismic reflection data by local differential semblance optimization: applications synthetic and real data sets: Geophysical Journal International, **144**, 14–26.

Chavent, G., 1991, New size times curvature conditions for strict quasiconvexity of sets: SIAM Journal on Control and Optimization, **29**, 1348–1372.

Gauthier, O., A. Tarantola, and J. Virieux, 1986, Two-dimensional nonlinear inversion of seismic waveforms: Geophysics, **51**, 1387–1403.

Kolb, P., F. Collino, and P. Lailly, 1986, Prestack inversion of a 1D medium: Proceedings of the IEEE, **74**, 498–506.

Lailly, P., 1983, The seismic inverse problem as a sequence of before-stack migrations, *in* Bednar, J., ed., Conference on Inverse Scattering: Theory and Applications, 206–220. Society for Industrial and Applied Mathematics.

Li, J. and W. Symes, 2007, Interval velocity estimation via nmo-based differential semblance: Geophysics.

Sacks, P. and F. Santosa, 1987, A simple computational scheme for determining the sound speed of an acoustic medium from its surface impulse response: SIAM J. SCI. STAT. COMPUT., **8**, 501–520.

Santosa, F. and W. Symes, 1989, An analysis of least-squares velocity inversion. Number 4 *in* Geophysical monograph series: SEG.

Sen, M. and P. Stoffa, 1991a, Nonlinear multiparameter optimization using genetic algorithms: Inversion of plane wave seismograms: Geophysics, **56**, 1794–1810.

———, 1991b, Nonlinear one-dimensional seismic waveform inversion using simulated annealing: Geophysics, **56**, 1624–1636.

Shin, C. and D.-J. Min, 2006, Waveform inversion using a logarithmic wavefield: Geophysics, **71**, R31–R42.

Symes, W., 1981, The inverse reflection problem for a smoothly stratified elastic medium: SIAM Journal on Mathematical Analysis, **12**, 421–453.

———, 1986a, Linearization stability for an inverse problem in several-dimensional wave propagation: SIAM Journal on Mathematical Analysis, **17**, 132–151.

———, 1986b, On the relation between coefficient and boundary values for solutions of webster's horn equation: SIAM Journal on Mathematical Analysis, **17**, 1400–1420.

———, 1991a, A differential semblance algorithm for the inverse problem of reflection seismology: Computers and Mathematics with Applications, **22**, 147–178.

———, 1991b, Layered velocity inversion: a model problem from reflection seismology:

SIAM Journal on Mathematical Analysis, **22**, 680–716.

———, 1993, A differential semblance criterion for inversion of multioffset seismic reflection data: J. Geoph. Res., **98**, 2061–2073.

Symes, W. and J. Carazzone, 1991, Velocity inversion by differential semblance optimization: Geophysics, **56**, 654–663.

———, 1992, Velocity inversion by coherency optimization: Geophysical Inversion, 59–89.

Symes, W. and R. Versteeg, 1993, Velocity model determination using differential semblance optimization: 63rd annual international meeting: Expanded abstracts, 696-699, Society of Exploration Geophysicists.

Symes, W. W., 1999, All stationary points of differential semblance are asymptotic global minimizers: layered acoustics: Technical Report 99-29, Department of Computational and Applied Mathematics, Rice University, Houston, Texas, USA.

———, 2007, Migration velocity analysis and waveform inversion: Technical Report 07-05, Department of Computational and Applied Mathematics, Rice University, Houston, Texas, USA.

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

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

———, 1986, A strategy for nonlinear elastic inversion of seismic reflection data: Geophysics, **51**, 1893–1903.

Tarantola, A. and B. Valette, 1982, Generalized nonlinear inverse problems solved using the least squares criterion: Peviews of Geophysics and Space Physics, **20**, 219–232.

Tarantola, A. and B. Vallette, 1982, Inverse problems: quest for information: J. Geophysics, **50**, 159–170.

Verm, R. and W. Symes, 2006, Practice and pitfalls in nmo-based differential semblance velocity analysis: 75th Annual International Meeting, Expanded Abstracts, SEG.