

Acoustic Based Crosshole Full Waveform Slowness Inversion in the Time Domain

Wensheng Zhang^{1,2}, Atish Kumar Joardar^{1,2,3}

¹LSEC, ICMSEC, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, China

²School of Mathematical Sciences, University of Chinese Academy of Sciences, Chinese Academy of Sciences, Beijing, China

³Department of Mathematics, Islamic University, Kushtia, Bangladesh

Email: zws@lsec.cc.ac.cn; atishjoardar@gmail.com

How to cite this paper: Zhang, W.S. and Joardar, A.K. (2018) Acoustic Based Crosshole Full Waveform Slowness Inversion in the Time Domain. *Journal of Applied Mathematics and Physics*, 6, 1086-1110.
<https://doi.org/10.4236/jamp.2018.65094>

Received: April 16, 2018

Accepted: May 28, 2018

Published: May 31, 2018

Copyright © 2018 by authors and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

We develop a new full waveform inversion (FWI) method for slowness with the crosshole data based on the acoustic wave equation in the time domain. The method combines the total variation (TV) regularization with the constrained optimization together which can inverse the slowness effectively. One advantage of slowness inversion is that there is no further approximation in the gradient derivation. Moreover, a new algorithm named the skip method for solving the constrained optimization problem is proposed. The TV regularization has good ability to inverse slowness at its discontinuities while the constrained optimization can keep the inversion converging in the right direction. Numerical computations both for noise free data and noisy data show the robustness and effectiveness of our method and good inversion results are yielded.

Keywords

Acoustic Wave Equation, Crosshole, Full Waveform Inversion, Slowness, Bound Constraints, TV Regularization

1. Introduction

Seismic exploration is one of the ways of identifying media properties and structures by propagation of waves. The wave is ignited by sources on the surface and propagates into underground. Due to different properties of media, various waves such as reflection wave and refraction wave are produced. The nature of wave reflection and refraction gives the physical information of media. So the inverse of seismic waves can give geological properties of underground

materials. Unlike the observation on the surface, crosshole configuration arranges sources in one well or hole and receivers in another well. Most approaches to the inversion of crosshole data involve travelttime tomography. Travelttime tomography is efficient and robust. However, the resolution of travelttime tomography is limited by its foundation on ray theory and its typical use of first arrivals rather than the full waveform.

The full waveform inversion (FWI) is a high resolution method to inverse the media parameter by using the whole wavefield information such as amplitude, phase and arrival time. The FWI can be divided into two categories, the time-domain method and the frequency-domain method. The FWI was originally developed in the time domain [1] [2] [3] [4] [5]. In 1984, Tarantola [1] laid the foundations for full nonlinear waveform inversion using the formalism of least-squares optimization for the wavefield misfit function. Because of the computational intensity of the forward modelling process, the global optimization methods are impractical. Instead, the gradient-based algorithms are applied to find the local minimum of the misfit function. In 1986, Tarantola [3] extended the theory to elastic wave inversion. The idea was applied to two dimensional crosshole data inversion successfully in the frequency domain [6] [7] [8] [9]. Later much more attentions are paid to the frequency domain method, and the frequency domain method has the same inversion ability like time domain method. Its advantage is that it can be implemented with one selected frequency. However, it needs to solve a large-scale system unlike the time domain method which the forward problem can be implemented effectively in the time domain.

The FWI is a typical ill-posed problem for its nonlinearity and cycle-skipping problem. To overcome the ill-posedness, various methods have been developed, for example the multiscale method [10] [11] [12] and the regularization method [13] [14]. Other methods such as the Laplace-domain method [15] [16] [17] and the envelop FWI method [18] are proposed to improve the inversion robustness. Both the Laplace-domain FWI and envelop FWI are designed to make best of low frequency to enhance the robustness of the inversion. An overview of full waveform inversion can be found in [19]. Driven by potential application, the development of FWI is very rapidly, for example, see [20] [21] [22].

In this paper, we develop an effective full waveform slowness inversion method for crosshole data by combing the total variation regularization with the constrained optimization together. We also propose the new skip algorithm to implement the constrained optimization problem. The rest of this paper is arranged as follows. The theoretical method is described in Section 2 in detail. It includes the forward method and the inverse method. In Section 3, numerical computations both for the noise free data and noisy data are implemented. Finally, the conclusion is drawn in Section 4.

2. Theory

There are three subsections in this section. In Section 2.1, the forward problem

and the staggered-grid scheme with the perfectly matched layer are described. In Section 2.2, the inversion method and the corresponding algorithm by combing TV regularization with bound constraints are described. Section 2.3 involves the computations for the gradient of objective function and the step length, which is important to the success of inversion.

2.1. Forward Method

We consider the acoustic wave equation excited by the source $f(T)$ located at the point (x_s, z_s) , which the mathematical form of the pressure $u(x, z, t)$ satisfies

$$\sigma \frac{\partial^2 u}{\partial t^2} - \nabla \cdot (\nabla u) = f(t) \delta(x - x_s) \delta(z - z_s), \text{ in } \Omega \times (0, T) \tag{1}$$

$$u(x, z, t)|_{t=0} = 0, \left. \frac{\partial u(x, z, t)}{\partial t} \right|_{t=0} = 0, \text{ on } \Omega \tag{2}$$

where $\Omega \subset R^2$ is the area of computational domain, T is the final observation time, $\sigma = 1/v^2$ is the square of slowness parameter, v is the wave velocity. The system (1)-(2) is the forward problem. In this paper, we solve the pressure u numerically with the finite difference method for its high computational efficiency. Among various of finite difference schemes, the staggered-grid scheme is a typical one [23]. In order to apply the staggered-grid method, we introduce new variables w_x and w_z defined by $\partial u / \partial x$ and $\partial u / \partial z$ respectively. Then system (1)-(2) can be written as the following hyperbolic form

$$\frac{\partial w_x}{\partial t} = \frac{\partial u}{\partial x}, \tag{3}$$

$$\frac{\partial w_z}{\partial t} = \frac{\partial u}{\partial z}, \tag{4}$$

$$\sigma \frac{\partial u}{\partial t} = \frac{\partial w_x}{\partial x} + \frac{\partial w_z}{\partial z} + \int_r f(t) \cdot \delta(x - x_s) \delta(z - z_s) dt. \tag{5}$$

Now we construct the difference scheme of (3)-(5) on the staggered grids. The schematic map of grid points for different variables is shown in **Figure 1**. Let $u_{i,j}^n$ be the wavfield at time $n\Delta t$ and position $(i\Delta x, j\Delta z)$, and $(w_x)_{i+1/2,j}^{1/2}$ be the value at time $(n+1/2)\Delta t$ and position $((n+1/2)\Delta x, j\Delta z)$, and so on. Here Δt is the time step, and Δx and Δz are the spatial steps in x and z directions respectively. We approximate Equation (3) at time n and position $(i+1/2, j)$, Equation (5) at time n and position $(i, j+1/2)$, and Equation (4) at time $(n+1/2)$ and position (i, j) . Thus we have

$$(w_x)_{i+1/2,j}^{n+1/2} - (w_x)_{i+1/2,j}^{n-1/2} = \frac{\Delta t}{\Delta x} (u_{i+1,j}^n - u_{i,j}^n), \tag{6}$$

$$(w_z)_{i,j+1/2}^{n+1/2} - (w_z)_{i,j+1/2}^{n-1/2} = \frac{\Delta t}{\Delta z} (u_{i,j+1}^n - u_{i,j}^n), \tag{7}$$

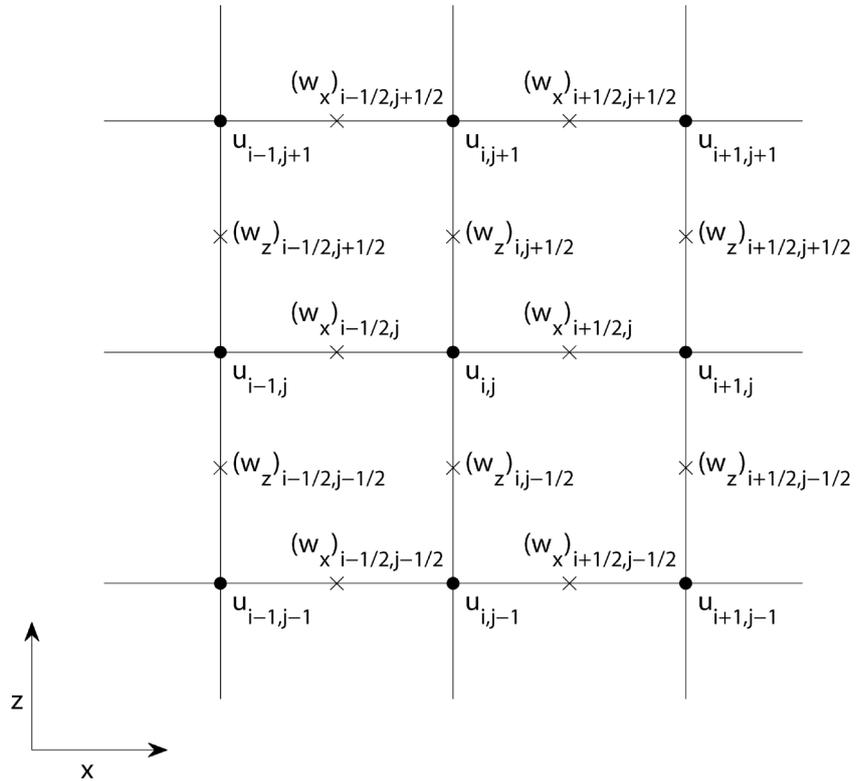


Figure 1. The schematic map of grid points for the staggered-grid scheme.

$$\begin{aligned}
 u_{i,j}^{n+1} - u_{i,j}^n &= \frac{\Delta t}{\sigma_{i,j} \Delta x} \left((w_x)_{i+\frac{1}{2},j}^{n+\frac{1}{2}} - (w_x)_{i-\frac{1}{2},j}^{n+\frac{1}{2}} \right) \\
 &+ \frac{\Delta t}{\sigma_{i,j} \Delta z} \left((w_z)_{i,j+\frac{1}{2}}^{n+\frac{1}{2}} - (w_z)_{i,j-\frac{1}{2}}^{n+\frac{1}{2}} \right) + \sigma_{i,j} \Delta t^2 \sum_{s=0}^n f_{i,j}^s,
 \end{aligned} \tag{8}$$

with initial conditions

$$u_{i,j}^{-1} = 0, \quad u_{i,j}^0 = 0, \tag{9}$$

$$(w_x)_{i+\frac{1}{2},j}^{\frac{1}{2}} = 0, \quad (w_z)_{i,j+\frac{1}{2}}^{\frac{1}{2}} = 0. \tag{10}$$

The sufficient and necessary stability for the scheme (6)-(8) is (see **Appendix A**)

$$\Delta t \leq \sqrt{\sigma \left(\frac{1}{\Delta x^2} + \frac{1}{\Delta z^2} \right)}. \tag{11}$$

Since the computational domain is limited, we need to use absorbing boundary conditions to eliminate the boundary reflections. Several methods can be applied, for example, the paraxial approximation [24], the exact nonreflecting conditions [25] [26] and the perfectly matched layer (PML) method [27]. Here, we adopt the PML method in its split formulation. The idea of PML is to design a particular layer around the computational domain in which the waves will be attenuated satisfactory. In split PML formulation, the wavefield u is split into

two components, x -component and z -component, *i.e.*, $u = u_x + u_z$. To remove the boundary reflections in the absorbing layer, a decaying coefficient $d(x)$ or $d(z)$ is introduced. Then system (3)-(5) in the case of the source free can be written as the following PML form

$$\frac{\partial u_x}{\partial t} + d(x)u_x = \frac{1}{\sigma} \frac{\partial w_x}{\partial x}, \tag{12}$$

$$\frac{\partial u_z}{\partial t} + d(z)u_z = \frac{1}{\sigma} \frac{\partial w_z}{\partial z}, \tag{13}$$

$$\frac{\partial w_x}{\partial t} + d(x)w_x = \frac{\partial u}{\partial x}, \tag{14}$$

$$\frac{\partial w_z}{\partial t} + d(z)w_z = \frac{\partial u}{\partial z}, \tag{15}$$

where $d(x)$ and $d(z)$ are designated to attenuate the refractions in absorbing zone.

We use the following model for the damping parameter $d(p)$ [28] [29]

$$d(p) = d_0 \left(\frac{p}{L} \right)^2, \quad d_0 = -\frac{3\nu}{2L} \ln(R), \quad p = x, z, \tag{16}$$

where L is the thickness of PML absorbing layer, ν is the media velocity, x or z is the distance from current position to PML inner boundary, and R is a parameter chosen as $10^{-3} \sim 10^{-6}$. Note that the coefficient $d(p)$ is zero in the original computational domain. The staggered-grid scheme for system (12)-(15) is

$$(w_x)_{i+\frac{1}{2},j}^{n+\frac{1}{2}} = \frac{1-\Delta t d_i/2}{1+\Delta t d_i/2} (w_x)_{i+\frac{1}{2},j}^{n-\frac{1}{2}} + \frac{1}{1+\Delta t d_i/2} \frac{\Delta t}{\Delta x} (u_{i+1,j}^n - u_{i,j}^n), \tag{17}$$

$$(w_z)_{i,j+\frac{1}{2}}^{n+\frac{1}{2}} = \frac{1-\Delta t d_j/2}{1+\Delta t d_j/2} (w_z)_{i,j+\frac{1}{2}}^{n-\frac{1}{2}} + \frac{1}{1+\Delta t d_j/2} \frac{\Delta t}{\Delta z} (u_{i,j+1}^n - u_{i,j}^n), \tag{18}$$

$$(u_x)_{i,j}^{n+1} = \frac{1-\Delta t d_i/2}{1+\Delta t d_i/2} (u_x)_{i,j}^n + \frac{1}{1+\Delta t d_i/2} \frac{\Delta t}{\Delta x \sigma_{i,j}} \left((w_x)_{i+\frac{1}{2},j}^{n+\frac{1}{2}} - (w_x)_{i-\frac{1}{2},j}^{n+\frac{1}{2}} \right), \tag{19}$$

$$(u_z)_{i,j}^{n+1} = \frac{1-\Delta t d_j/2}{1+\Delta t d_j/2} (u_z)_{i,j}^n + \frac{1}{1+\Delta t d_j/2} \frac{\Delta t}{\Delta z \sigma_{i,j}} \left((w_z)_{i,j+\frac{1}{2}}^{n+\frac{1}{2}} - (w_z)_{i,j-\frac{1}{2}}^{n+\frac{1}{2}} \right), \tag{20}$$

$$u_{i,j}^{n+1} = (u_x)_{i,j}^{n+1} + (u_z)_{i,j}^{n+1}. \tag{21}$$

The forward computations can be extrapolated explicitly in the time direction based on (17)-(21). In **Figure 2**, the snapshot of wave propagation in a homogeneous media with velocity 3000 m/s at time 0.8 s is shown. **Figure 2(a)** is the result without PML while **Figure 2(b)** is the result with PML. We can see that the serious boundary reflections in **Figure 2(a)** are eliminated effectively in **Figure 2(b)**.

2.2. Inversion with Constraints

Consider the inversion of slowness fields in the vector $\sigma = (\sigma_1, \sigma_2, \dots, \sigma_s)^T$.

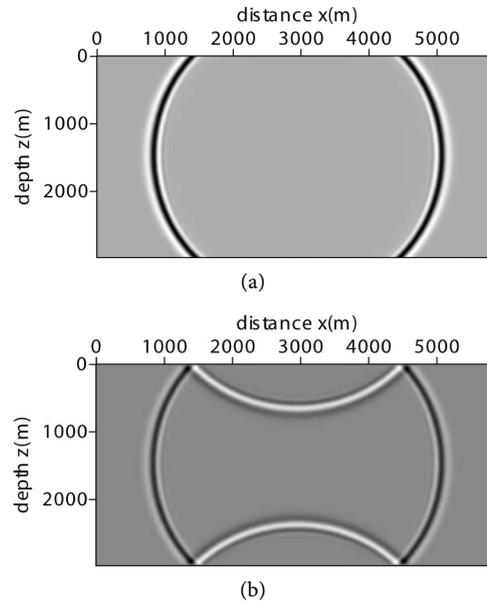


Figure 2. Snapshot of wave propagation at time 0.8 s. (a) No PML; (b) PML.

Define the objective function $J(\sigma)$ as

$$J(\sigma) = \frac{1}{2} \sum_{shot} \int_0^T \int_{\Omega} (u(x, z, t) - u_{obs}(x, z, t))^2 \delta(x - x_r) dx dz dt + \eta \int_{\Omega} \sqrt{\varepsilon^2 + |\nabla \sigma|^2} dx dz, \tag{22}$$

where the summation is for all the shots or sources in the well. In Equation (22), the first term comes from the misfit between the forward data u and the observed data u_{obs} , the second term is the TV regularization with positive parameter ε^2 , and $\eta > 0$ is the regularization parameter. The TV regularization is applied because it can yield better inversion result at jump discontinuities theoretically [30] [31]. For more theory about Tikhonov theory, reader may refer to references, for example [32] [33]. The inverse problem is to minimize the objective function, *i.e.*,

$$\min_{\sigma} J(\sigma). \tag{23}$$

It is a high dimensional complex optimization problem. The dimensions are $N_x N_z$, where N_x and N_z are the discretization points along x and z directions respectively. In this paper, the iterative decent method named the global Barzilai-Borwein (GBB) method is applied because it is suitable for solving high-dimensional complex minimization problem [34] [35] [36]. Generally, for a given nonlinear function $f : R^n \rightarrow R$, let $\sigma^0 \in R^n$ be the initial slowness vector, then the $(m + 1)$ -th iteration is obtained by

$$\sigma^{m+1} = \sigma^m - \alpha^m \nabla f(\sigma^m), \tag{24}$$

where the superscript m denotes the iteration number, α^m is the step length and ∇f is the gradient of f at the point σ^m .

Now we propose to solve the minimization problem (23) as a constrained optimization problem subject to the constraints $a \leq \sigma \leq b$. Here a and b are the lower bound and upper bound for each component of the vector σ respectively. The solution of constraint problem for nonlinear optimization is challenging. Usually, the penalty function technique is used to change the constrained optimization to unconstrained optimization problem [37]. In this paper, we develop a new effective algorithm called the skip method to solve it. The concept of the skip method is to skip the current inverse result if it goes outside of the bound constraints and simply store previous results. For our experience, the skip method performs better than the penalty method. Numerical computations in Section 3 will show the effectiveness of the skip method.

2.3. Computations for the Gradient and Step Length

For minimizing the problem (23), the gradient and the step length are required. The gradient of the objection functional (22) with respect to σ is

$$\frac{\partial J(\sigma)}{\partial \sigma} = -\sum_{shot} \int_0^T \int_{\Omega} \phi \frac{\partial^2 u}{\partial t^2} dx dz dt - \eta \int_{\Omega} \nabla \cdot \frac{\nabla \sigma}{\sqrt{\varepsilon^2 + |\nabla \sigma|^2}} dx dz, \tag{25}$$

where ϕ is the backward wavefield satisfying

$$\sigma \frac{\partial^2 \phi}{\partial t^2} - \nabla \cdot \nabla \phi = -(u - u_{obs}) \cdot \delta(x - x_r), \tag{26}$$

$$\phi|_{t=T} = 0, \quad \frac{\partial \phi}{\partial t} \Big|_{t=T} = 0, \tag{27}$$

which can be solved effectively like the forward problem in Section 2.1. The backward problem (26)-(27) is introduced in the gradient derivation naturally in **Appendix B**. There are two terms on the right hand side of Equation (25). The first is the gradient of the misfit functional between the forward data u and the observed data u_{obs} , which is derived in **Appendix B**. We remark that there is no Born approximation in the gradient derivation in **Appendix B**. The second is the gradient of TV regularization term and its discrete form can be obtained by the approximation of the gradient operator, *i.e.*,

$$|\nabla \sigma|^2 \approx (D_x \sigma_{i,j})^2 + (D_z \sigma_{i,j})^2, \tag{28}$$

where

$$D_x \sigma_{i,j} = \frac{\sigma_{i+1,j} - \sigma_{i-1,j}}{2\Delta x}, \quad D_z \sigma_{i,j} = \frac{\sigma_{i,j+1} - \sigma_{i,j-1}}{2\Delta z}. \tag{29}$$

The parameter ε in Equation (25) is a small number to prevent the singularity of the denominator.

Now we consider the step length which is important to the success of inversion. For a general objective function ϕ :

$$\phi(\alpha_k) = f(x_k + \alpha_k p_k), \tag{30}$$

the one dimensional line search is applied to find the step length, where p_k is the search direction and α_k is the step length. Line search condition gives the

step length α_k which guarantees the sufficient decrease in the objective function f by inequality

$$f(x_k + \alpha_k p_k) \leq f(x_k) + c_1 \alpha_k \nabla f_k^T p_k$$

for some constant $c_1 \in (0, 1)$. The reduction of the objective function is proportional to both the search direction and the step length. The sufficient condition states that α_k is the step length only if

$$\phi(\alpha_k) \leq f(x_k) + c_1 \alpha_k \nabla f_k^T p_k.$$

The decrease is not enough if it only satisfies the sufficient condition because it can satisfy for the very small value of α . For this reason it requires another essential condition called the curvature condition. If α_k satisfies

$$\nabla f(x_k + \alpha_k p_k)^T p_k \geq c_2 \nabla f_k^T p_k$$

for some constant $c_2 \in (c_1, 1)$, it is called to satisfy the curvature condition. The sufficient decrease condition and the curvature condition together are called Wolfe condition [37]. It can also be possible a step length satisfies the Wolfe condition but it closes to a minimizer of ϕ . Due to this reason we use the modified curvature condition to enlarge the step length which lies in a broad neighbourhood of local minimizer. Therefore, α_k satisfies the strong Wolfe condition if it satisfies the following two conditions

$$f(x_k + \alpha_k p_k) \leq f(x_k) + c_1 \alpha_k \nabla f_k^T p_k, \tag{31}$$

$$\left| \nabla f(x_k + \alpha_k p_k)^T p_k \right| \leq c_2 \left| \nabla f_k^T p_k \right|, \tag{32}$$

where $0 < c_1 < c_2 < 1$. For the importance of line search in inversion, the algorithm for line search is given by Algorithm 1 and Algorithm 2. The FWI algorithm with both TV regularization and bound constraints is given by Algorithm 3.

Algorithm 1: Line search for the step length
 Let $\alpha_0 = 0$, choose $\alpha_{\max} > 0$ and $\alpha_1 \in (0, \alpha_{\max})$; $i \leftarrow 1$;
Repeat
 Compute $J(\alpha_i)$ based on equation (22);
 If
 $J(\alpha_i) > J(0) + c_1 \alpha_i J'(0)$ or $J(\alpha_i) \geq J(\alpha_{i-1})$ for $i > 1$
 then
 $\alpha_* \leftarrow \text{Zoom}(\alpha_{i-1}, \alpha_i)$ and stop;
 End if
 Compute $J'(\alpha_i)$;
 If $|J'(\alpha_i)| \leq -c_2 J'(0)$ **then**
 Set $\alpha_* \leftarrow \alpha_i$ and stop;
 End if
 If $J'(\alpha_i) \geq 0$ **then**
 Set $\alpha_* \leftarrow \text{Zoom}(\alpha_i, \alpha_{i-1})$ and stop;
 End if
 Choose $\alpha_{i+1} \in (\alpha_i, \alpha_{\max})$; $i \leftarrow i + 1$
Until (The step length yields)

Algorithm 2: Zoom step in Algorithm 1**Repeat**Find a trial step length α_k between α_{lo} and α_{hi} by interpolation;Compute $J(\alpha_k)$ based on equation (22);**If** $J(\alpha_k) > J(0) + c_1\alpha_k J'(0)$ or $J(\alpha_k) \geq J(\alpha_{lo})$ **then** $\alpha_{hi} \leftarrow \alpha_k$;**else**Compute $J'(\alpha_k)$;**If** $|J'(\alpha_k)| \leq -c_2 J'(0)$ **then**Set $\alpha_* \leftarrow \alpha_k$ and stop;**End if****If** $J'(\alpha_k)(\alpha_{hi} - \alpha_{lo}) \geq 0$ **then** $\alpha_{hi} \leftarrow \alpha_{lo}$; $\alpha_{lo} \leftarrow \alpha_k$;**End if****End if****Until** (The step length yields)**Algorithm 3:** FWI with both TV regularization and bound constraintsLet $\sigma = \sigma^0$ and set the tolerance**For** $m = 0$ and satisfy the tolerance condition **do****For** all shots **do**

Solve the forward problem with (17)-(21);

Solve the backward problem (26)-(27);

End for

Compute the gradient from equation (25);

Compute the step length by line search Algorithm 1;

If σ^{m+1} is in the bound interval **then**Update the σ^{m+1} ;**Else**Not update $\sigma^m \rightarrow \sigma^{m+1}$;**End if****End for**

3. Numerical Computations

We implement numerical computations with C-language code. The computational domain is $\Omega = X \times Z$, where $X = [0, 0.25 \text{ km}]$ and $Z = [0, 0.25 \text{ km}]$. The spatial increment is $\Delta x = \Delta z = 8.33 \text{ m}$ and time step $\Delta t = 0.001 \text{ s}$. In numerical computations, the regularization parameter η is around 1.0×10^{-9} . The crosshole configuration is shown in **Figure 3**. The sources are located in the left well marked as * while the receivers are located in the right well marked as \circ . There are 27 sources in the left well and 29 receivers in the right well. The source function is the Ricker wavelet given by

$$f(t) = A \left(2 \left(\pi f_0 (t - t_0) \right)^2 - 1 \right) e^{-\pi f_0 (t - t_0)^2}, \quad (33)$$

where $f_0 = 25 \text{ Hz}$ is the peak frequency, A is the maximum amplitude and

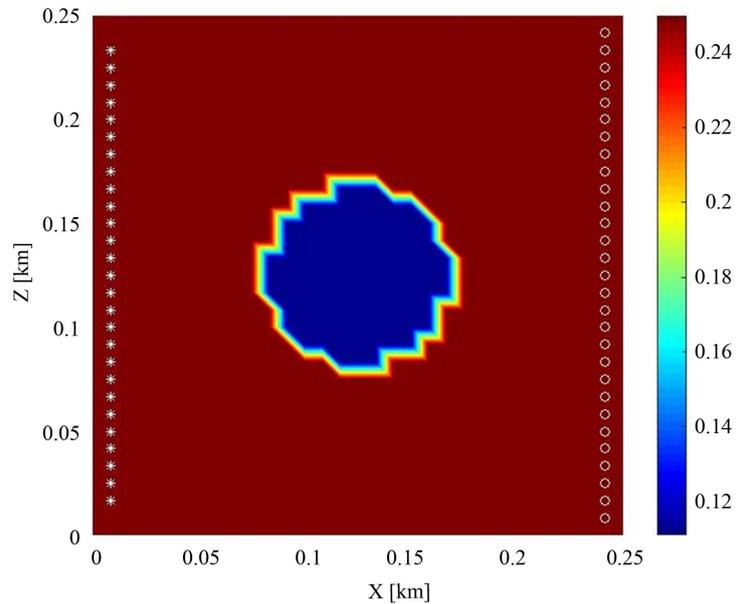


Figure 3. The exact slowness model. The sources are in the left well marked as * and the receivers in the right well marked as o.

$t_0 = 0.01$ s is the time when maximum amplitude occurs. The squared slowness of the model exact is

$$\sigma = \begin{cases} 0.111, & (x, z) \in C \\ 0.250, & (x, z) \in \Omega \setminus C \end{cases} \quad (34)$$

where

$$C = \{(x, z) : (x - 0.125)^2 + (z - 0.125)^2 < 0.000851\}.$$

As shown in **Figure 3**, this exact model contains a roughly spherical body with lower slowness in the middle on the background media with high slowness. We will consider the inversion in two cases: the noise free data and nosy data.

3.1. Noise Free Data

First we consider the inverse for the noise free data. The inverse result after 54 iterations without TV regularization and bound constraints in Algorithm 3 is shown in **Figure 4**. We remark that more iterations in **Figure 4** is impossible because no step length can be found and the iteration is automatically stopped. We can see the inverse result differs from the exact model very much, which shows the inversion does not converge correctly. The inverse results with three different methods after 200 iterations are shown in **Figure 5**. For contrast, the exact model is shown in **Figure 5(a)** again. **Figures 5(b)-(d)** are the inverse results with TV regularization, with bound constraints, and with both TV regularization and bound constraints, respectively. Comparisons show that **Figures 5(b)-(d)** are much better than **Figure 4**. Moreover, **Figure 5(d)** with both TV regularization and bound constraints is best. **Figure 6** is the convergence history

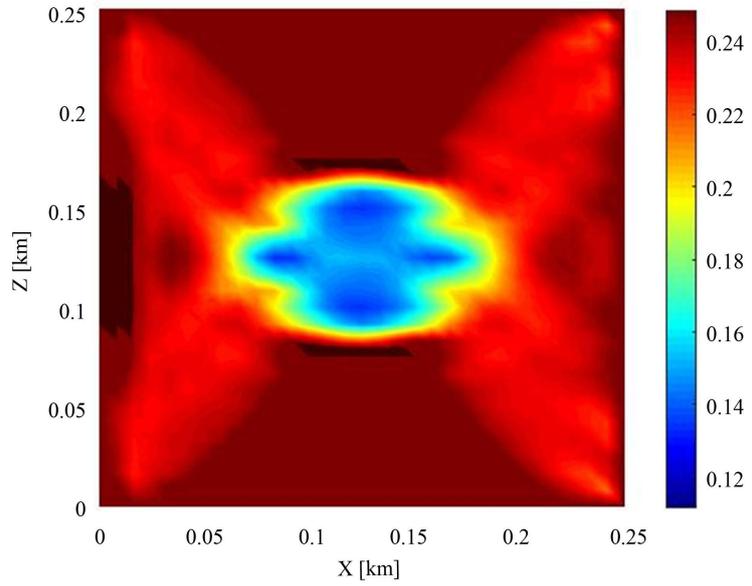


Figure 4. Inverse result without TV regularization and bound constraints. The inversion is stopped automatically after 54 iterations.

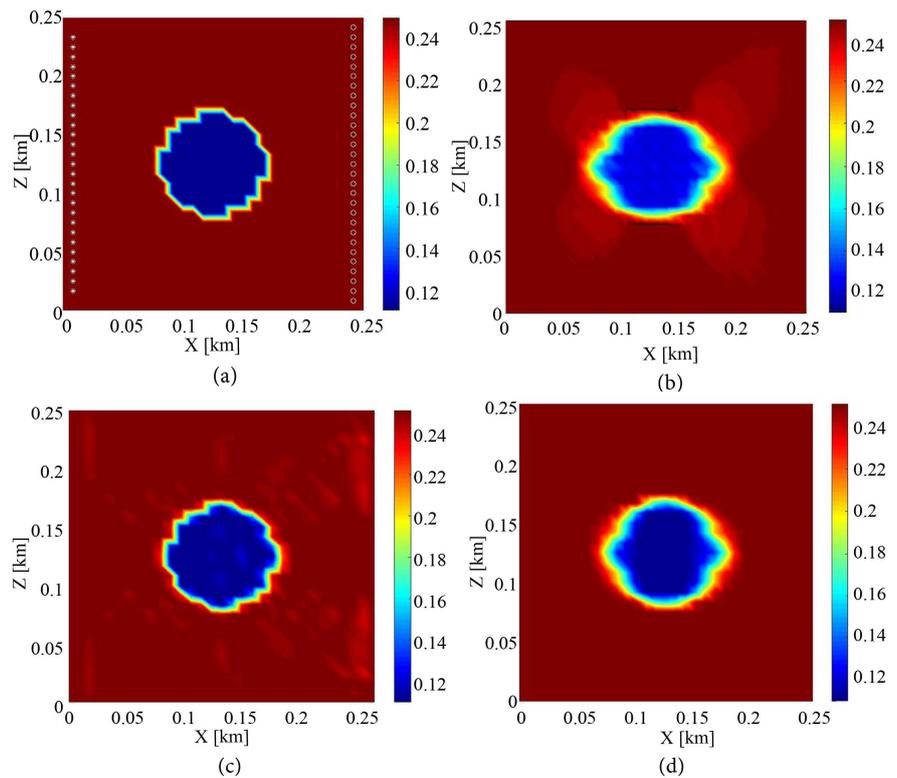


Figure 5. Exact model and inverse results with three different methods after 200 iterations. (a) Exact model; (b) TV regularization; (c) Bound constraints; (d) Both TV regularization and bound constraints.

of the objective function for the first 200 iterations. In **Figure 6**, we can see the curve for the method with both TV regularization and bound constraints

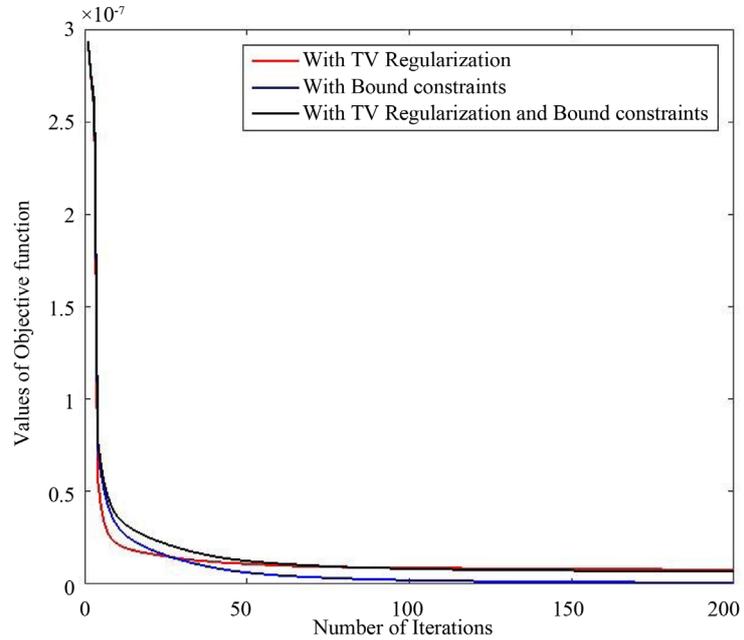


Figure 6. The objective function values for the first 200 iterations.

decreases fastest. The inverse results can be improved further if more iterations are carried out. **Figure 7** is a comparison between the exact model and three inverse results after 1000 iterations. Every inversion method takes approximately 12 hours for 1000 iterations with single processor (cpu MHz: 20000.021). **Figure 7(a)** is the exact model. **Figures 7(b)-(d)** are the inverse results with TV regularization, with bound constraints, and with both TV regularization and bound constraints, respectively. We can clearly see that **Figures 7(b)-(d)** are much better than **Figures 5(b)-(d)**, respectively. And **Figure 7(d)** obtained with both TV regularization and bound constraints performs best. Therefore, the TV regularization and bound constraints are necessary and its effectiveness in inversion is obvious.

3.2. Noisy Data

The recorded data are usually affected by noise due to error in measurements or the influence of natural environment. So it is necessary to consider noise analysis. The Gaussian white noise is one of the effective way to add noise. It has the normal distribution and standard deviation of the data. Let the discrete signal be $\{S^i\}_{i=1}^{N_t}$ and the noisy signal $\{\tilde{S}^i\}_{i=1}^{N_t}$. To obtain $\{\tilde{S}^i\}_{i=1}^{N_t}$, we simply add $\{S^i\}_{i=1}^{N_t}$ to a white Gaussian noise which is random number with the mean zero and standard deviation under different signal-to-noise ratio (SNR)

$$\tilde{S}^i = S^i + N^i, \quad N^i \sim \mathcal{N}\left(0, \frac{\|S\|^2/N_t}{\text{SNR}}\right), \quad (35)$$

where $\mathcal{N}(m, s^2)$ denotes a normal distribution with mean m and standard deviation s^2 , and $\|\cdot\|$ is the Euclidean mean. We first present some noisy data

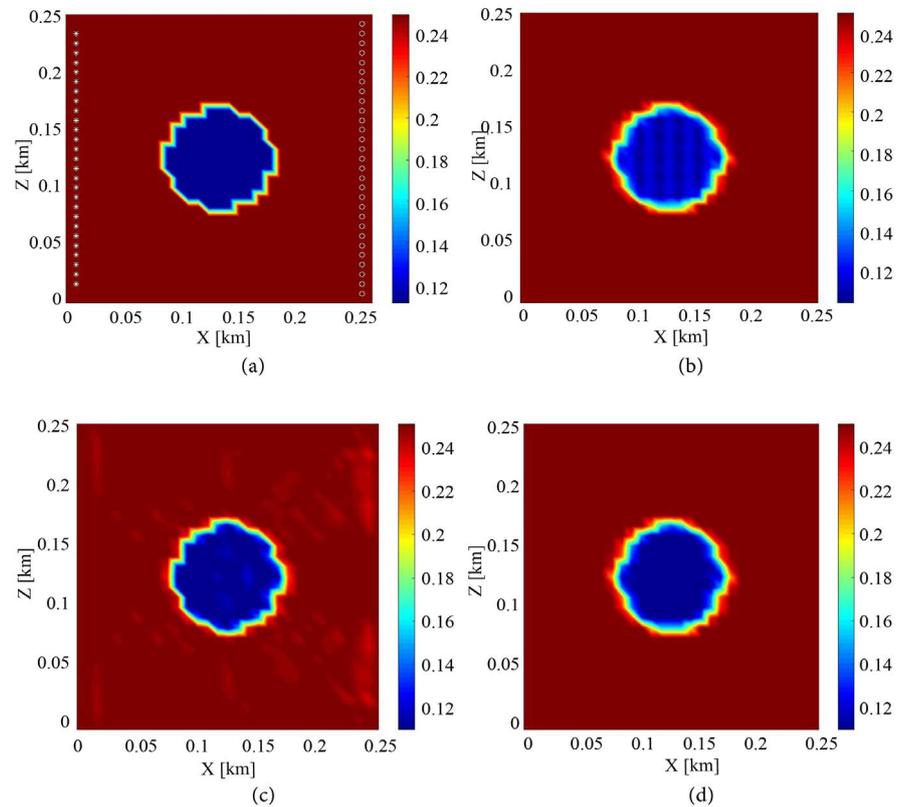
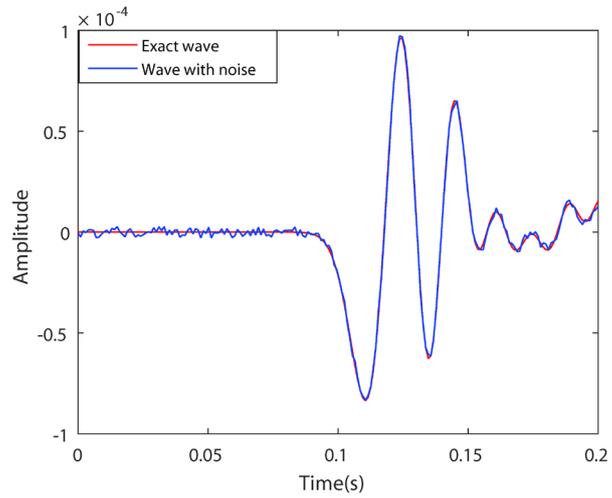
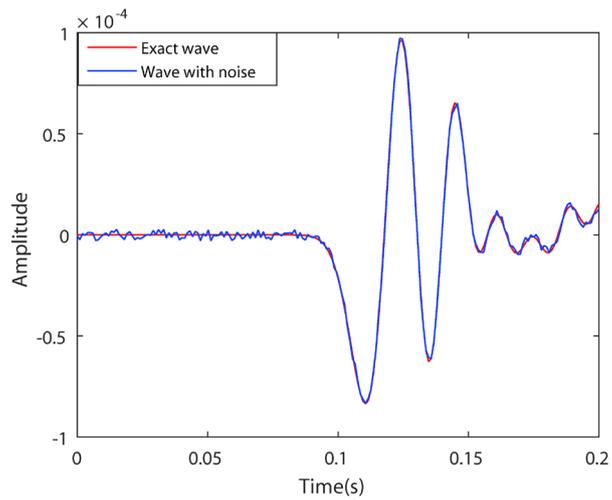


Figure 7. Exact model and inverse results with three different methods after 1000 iterations. (a) Exact model; (b) TV regularization; (c) Bound constraints; (d) Both TV regularization and bound constraints.

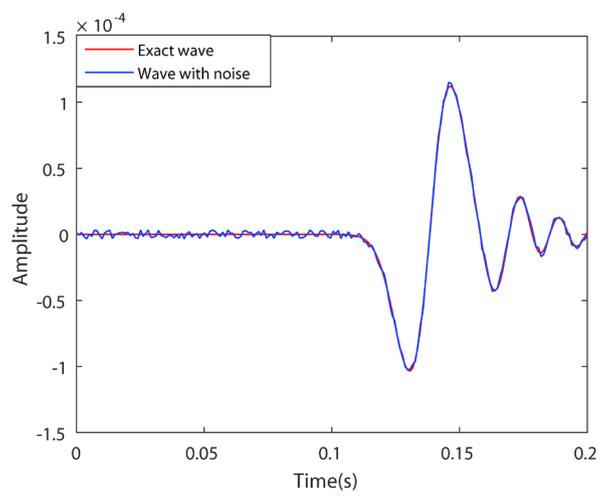
at different receivers with different SNR levels. **Figures 8-10** are the noisy data received from a source with SNR = 100, SNR = 10, SNR = 1, respectively. In each figure, the data received at three different receivers z_r , *i.e.*, $z_r = N_z/2$, $z_r = 3N_z/4$ and $z_r = N_z$ are shown. Here N_z is the number of discretization points in z direction. In these figures, the red line denotes the noise free data while the blue line represents the noisy data and we can see clearly that the lower SNR is the higher noise in the data. Generally, the inversion for noisy data is challenge because the problem is nonlinear and ill-posed. However, numerical inverse results by our proposed method are still good. In **Figure 11**, the inverse results by three different methods for noisy data with SNR = 100 are given. **Figure 11(a)** is the exact model. **Figures 11(b)-(d)** are the inverse results after 1000 iterations with TV regularization, with bound constraints, and with both TV regularization and bound constraints, respectively. Comparing the results in **Figure 11**, we can clearly see that **Figure 11(d)** obtained with both TV regularization and bound constraints performs best. Correspondingly, the inverse results for the noisy data with SNR = 10 and SNR = 1 are shown in **Figure 12** and **Figure 13** respectively. After comparing these results, we can see that even in the case of lower SNR with SNR = 1, the inverse result with both TV regularization and bound constraints are still good, which shows the robustness of our proposed method.



(a)

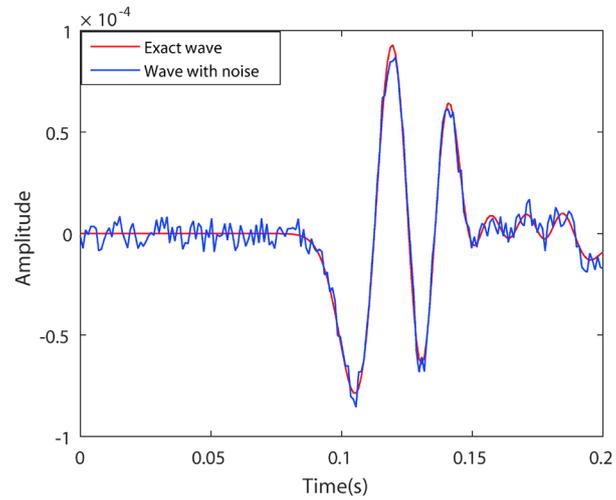


(b)

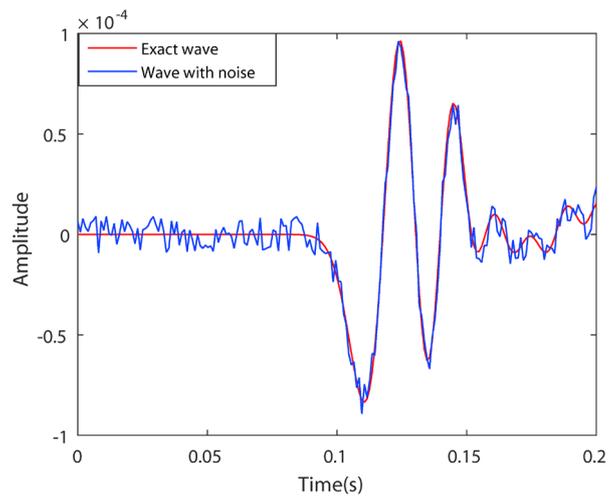


(c)

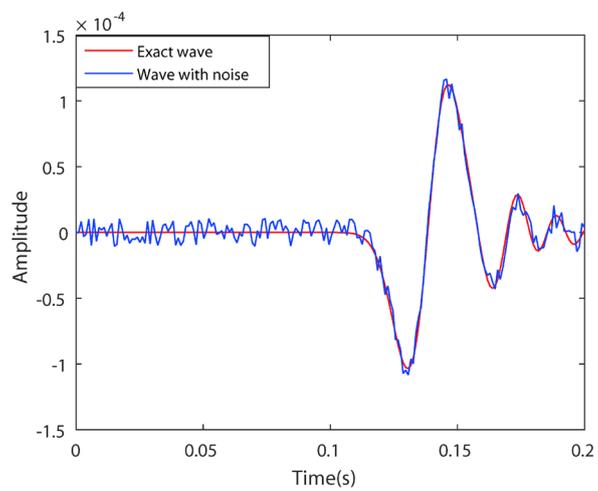
Figure 8. Noisy data with SNR = 100 received in the right well at three different positions z_r . (a) $z_r = N_z/2$, (b) $z_r = 3N_z/4$, (c) $z_r = N_z$.



(a)

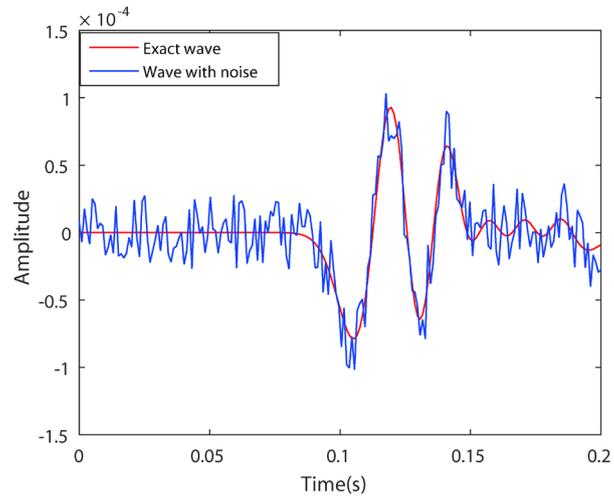


(b)

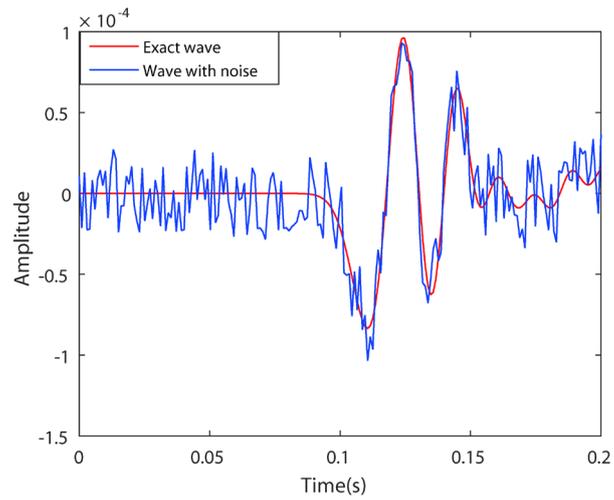


(c)

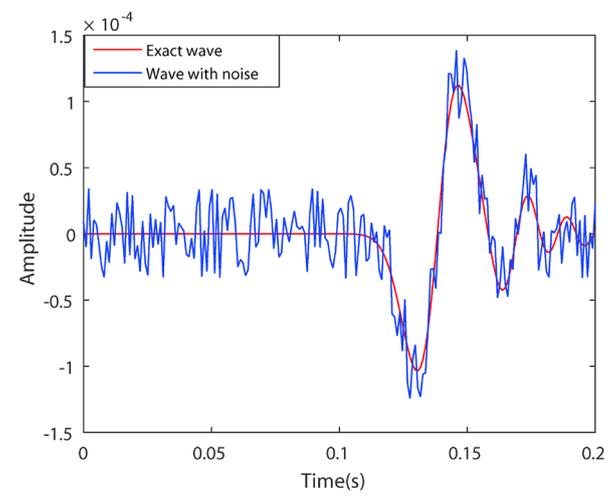
Figure 9. Noisy data with SNR = 10 received in the right well at three different positions z_r . (a) $z_r = N_z/2$; (b) $z_r = 3N_z/4$; (c) $z_r = N_z$.



(a)



(b)



(c)

Figure 10. Noisy data with SNR = 1 received in the right well at three different positions z_r . (a) $z_r = N_z/2$; (b) $z_r = 3N_z/4$; (c) $z_r = N_z$.

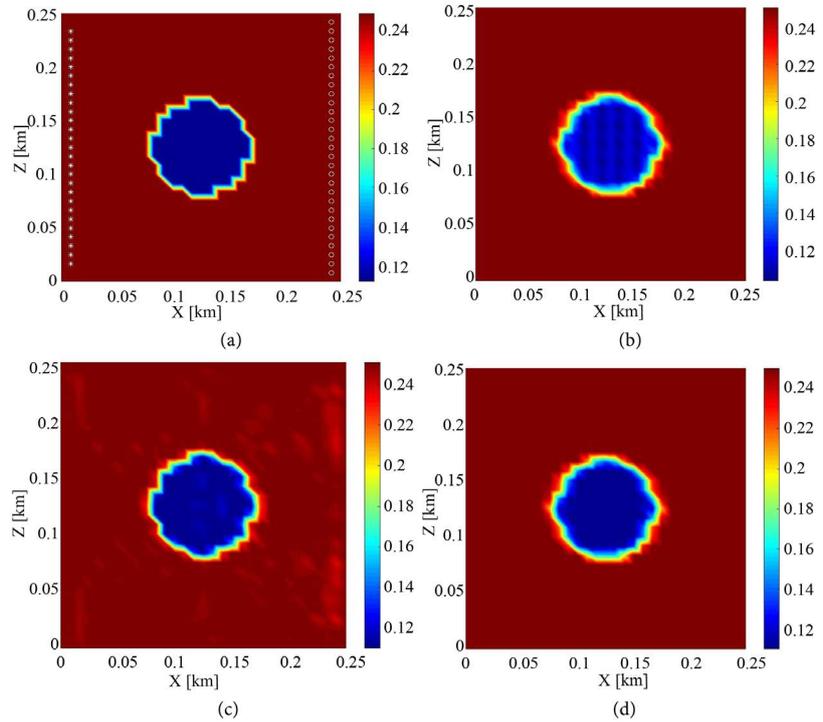


Figure 11. Exact model and inverse results with three different methods after 1000 iterations for noisy data with SNR = 100. (a) Exact model; (b) TV regularization; (c) Bound constraints; (d) Both TV regularization and bound constraints.

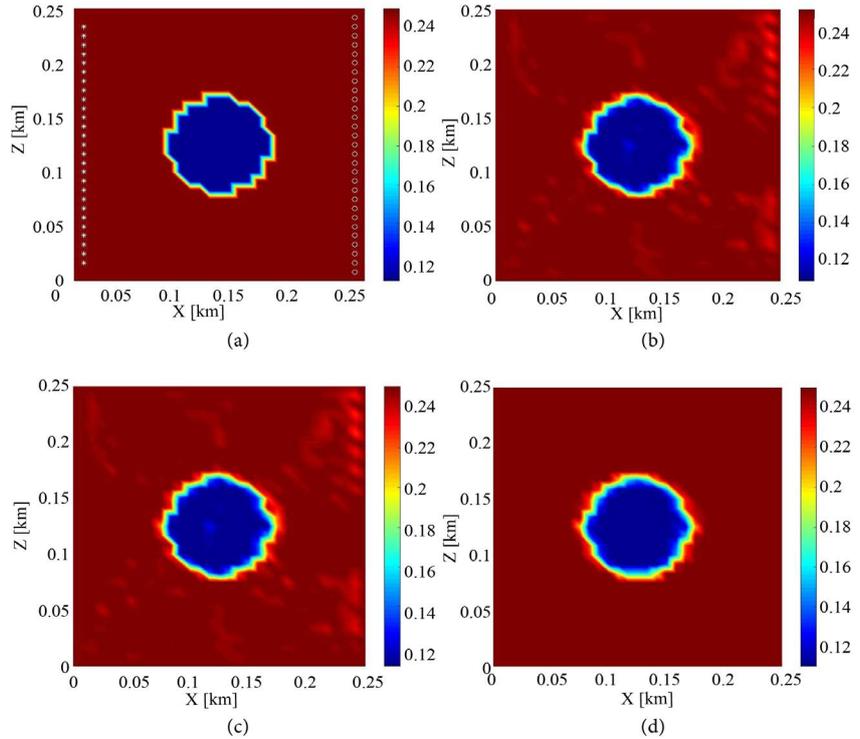


Figure 12. Exact model and inverse results with three different methods after 1000 iterations for noisy data with SNR = 10. (a) Exact model; (b) TV regularization; (c) Bound constraints; (d) Both TV regularization and bound constraints.

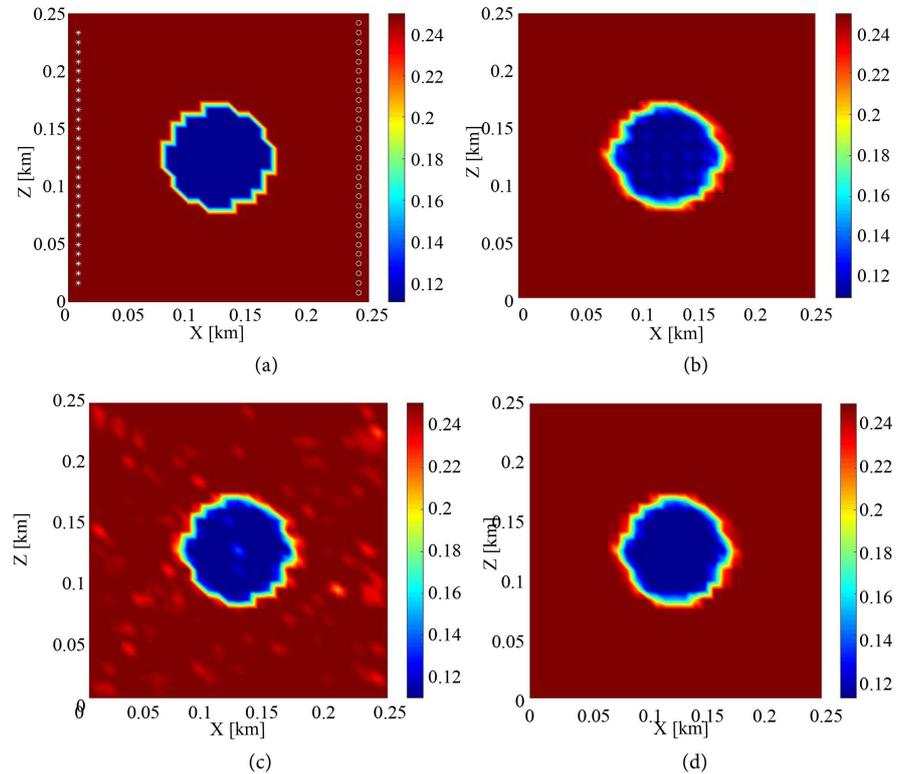


Figure 13. Exact model and inverse results with three different methods after 1000 iterations for noisy data with $\text{SNR} = 1$. (a) Exact model; (b) TV regularization; (c) Bound constraints; (d) Both TV regularization and bound constraints.

4. Conclusion

Full waveform inversion is an effective method for recovering the media parameter. It is an optimization iterative process by minimizing the misfit between the recorded and synthesized data. In this paper, we have developed a new numerical method for crosshole waveform slowness inversion. This method combines TV regularization with constrained optimization together. The TV regularization method can yield better images of slowness at its discontinuities while the bound constraints from logging data can give more reliable constraints to keep the inversion converging to the right result. The combination is necessary to improve the robustness and inversion precision. The novel skip method is proposed to implement the constrained optimization problem effectively. Numerical computations both for noise free data and noisy data with lower SNR show the effectiveness and robustness of the proposed method. In the future, we will design preconditioners to improve the computational efficiency further.

Acknowledgements

This work is supported by China National Natural Science Foundation under grand number 11471328. It is also supported by the Chinese Academy of Sciences and World Academy of Sciences for CAS-TWAS fellowship and the National Center for Mathematics and Interdisciplinary Sciences, Chinese

Academy of Sciences. The computations are completed in the State Key Laboratory of Scientific and Engineering, ICMSEC.

References

- [1] Tarantola, A. (1984) Inversion of Seismic Reflection Data in the Acoustic Approximation. *Geophysics*, **49**, 1259-1266. <https://doi.org/10.1190/1.1441754>
- [2] Gauthier, O., Virieux, J. and Tarantola, A. (1986) Two-Dimensional Nonlinear Inversion of Seismic Waveforms: Numerical Results. *Geophysics*, **51**, 1387-1403. <https://doi.org/10.1190/1.1442188>
- [3] Tarantola, A. (1986) A Strategy for Nonlinear Inversion of Seismic Reflection Data. *Geophysics*, **51**, 1893-1903. <https://doi.org/10.1190/1.1442046>
- [4] Mora, P. (1987) Nonlinear Two-Dimensional Elastic Inversion of Multioffset Seismic Data. *Geophysics*, **52**, 1211-1228. <https://doi.org/10.1190/1.1442384>
- [5] Mora, P. (1988) Elastic Wavefield Inversion of Reflection and Transmission Data. *Geophysics*, **53**, 750-759. <https://doi.org/10.1190/1.1442510>
- [6] Pratt, R.G. (1990) Frequency-Domain Elastic Wave Modeling by Finite Differences: A Tool for Crosshole Seismic Imaging. *Geophysics*, **55**, 626-632. <https://doi.org/10.1190/1.1442874>
- [7] Pratt, R.G. and Worthington, M.H. (1990) Inverse Theory Applied to Multi-Source Cross-Hole Tomography, Part 1: Acoustic Wave Equation Method. *Geophysical Prospecting*, **38**, 287-310. <https://doi.org/10.1111/j.1365-2478.1990.tb01846.x>
- [8] Pratt, R.G. (1990) Inverse Theory Applied to Multi-Source Cross-Hole Tomography, Part 2: Elastic Wave Equation Method. *Geophysical Prospecting*, **38**, 311-329. <https://doi.org/10.1111/j.1365-2478.1990.tb01847.x>
- [9] Song, Z.M., Williamson, P.R. and Pratt, R.G. (1995) Frequency-Domain Acoustic-Wave Modeling and Inversion of Corsshole Data: Part II Inversion Method, Synthetic Experiments and Real-Data Results. *Geophysics*, **3**, 796-809. <https://doi.org/10.1190/1.1443818>
- [10] Bunks, C., Saleck, F., Zaleski, S. and Chavent, G. (1995) Multiscale Seismic Waveform Inversion. *Geophysics*, **60**, 1457-1473. <https://doi.org/10.1190/1.1443880>
- [11] Sirgue, L. and Pratt, R.G. (2004) Efficient Waveform Inversion and Imaging: A Strategy for Selecting Temporal Frequencies. *Geophysics*, **69**, 231-248. <https://doi.org/10.1190/1.1649391>
- [12] Zhang, W., Luo, J. and Teng, J. (2015) Frequency Multiscale Full-Waveform Velocity Inversion. *Chinese Journal of Geophysics*, **58**, 216-228.
- [13] Signorini, M., Micheletti, S. and Perotto, S. (2016) CMFWI: Coupled Multiscenario Full Waveform Inversion. *Inverse Problems in Science and Engineering*, **25**, 939-964. <https://doi.org/10.1080/17415977.2016.1209746>
- [14] Guitton, A. (2012) Blocky Regularization Schemes for Full-Waveform Inversion. *Geophysical Prospecting*, **60**, 870-884.
- [15] Shin, C. and Cha, Y.H. (2008) Waveform Inversion in the Laplace Domain. *Geophysical Journal International*, **173**, 922-931. <https://doi.org/10.1111/j.1365-246X.2008.03768.x>
- [16] Shin, C. and Cha, Y.H. (2009) Waveform Inversion in the Laplace-Fourier Domain. *Geophysical Journal International*, **177**, 1067-1079. <https://doi.org/10.1111/j.1365-246X.2009.04102.x>
- [17] Ha, T. and Shin, C. (2012) Laplace-Domain Full-Waveform Inversion of Seismic

- Data Lacking Low-Frequency Information. *Geophysics*, **77**, R199-R206.
<https://doi.org/10.1190/geo2011-0411.1>
- [18] Wu, R.S., Luo, J. and Wu, B. (2014) Seismic Envelop Inversion and Modulation Signal Model. *Geophysics*, **79**, WA13-WA24.
<https://doi.org/10.1190/geo2013-0294.1>
- [19] Virieux, J. and Operto, S. (2009) An Overview of Full-Waveform Inversion in Exploration Geophysics. *Geophysics*, **74**, WCC127-WCC152.
<https://doi.org/10.1190/1.3238367>
- [20] Choi, Y. and Alkhalifah, T. (2018) Time-Domain Full-Waveform Inversion of Exponentially Damped Wavefield Using the Deconvolution-Based Objective Function. *Geophysics*, **83**, R77-R88. <https://doi.org/10.1190/geo2017-0057.1>
- [21] Manukyan, E., Maurer, H. and Nuber, A. (2018) Improvements to Elastic Full-Waveform Inversion Using Cross-Gradient Constraints. *Geophysics*, **83**, R105-R115. <https://doi.org/10.1190/geo2017-0266.1>
- [22] Capdeville, Y. and Metivier, L. (2018) Elastic Full Waveform Inversion Based on the Homogenization Method: Theoretical Framework and 2-D Numerical Illustrations. *Geophysical Journal International*, **213**, 1093-1112.
<https://doi.org/10.1093/gji/ggy039>
- [23] Virieux, J. (1986) P-SV Wave Propagation in Heterogeneous Media: Velocity-Stress Finite-Difference Method. *Geophysics*, **51**, 889-901.
<https://doi.org/10.1190/1.1442147>
- [24] Clayton, R. and Engquist, B. (1977) Absorbing Boundary Conditions for Acoustic and Elastic Wave Equations. *Bulletin of the Seismological Society of America*, **67**, 1529-1540.
- [25] Grote, M.J. and Keller, J.B. (2000) Exact Nonreflecting Boundary Condition for Elastic Waves. *SIAM Journal on Applied Mathematics*, **60**, 803-819.
<https://doi.org/10.1137/S0036139998344222>
- [26] Zhang, W., Tong, L. and Chung, E.T. (2014) Exact Nonreflecting Boundary Conditions for Three Dimensional Poroelastic Wave Equations. *Communications in Mathematical Sciences*, **12**, 61-98. <https://doi.org/10.4310/CMS.2014.v12.n1.a4>
- [27] Berenger, J.P. (1994) A Perfectly Matched Layer for the Absorption of Electromagnetic Waves. *Journal of Computational Physics*, **114**, 185-200.
<https://doi.org/10.1006/jcph.1994.1159>
- [28] Collino, F. and Tsogka, C. (2001) Application of the Perfectly Matched Absorbing Layer Model to the Linear Elastodynamic Problem in Anisotropic Heterogeneous Media. *Geophysics*, **66**, 294-307. <https://doi.org/10.1190/1.1444908>
- [29] Komatitsch, D. and Tromp, J. (2003) A Perfectly Matched Layer Absorbing Boundary Condition for the Second-Order Seismic Wave Equation. *Geophysical Journal International*, **154**, 146-153. <https://doi.org/10.1046/j.1365-246X.2003.01950.x>
- [30] Royden, R.L. (1968) Real Analysis. 2nd Edition, MacMillan, New York.
- [31] Acar, R. and Vogel, C.R. (1994) Analysis of Total Variation Penalty Methods for Ill-Posed Problems. *Inverse Problem*, **10**, 1217-1229.
<https://doi.org/10.1088/0266-5611/10/6/003>
- [32] Vogel, C.R. (2002) Computational Methods for Inverse Problems. SIAM, Philadelphia.
- [33] Engl, H.W., Hanke, M. and Neubauer, A. (2000) Regularization of Inverse Problems. Kluwer Academic Publishers, Dordrecht/Boston/London.
- [34] Barzilai, J. and Borwein, J. (1988) Two-Point Step Size Gradient Method. *IMA*

Journal of Numerical Analysis, **8**, 141-148. <https://doi.org/10.1093/imanum/8.1.141>

- [35] Raydan, M. (1997) The Barzilai and Borwein Gradient Method for the Large Scale Unconstrained Minimization Problem. *SIAM Journal on Optimization*, **7**, 26-33. <https://doi.org/10.1137/S1052623494266365>
- [36] Qi, L., Teo, K. and Yang, X. (2005) *Optimization and Control with Applications*. Springer Science and Business Media, Inc., New York. <https://doi.org/10.1007/b104943>
- [37] Nocedal, J. and Wright, S. (1999) *Numerical Optimization*. Springer Science and Business Media, Inc., New York. <https://doi.org/10.1007/b98874>

Appendix A. The Stability for the Scheme (6)-(8)

We apply the Fourier method to analyze the stability. Since the source term in Equation (8) has no influence on stability, we consider the source-free formulation of the scheme (6)-(8) for brevity, *i.e.*,

$$(w_x)_{i+\frac{1}{2},j}^{n+\frac{1}{2}} - (w_x)_{i+\frac{1}{2},j}^{n-\frac{1}{2}} = r_x \delta_x u_{i+\frac{1}{2},j}^n, \tag{36}$$

$$(w_z)_{i,j+\frac{1}{2}}^{n+\frac{1}{2}} - (w_z)_{i,j+\frac{1}{2}}^{n-\frac{1}{2}} = r_z \delta_z u_{i,j+\frac{1}{2}}^n, \tag{37}$$

$$u_{i,j}^{n+1} - u_{i,j}^n = \frac{1}{\sigma} \left(r_x \delta_x (w_x)_{i,j}^{n+\frac{1}{2}} + r_z \delta_z (w_z)_{i,j}^{n+\frac{1}{2}} \right), \tag{38}$$

where $r_x = \Delta t / \Delta x$, $r_z = \Delta t / \Delta z$, and δ_x and δ_z are the central difference operators in x and z directions respectively. Noting that at time $n - 1/2$ we have

$$u_{i,j}^n - u_{i,j}^{n-1} = \frac{1}{\sigma} \left(r_x \delta_x (w_x)_{i,j}^{n-\frac{1}{2}} + r_z \delta_z (w_z)_{i,j}^{n-\frac{1}{2}} \right). \tag{39}$$

Subtracting (39) from (38) and applying (36)-(37), we obtain

$$u_{i,j}^{n+1} - 2u_{i,j}^n + u_{i,j}^{n-1} = \frac{1}{\sigma} \left(r_x^2 \delta_x (\delta_x u_{i,j}^n) + r_z^2 \delta_z (\delta_z u_{i,j}^n) \right). \tag{40}$$

Let $I = \sqrt{-1}$ and suppose k_x and k_z are the wavenumber in the x and z directions respectively. By applying $u_{i,j}^n = \tilde{u}^n e^{ik_x i} e^{jk_z j}$ to Equation (40), we have

$$Z^{n+1} = AZ^n, \tag{41}$$

where

$$A = \begin{pmatrix} 2 + (p_x r_x^2 + p_z r_z^2) / \sigma & -1 \\ 1 & 0 \end{pmatrix}, \tag{42}$$

$$Z^{n+1} = \begin{pmatrix} \tilde{u}^{n+1} \\ \tilde{u}^n \end{pmatrix}, \tag{43}$$

and

$$p_x = -4 \sin^2 \left(\frac{k_x}{2} \right), \quad p_z = -4 \sin^2 \left(\frac{k_z}{2} \right). \tag{44}$$

The sufficient and necessary condition for stability is that the eigenvalues of matrix A is within or on the unit circle. The characteristic equation of matrix A is

$$\lambda^2 - \left(2 + \frac{p_x r_x^2 + p_z r_z^2}{\sigma} \right) \lambda + 1 = 0. \tag{45}$$

The stability requires

$$\left| 2 + \frac{p_x r_x^2 + p_z r_z^2}{\sigma} \right| \leq 2, \tag{46}$$

which gives

$$\Delta t^2 \leq -4\sigma \left(\frac{p_x}{\Delta x^2} + \frac{p_z}{\Delta z^2} \right)^{-1}, \tag{47}$$

or

$$\Delta t \leq \sqrt{\sigma \left(\frac{1}{\Delta x^2} + \frac{1}{\Delta z^2} \right)}, \tag{48}$$

which is just the stability (11).

Appendix B. The Derivation for the Gradient

For convenience, we denote the first term in Equation (22) by

$$\chi(\sigma) = \frac{1}{2} \sum_{shot} \int_T \int_{\Omega} (u - u_{obs})^2 \delta(x - x_r) dx dz dt, \tag{49}$$

and the forward operator A by

$$Au = f \cdot \delta(x - x_s) \delta(z - z_s), \tag{50}$$

where A is given by

$$A = \sigma \frac{\partial^2}{\partial t^2} - \nabla \cdot \nabla. \tag{51}$$

The operator A is a self-conjugate operator. See the following **Lemma 1**.

Lemma 1. For any $\phi, u \in L^2(0, T; H_0^1(\Omega))$ satisfying

$$\phi|_{t=T} = 0, \frac{\partial \phi}{\partial t}|_{t=T} = 0; u|_{t=0} = 0, \frac{\partial u}{\partial t}|_{t=0} = 0, \tag{52}$$

where H_0^1 is the Sobolev space, then operator A is the self-adjoint operator, i.e.,

$$\int_T \int_{\Omega} \left(\sigma \frac{\partial^2 u}{\partial t^2} - \nabla \cdot \nabla u \right) \phi dx dz dt = \int_T \int_{\Omega} u \left(\sigma \frac{\partial^2 \phi}{\partial t^2} - \nabla \cdot \nabla \phi \right) dx dz dt. \tag{53}$$

Proof. Based on the definition of the operator A, we have

$$\begin{aligned} \int_T \int_{\Omega} (Au) \phi dx dz dt &= \int_T \int_{\Omega} \left(\sigma \frac{\partial^2 u}{\partial t^2} - \nabla \cdot \nabla u \right) \phi dx dz dt \\ &= \int_T \int_{\Omega} \sigma \frac{\partial^2 u}{\partial t^2} \phi dx dz dt - \int_T \int_{\Omega} (\nabla \cdot \nabla u) \phi dx dz dt. \end{aligned} \tag{54}$$

Applying the Green formula in time and using the conditions (52), the first term in Equation (54) becomes

$$\begin{aligned} &\int_{\Omega} dx dz \int_T \sigma \frac{\partial^2 u}{\partial t^2} \phi dt \\ &= \int_{\Omega} \sigma dx dz \left(\frac{\partial u}{\partial t} \phi|_0^T - \int_T \frac{\partial u}{\partial t} \frac{\partial \phi}{\partial t} dt \right) \\ &= \int_{\Omega} \sigma dx dz \left(\frac{\partial u}{\partial t} \phi|_0^T - u \frac{\partial \phi}{\partial t} \Big|_0^T + \int_T u \frac{\partial^2 \phi}{\partial t^2} dt \right) \\ &= \int_{\Omega} \sigma dx dz \left(\int_T u \frac{\partial^2 \phi}{\partial t^2} dt \right). \end{aligned} \tag{55}$$

Similarly, applying the Green formula in space and noting boundary conditions of u and ϕ , the second term in Equation (54) becomes

$$\begin{aligned} & \int_T dt \int_{\Omega} (\nabla \cdot \nabla u) \phi dx dz \\ &= \int_T dt \left(\frac{\partial u}{\partial n} \phi \Big|_{\partial \Omega} - \int_{\Omega} \nabla u \cdot \nabla \phi dx dz \right) \\ &= \int_T dt \left(\frac{\partial u}{\partial n} \phi \Big|_{\partial \Omega} - \frac{\partial \phi}{\partial n} u \Big|_{\partial \Omega} + \int_{\Omega} u \nabla \cdot \nabla \phi dx dz \right) \\ &= \int_T dt \left(\int_{\Omega} u \nabla \cdot \nabla \phi dx dz \right). \end{aligned} \tag{56}$$

Inserting Equations (55) and (56) into Equation (54), we have

$$\int_T \int_{\Omega} \left(\sigma \frac{\partial^2 u}{\partial t^2} - \nabla \cdot \nabla u \right) \phi dx dz dt = \int_T \int_{\Omega} u \left(\sigma \frac{\partial^2 \phi}{\partial t^2} - \nabla \cdot \nabla \phi \right) dx dz dt.$$

The proof is complete. Making variation for the objective function (49), we obtain

$$\begin{aligned} \delta \chi(\sigma) &= \chi(\sigma + \delta \sigma) - \chi(\sigma) \\ &= \sum_{shot} \int_T \int_{\Omega} (u - u_{obs}) \delta(x - x_r) \delta u dx dz dt, \end{aligned} \tag{57}$$

where

$$\delta u = u(\sigma + \delta \sigma) - u(\sigma). \tag{58}$$

Note that $u(\sigma + \delta \sigma), u(\sigma) \in L^2(0, T; H_0^1(\Omega))$ and satisfy initial conditions

$$\begin{aligned} u(\sigma) \Big|_{t=0} &= 0, \quad \frac{\partial u(\sigma)}{\partial t} \Big|_{t=0} = 0, \\ \frac{\partial u(\sigma + \delta \sigma)}{\partial t} \Big|_{t=0} &= 0, \quad u(\sigma + \delta \sigma) \Big|_{t=0} = 0, \end{aligned} \tag{59}$$

we obtain $\delta u \in L^2(0, T; H_0^1(\Omega))$ and δu satisfies $\delta u \Big|_{t=0} = \frac{\partial \delta u}{\partial t} \Big|_{t=0} = 0$.

Theorem 2 If ϕ is the solution of the following backward problem

$$\sigma \frac{\partial^2 \phi}{\partial t^2} - \nabla \cdot \nabla \phi = -(u - u_{obs}) \cdot \delta(x - x_r), \tag{60}$$

$$\phi \Big|_{t=T} = 0, \quad \frac{\partial \phi}{\partial t} \Big|_{t=T} = 0, \tag{61}$$

then the gradient of the objective function (49) is

$$\frac{\partial \chi}{\partial \sigma} = - \sum_{shot} \int_T \int_{\Omega} \phi \frac{\partial^2 u}{\partial t^2} dx dz dt. \tag{62}$$

Proof Based on Equation (57) and the definition of ϕ by Equations (60)-(61), we have

$$\begin{aligned} \delta \chi(\sigma) &= \sum_{shot} \int_T \int_{\Omega} (u - u_{obs}) \delta(x - x_r) \delta u dx dz dt \\ &= \sum_{shot} \int_T \int_{\Omega} \left(-\sigma \frac{\partial^2 \phi}{\partial t^2} + \nabla \cdot \nabla \phi \right) \delta u dx dz dt, \end{aligned} \tag{63}$$

Using boundary conditions and **Lemma 1**, we obtain

$$\delta\chi(\sigma) = \sum_{shot} \int_T \int_{\Omega} \phi \left(-\sigma \frac{\partial^2 \delta u}{\partial t^2} + \nabla \cdot \nabla \delta u \right) dx dz dt. \quad (64)$$

From the definition we know $u(\sigma)$ and $u(\sigma + \delta\sigma)$ and satisfy

$$-(\sigma + \delta\sigma) \frac{\partial^2 u(\sigma + \delta\sigma)}{\partial t^2} + \nabla \cdot \nabla u(\sigma + \delta\sigma) = -f(t) \cdot \delta(x - x_s) \delta(z - z_s), \quad (65)$$

$$-\sigma \frac{\partial^2 u(\sigma)}{\partial t^2} + \nabla \cdot \nabla u(\sigma) = -f(t) \cdot \delta(x - x_s) \delta(z - z_s). \quad (66)$$

Subtracting Equation (66) from equation (65), we have

$$\left(-\sigma \frac{\partial^2 \delta u}{\partial t^2} + \nabla \cdot \nabla \delta u \right) + \delta\sigma \frac{\partial^2 u(\sigma + \delta\sigma)}{\partial t^2} = 0. \quad (67)$$

Inserting Equation (67) into Equation (64), we have

$$\delta\chi(\sigma) = \sum_{shot} \int_T \int_{\Omega} \phi \left(-\delta\sigma \frac{\partial^2 u(\sigma + \delta\sigma)}{\partial t^2} \right) dx dz dt. \quad (68)$$

Let $\delta\sigma \rightarrow 0$, we obtain the gradient

$$\frac{\partial \chi}{\partial \sigma} = - \sum_{shot} \int_T \int_{\Omega} \phi \frac{\partial^2 u}{\partial t^2} dx dz dt. \quad (69)$$

The proof is completed. The result (69) is just the first term on the right hand side of Equation (25). We can see that there is no Born approximation in the derivation above.