

Full-waveform Velocity Inversion Based on the Acoustic Wave Equation

Wensheng Zhang, Jia Luo

Institute of Computational Mathematics and Scientific/Engineering Computing, LSEC
Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, P. R. China
Email: zws@lsec.cc.ac.cn, luojia@lsec.cc.ac.cn

Received June, 2013

ABSTRACT

Full-waveform velocity inversion based on the acoustic wave equation in the time domain is investigated in this paper. The inversion is the iterative minimization of the misfit between observed data and synthetic data obtained by a numerical solution of the wave equation. Two inversion algorithms in combination with the CG method and the BFGS method are described respectively. Numerical computations for two models including the benchmark Marmousi model with complex structure are implemented. The inversion results show that the BFGS-based algorithm behaves better in inversion than the CG-based algorithm does. Moreover, the good inversion result for Marmousi model with the BFGS-based algorithm suggests the quasi-Newton methods can provide an important tool for large-scale velocity inversion. More computations demonstrate the correctness and effectiveness of our inversion algorithms and code.

Keywords: Finite Difference; Acoustic Wave Equation; Full-waveform Inversion; CG Method; BFGS Method; Marmousi Model

1. Introduction

The velocity of underground media is an important parameter for resources exploration. It may indicate the structure and position of reservoir. Full-waveform inversion uses travel time, amplitude and phase information simultaneously to inverse media velocity. It is a powerful way in reconstructing complex velocity structures. The inversion can be performed in the time-space domain [1-4] or in the frequency-space domain [5-9]. The frequency-domain inversion approach is equivalent to the time-domain inversion approach if all of the frequency data components are used in the inversion process [9].

Full-waveform inversion can be formulated as a nonlinear optimization problem with an objective function consisting of the difference between observed and synthetic wave field in a suitable norm, usually L_2 norm. This choice results in least-squares formulation. However, solving full-waveform inversion problems using nonlinear optimization methods remains a challenge. This is mostly because of mathematical and numerical difficulties, which include multiple minima and ill-posedness. The objective function is relatively insensitive to wave number components of the model that are shorter than wavelength. The eigenvalues of the linearized forward operator corresponding to those components are nearly zero. Recently, the technique of multifrequency inversion of full-waveform has been investigated. For frequency

domain inversion, the important issue is how to choose the various frequencies for inversion [10]. Moreover, the wave modeling in the frequency domain requires solving a large-scale system at each iteration which is time-consuming for large-scale problem. Comparing with the frequency-domain inversion methods, the time-domain methods have the advantage of high efficiency in forward modeling.

The algorithms used to solve the optimization problem include linearized approximation, the steepest descent algorithm, the conjugate gradient (CG) [11] and the quasi-Newton methods [12-15]. However, these methods behave differently as their own characters and the essential ill-posedness of the problem. In this paper, we implement full-waveform inversion with the CG method which is a typical gradient-based method, and with the BFGS method which is a representative of the quasi-Newton methods. As to the forward method, the finite-difference method is employed as its high efficiency. The discretization schemes including the absorbing boundary conditions and corner conditions are given.

2. Theoretical Methods

2.1. Finite Difference Simulation

The wave equations can be solved in either the time or the frequency domain. The usual numerical methods

such as the finite difference method, the finite volume method and the finite element method can all be employed. Here we use the finite difference method for its high efficiency. The 2-D acoustic wave equation in the time domain can be written as

$$\frac{1}{v(x, z)^2} \frac{\partial^2 u}{\partial t^2} - \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial z^2} \right) = f(t) \delta(x - x_s) \delta(z - z_s), \quad (1)$$

where $u(x, z, t)$ is the pressure, $v(x, z)$ is the velocity, $f(t)$ is the source time function, (x_s, z_s) is the source location. Usually the initial conditions are zero, i.e.

$$u(x, z, t=0) = 0, \quad u_t(x, z, t=0) = 0. \quad (2)$$

Let $u_{n,m}^l$ be the pressure at the time $l\Delta t$ and at the spatial position $(n\Delta x, m\Delta z)$, and $v_{n,m}$ be the velocity at $(n\Delta x, m\Delta z)$, then the difference scheme of (1) is

$$\begin{aligned} u_{n,m}^{l+1} &= u_{n,m}^{l-1} - 2u_{n,m}^l \\ &+ v_{n,m}^2 \Delta t^2 \left(\frac{u_{n+1,m}^l - 2u_{n,m}^l + u_{n-1,m}^l}{\Delta x^2 r} \right. \\ &\quad \left. + \frac{u_{n,m+1}^l - 2u_{n,m}^l + u_{n,m-1}^l}{\Delta z^2 r} \right) \\ &+ f(l\Delta t) \delta(n - n_s) \delta(m - m_s), \end{aligned} \quad (3)$$

with the initial conditions

$$u_{n,m}^{-1} = 0, \quad u_{n,m}^0 = 0 \quad (4)$$

The numerical stability condition for (3) and (4) is

$$\min(\Delta x, \Delta z) > \sqrt{2} \Delta t \max(v). \quad (5)$$

In the wave simulation, the unsuitable choice of time and spatial steps may cause serve dispersion and wave distortion. In order to suppress numerical dispersion, it is usually required that there are 10 sampling points in a wavelength for the second-order difference scheme, i.e.,

$$\min(\Delta x, \Delta z) < \frac{\min(v)}{10 f_{\max}}. \quad (6)$$

Obviously for the same frequency the numerical dispersion increases as the spatial steps rise. However, the dispersion can be depressed if the high-order accurate computational schemes are used.

In numerical computations, the computational domain is usually truncated to a finite region. So absorbing boundary conditions (ABCs) are required to absorb boundary reflections of outgoing waves. A number of ABCs have been introduced in wave simulation, for example, the paraxial approximation method [16] and the perfectly matched layer (PML) method [17]. The PML technique was proposed for electromagnetic fields equations, but can be applied to other wave equations and can fit irregular boundary.

Suppose the computational domain is a rectangular area

$$\Omega = \{(x, z), 0 \leq x \leq X, 0 \leq z \leq Z\}.$$

We set the boundary condition on surface ($z=0$) be the Dirichlet boundary condition as the data are gathered near the surface in our problem. For the boundary conditions of other three sides and four corners, we select the following Clayton and Engquist [16] conditions.

For the top boundary:

$$u(x, z=0, t) = 0. \quad (7)$$

For the bottom boundary:

$$\left(\frac{\partial^2 u}{\partial z \partial t} + \frac{1}{v} \frac{\partial^2 u}{\partial t^2} - \frac{v}{2} \frac{\partial^2 u}{\partial x^2} \right) (x, z=Z, t) = 0. \quad (8)$$

For the left boundary:

$$\left(\frac{\partial^2 u}{\partial x \partial t} - \frac{1}{v} \frac{\partial^2 u}{\partial t^2} + \frac{v}{2} \frac{\partial^2 u}{\partial z^2} \right) (x=0, z, t) = 0. \quad (9)$$

For the right boundary:

$$\left(\frac{\partial^2 u}{\partial x \partial t} + \frac{1}{v} \frac{\partial^2 u}{\partial t^2} - \frac{v}{2} \frac{\partial^2 u}{\partial z^2} \right) (x=X, z, t) = 0. \quad (10)$$

The absorbing conditions for four corners can be derived through the 45° rotation of the ABCs above. The results are the following:

For the left-upper corner:

$$\left(-\frac{\partial u}{\partial x} - \frac{\partial u}{\partial z} + \frac{\sqrt{2}}{v} \frac{\partial u}{\partial t} \right) = 0. \quad (12)$$

For the right-upper corner:

$$\left(\frac{\partial u}{\partial x} - \frac{\partial u}{\partial z} + \frac{\sqrt{2}}{v} \frac{\partial u}{\partial t} \right) = 0. \quad (13)$$

For the left-down corner:

$$\left(-\frac{\partial u}{\partial x} + \frac{\partial u}{\partial z} + \frac{\sqrt{2}}{v} \frac{\partial u}{\partial t} \right) = 0. \quad (14)$$

For the right-down corner:

$$\left(\frac{\partial u}{\partial x} + \frac{\partial u}{\partial z} + \frac{\sqrt{2}}{v} \frac{\partial u}{\partial t} \right) = 0. \quad (15)$$

The difference schemes for the absorbing boundary condition (7)-(10) can also be obtained by the second-order central finite difference method, i.e.,

$$u_{n,m}^{l+1} = 0, \quad (16)$$

$$\frac{\delta_z^- \delta_0^+ u_{n,m}^l}{2\Delta t \Delta z} + \frac{1}{2\Delta t^2 v_{i,j}} \delta_+^t \delta_-^t (u_{n,m}^l + u_{n,m-1}^l) \quad (17)$$

$$- \frac{v_{i,j}}{4\Delta x^2} \delta_+^x \delta_-^x (u_{n,m-1}^{l+1} + u_{n,m}^{l-1}) = 0,$$

$$\frac{\delta_+^x \delta_0^t u_{n,m}^l}{2\Delta t \Delta x} - \frac{1}{2\Delta t^2 v_{i,j}} \delta_+^t \delta_-^t (u_{n,m}^l + u_{n+1,m}^l) \quad (18)$$

$$+ \frac{v_{i,j}}{4\Delta z^2} \delta_+^z \delta_-^z (u_{n+1,m}^{l+1} + u_{n,m}^{l-1}) = 0,$$

$$\frac{\delta_+^x \delta_0^t u_{n,m}^l}{2\Delta t \Delta x} + \frac{1}{2\Delta t^2 v_{i,j}} \delta_+^t \delta_-^t (u_{n,m}^l + u_{n-1,m}^l) \quad (19)$$

$$- \frac{v_{i,j}}{4\Delta z^2} \delta_+^z \delta_-^z (u_{n-1,m}^{l+1} + u_{n,m}^{l-1}) = 0,$$

where δ_+ and δ_- are the forward and backward difference operators with first-order accuracy respectively, δ_0 is the central difference operator with second-order accuracy, Δx and Δz are the spatial steps in x and z directions respectively. Similarly, the difference schemes for four corners conditions (12)-(15) can be obtained in the following:

For the left-upper corner: $(n, m) = (0, 1)$,

$$u_{n,m}^{l+1} = \frac{r_1 u_{n+1,m}^{l+1} + r_2 u_{n,m+1}^{l+1} + r_3 u_{n,m}^l}{r_1 + r_2 + r_3}. \quad (20)$$

For the right-upper corner: $(n, m) = (N_x - 1, 1)$,

$$u_{n,m}^{l+1} = \frac{r_1 u_{n-1,m}^{l+1} + r_2 u_{n,m+1}^{l+1} + r_3 u_{n,m}^l}{r_1 + r_2 + r_3}. \quad (21)$$

For the left-down corner:

$$(n, m) = (0, N_z - 1), (1, N_z - 1), (0, N_z - 2), \quad (22)$$

$$u_{n,m}^{l+1} = \frac{r_1 u_{n+1,m}^{l+1} + r_2 u_{n,m-1}^{l+1} + r_3 u_{n,m}^l}{r_1 + r_2 + r_3}.$$

For the right-down corner:

$$(n, m) = (N_x - 1, N_z - 1), (N_x - 2, N_z - 1), \quad (23)$$

$$(N_x - 1, N_z - 2),$$

$$u_{n,m}^{l+1} = \frac{r_1 u_{n-1,m}^{l+1} + r_2 u_{n,m-1}^{l+1} + r_3 u_{n,m}^l}{r_1 + r_2 + r_3}.$$

where

$$r_1 = \frac{\Delta t}{\Delta x}, r_2 = \frac{\Delta t}{\Delta z}, r_3 = \frac{\sqrt{2}}{v_{n,m}}. \quad (24)$$

2.2. Two Full-waveform Inversion Algorithms

Full-waveform inversion is minimization of the following objective function

$$f(v) = \frac{1}{2} \|u_{obs} - u_{cal}\|^2 = \frac{1}{2} \|\Delta u\|^2, \quad (25)$$

where u_{obs} is the observed wave field and u_{cal} is the synthetic data. The functional (25) can be written the following discrimination form

$$f(v) = \sum_{n_s=0}^{N_s} \sum_{n_r=0}^{N_r} \sum_{n_t=0}^{N_t} \{u_{obs}(r, n_t, n_s) - u_{cal}(v; n_r, m_r, n_t, n_s)\}^2, \quad (26)$$

where the number of spatial points is $M = N_x \times N_z$, which is also the dimension of unknown velocity v . Assume that the receiver number is N_r and the source number is N_s then the dimension of $u_{obs}(r, n_t, n_s)$ is $N_r \times N_s \times N_t$, here N_t is the time sampling points.

The optimization methods such as nonlinear conjugate gradient and BFGS can be implemented. The gradient-based methods are the commonly used methods in multi-dimensional unconstrained optimization. The CG method is one and it uses the conjugate directions instead of the negative gradient direction as search direction, which can be written as

$$p_k = \begin{cases} -g_k, & k = 0, \\ -g_k + \beta_k p_{k-1}, & k \geq 1. \end{cases} \quad (27)$$

Different choices of β_k lead to different conjugate gradient methods. For example, the FR method [11], the HS method [18] and the PRP method [19, 20]:

$$\beta_k^{FR} = \frac{\|g_{k+1}\|^2}{\|g_k\|^2}, \quad \beta_k^{HS} = \frac{g_{k+1}^T y_k}{p_k^T y_k}, \quad (28)$$

$$\beta_k^{PRP} = \frac{g_{k+1}^T y_k}{\|g_k\|^2}, \quad (29)$$

where $y_k = g_{k+1} - g_k$. Here we adopt the PRP method.

The partial derivative of $f(v)$ with respect to unknown velocity v is

$$\frac{\partial f(v)}{\partial v} = J^T \Delta u, \quad (30)$$

where J is the Fréchet derivative matrix with dimension $M \times (N_r \times N_s)$:

$$J = \left[\frac{\partial u_{cal}}{\partial v} \right] = \left[\frac{\partial u_{cal}}{\partial v_0}, \frac{\partial u_{cal}}{\partial v_1}, \dots, \frac{\partial u_{cal}}{\partial v_{M-1}} \right]. \quad (31)$$

The gradient g is a vector with dimension M , which can be written as

$$g = -J^T \Delta u. \quad (32)$$

For the line search method we may use the Aimijo algorithm [21] or the strong Wolfe algorithm [22, 23]. The Aimijo algorithm requires

$$J(v + \alpha p) \leq J(v) + b_1 \alpha p^T \nabla J(v), \quad (33)$$

where $b_1 \in (0, 1)$ and we select $b_1 = 0.001$ in computations. And the strong Wolfe algorithm requires the following condition besides condition (33) above

$$|p^T \nabla J(v + \alpha p)| \leq -b_2 \alpha p^T \nabla J(v), \quad (34)$$

where $b_2 \in (b_1, 1)$ and we select $b_2 = 0.9$ in computations. Now we give the following full-waveform inversion algorithm combining with the CG method.

Algorithm 1:

1: Give the initial velocity model v_0 and the iterative

stopping condition of the gradient: $\|g\| \leq \varepsilon_{stop}$. Set iteration number $k = 0$;

2: Compute the objective function f_0 and the gradient g . If the gradient satisfies the stopping condition $\|g\| \leq \varepsilon_{stop}$ then stop the iteration. Otherwise set $p_0 = -g_0$;

3: Set $k = k + 1$. If $k < k_{max}$ then continue the iteration;

3.1: Use the strong Wolfe line search algorithm to get the step α_k ;

3.2: Update the velocity model: $v_{k+1} = v_k + \alpha_k p_k$;

3.3: Compute the objective function $f_{k+1}(v_{k+1})$ and the gradient g_{k+1} ;

3.4: If the gradient satisfies the stopping criterion then go out of the loop;

3.5: If the updating quantity of velocity satisfies the stopping condition then goes out of the loop;

3.6: Calculate, $s_k = v_{k+1} - v_k$, $y_k = g_{k+1} - g_k$;

3.7: Calculate β_k^{PRP} and update

$$p_{k+1} : p_{k+1} = -g_{k+1} + \beta_k^{PRP} p_k.$$

Set $k = k + 1$, return step 3.

The algorithm above doesn't use the information of second-order derivate of objective function. The BFGS method is a quasi-Newton method which proposed dependently by Broyden, Fletcher and Goldfarb and Shanno [12-15]. The key of the quasi-Newton method is how to construct B_{k+1} satisfying the following so-called quasi-Newton condition:

$$B_{k+1} s_k = y_k, \quad (35)$$

where $s_k = v_{k+1} - v_k$ and $y_k = g_{k+1} - g_k$. The BFGS method modifies matrix B_{k+1} according to the following expression

$$B_{k+1} = B_k - \frac{B_k s s^T B_k}{s^T B_k s} + \frac{y y^T}{s^T y}, \quad (36)$$

and its inverse is given by

$$H_{k+1} = H_k + \frac{s s^T}{y^T s} \left[\frac{y^T H_k y}{y^T s} + 1 \right] - \frac{1}{y^T s} [s y^T H_k + H_k y s^T], \quad (37)$$

where $H_k = B_k^{-1}$. If H_k is positive definite then H_{k+1} is positive definite as long as $y^T s > 0$. The following algorithm is the full-waveform inversion algorithm based on the BFGS method.

Algorithm 2:

1: Give the initial velocity model v_0 and the iterative terminal condition of the gradient: $\|g\| \leq \varepsilon_{stop}$. Set iteration number $k = 0$;

2: Compute the objective function f_0 and the gradient g . If the gradient satisfies the stopping rule $\|g\| \leq \varepsilon_{stop}$ then stop the iteration. Otherwise set $p_0 = -g_0$;

3: Set $k = k + 1$. If $k < k_{max}$ then continue the iteration;

3.1: Use the strong Wolfe line search algorithm to obtain the step α_k ;

3.2: Update the velocity model: $v_{k+1} = v_k + \alpha_k p_k$;

3.3: Compute the objective function $f_{k+1}(v_{k+1})$ and the gradient g_{k+1} ;

3.4: If the gradient satisfies the stopping condition, then go out of the loop;

3.5: If the updating quantity of velocity satisfies the stopping condition, then go out of the loop;

3.6: Calculate $s_k = v_{k+1} - v_k$, $y_k = g_{k+1} - g_k$;

3.7: Calculate variable scaling factor with

$$\tau_k = \frac{y_k^T s_k}{s_k^T B_k s_k} \quad \text{and let } B_k = \tau_k B_k;$$

3.8: Update B_{k+1} with (36).

3.9: Use the Cholesky method to solve the equation

$$B_{k+1} p_{k+1} = -g_{k+1} \quad \text{to get } p_{k+1}.$$

Set $k = k + 1$, continue with step 3.

In step 3.7 above, the scaling factor τ_k is introduced which can improve convergence [24]. The CG method is only linearly convergent while the BFGS method is superlinear convergent but requires solving a system at each iteration.

3. Numerical Computations

First of all, we test the validity of the finite difference method for wave modeling. **Figure 1** is the wave field received near the surface for a fixed source. The boundary reflections are very serious in **Figure 1** since only Dirichlet boundary conditions are used. **Figure 2** is the wave field corresponding to **Figure 1** but the ABCs are used. We can see the serious boundary reflections in **Figure 1** are eliminated obviously in **Figure 2**.

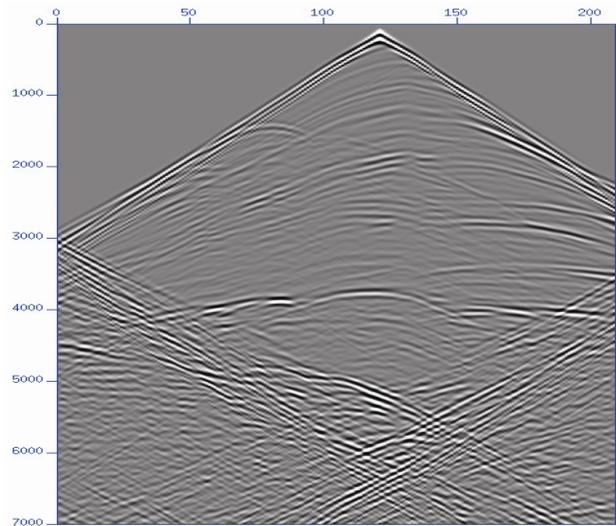


Figure 1. Wavefield received near the surface by the finite difference method with Dirichlet boundary conditions.

Next we test two inversion algorithms for a simple model shown in **Figure 3**. The model has a homogeneous background media with velocity 2000 m/s and two square areas with greater abnormal velocity. The left square media has velocity 2600 m/s and the right square media has the velocity 2100 m/s . The configuration of sources and receivers are shown in **Figure 4**. There are 16 sources totally and 31 receivers for each source. The source step is 10 m and the receiver step is 5 m . The source line and receiver line are 10 m and 5 m below the surface respectively. The spatial discretization points are $N_x = 17$ and $N_z = 33$. The spatial steps are $\Delta x = \Delta z = 5\text{ m}$. The parameter ε_{stop} is 10^{-7} . The maximum number of iterations k_{max} is 200. The initial velocity v_0 is 2000 m/s .

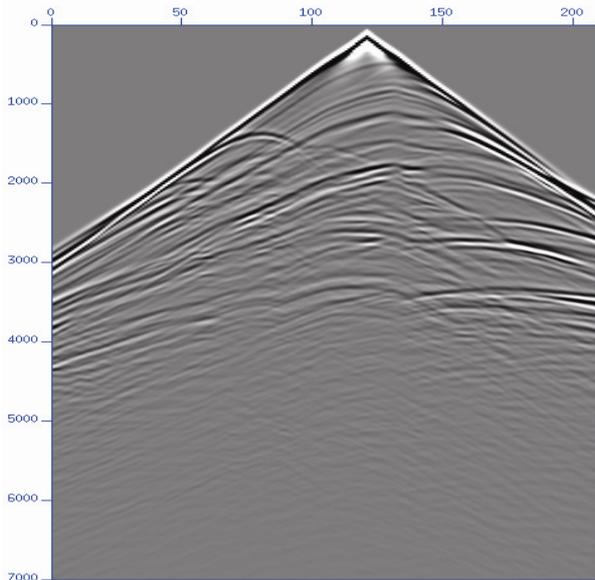


Figure 2. Wavefield received near the surface by the finite difference method with absorbing boundary conditions.



Figure 3. True model for testing two inversion algorithms.

The inversion results with the CG-based algorithm are shown in **Figures 5** and **6**. **Figure 5** is the inversion result after 20 iterations and **Figure 6** is that after 200 iterations. For the BFGS-based algorithm the inversion results are shown in **Figures 7-9**, which are the results after 20 iterations, 100 iterations and 200 iterations, respectively. Comparing **Figures 7-9** with **Figures 5** and **6**, we can see that the BFGS-based algorithm has higher inverse accuracy than the CG-based algorithm. In order to test the ability of the BFGS-based algorithm further, we inverse the Marmousi model next.

The Marmousi model is a benchmark model for testing the ability of migration or inversion methods [25]. The velocity model is shown in **Figure 10**. We can see it has many complex structures. The synthetic data is also generated by the finite difference code. The following computations parameters are adopted: $N_x = 493$ $N_z = 249$,

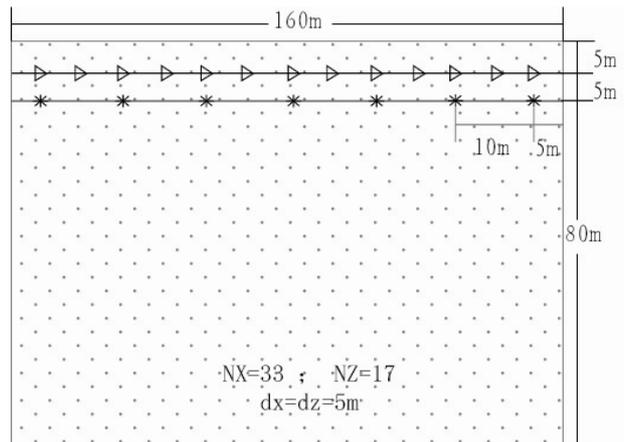


Figure 4. The configuration of sources and receivers for full-waveform inversion.

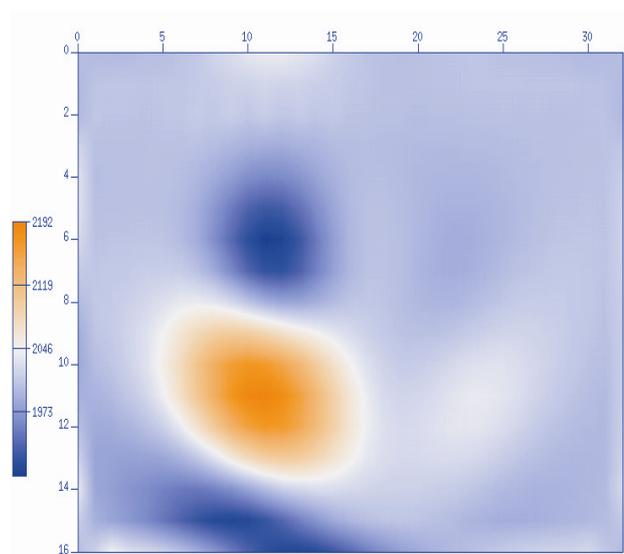


Figure 5. Inversion result after 20 iterations with the CG-based Algorithm 1.

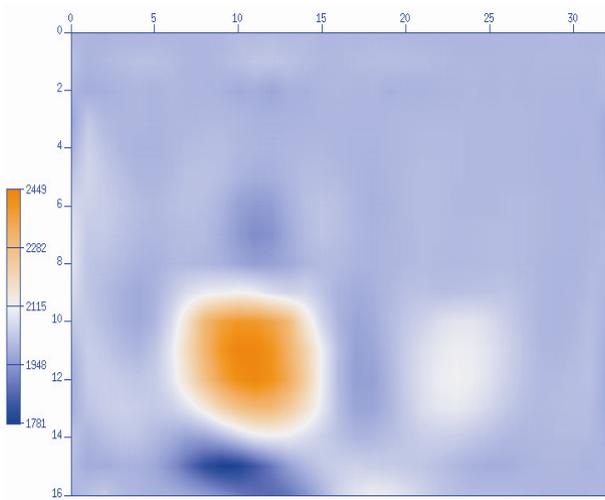


Figure 6. Inversion result after 200 iterations with the CG-based Algorithm 1.

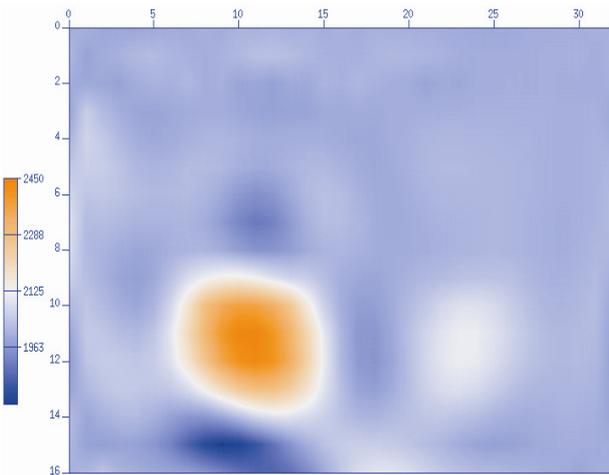


Figure 7. Inversion result after 20 iterations with the BFGS-based Algorithm 2.

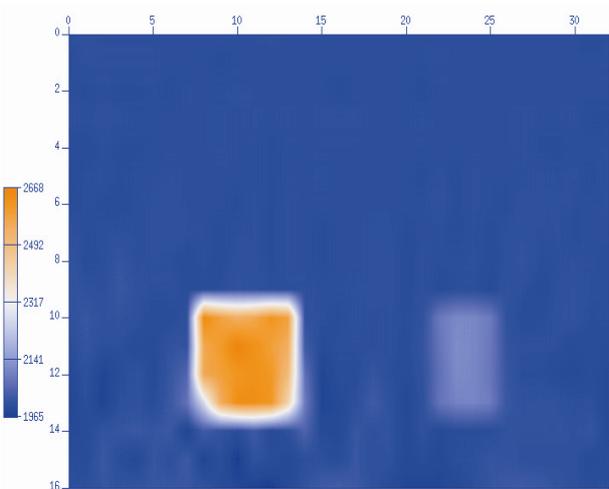


Figure 8. Inversion result after 100 iterations with the BFGS-based Algorithm 2.

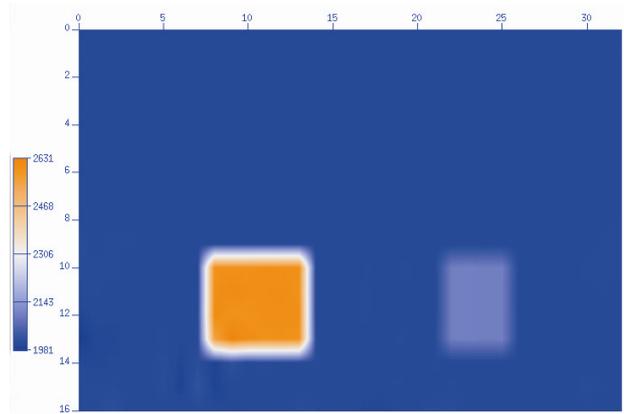


Figure 9. Inversion result after 200 iterations with the BFGS-based Algorithm 2.

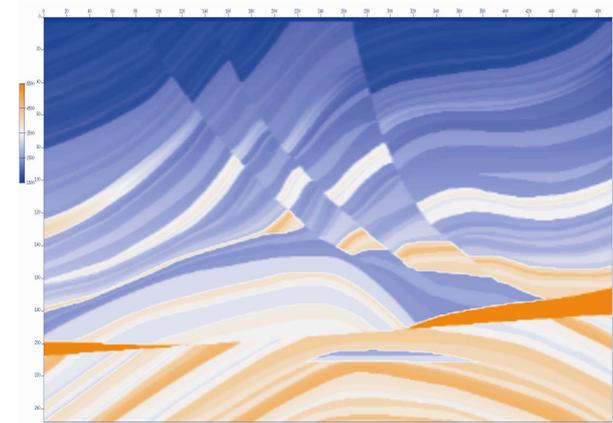


Figure 10. Marmousi model with complex structures.

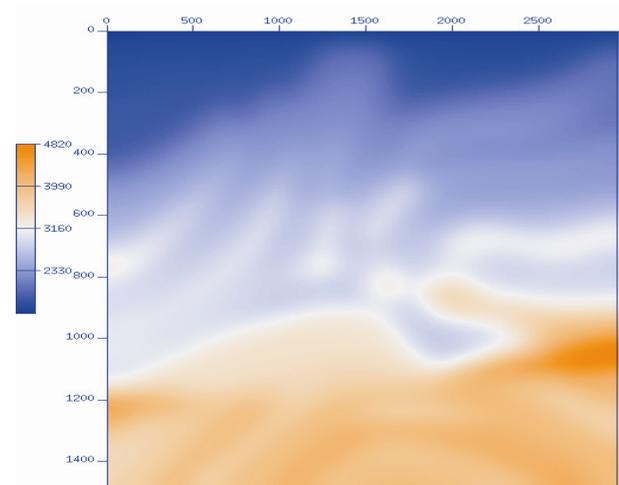


Figure 11. Initial velocity model for inversion.

$\Delta x = \Delta z = 6m$, $\Delta t = 0.6ms$, $N_t = 3501$. **Figure 11** is the initial model and **Figure 12** is the inversion result after 200 iterations. This result clearly shows that BFGS-based algorithm has obtained very good inversion accuracy.

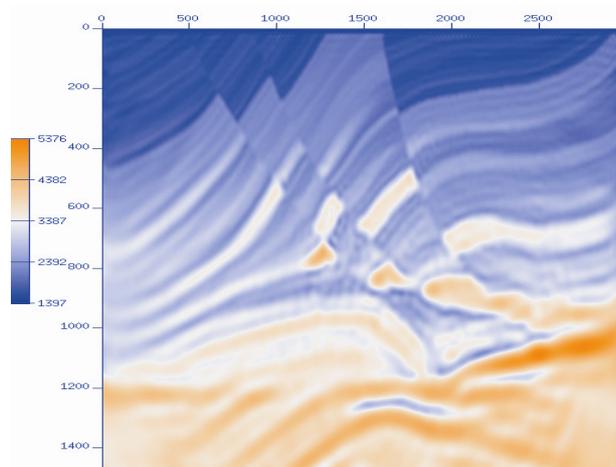


Figure 12. Inversion result after 200 iterations with the BFGS-based Algorithm 2.

4. Conclusions

We consider full-waveform inversion as an iterative minimization of the misfit between observed data and synthetic data. The algorithm is independent of the dimensionality of the problem. Two full-waveform inversion algorithms based on the CG method and BFGS method are proposed. Numerical computations for two models including complicated Marmousi model show that the BFGS-based method has higher inversion accuracy than the CG-based algorithm because the later uses the second-order derivative information in computations. The good inversion results for complicated Marmousi model show that the BFGS-based algorithm can be expected to inverse real data in the future.

5. Acknowledgements

This research is supported by the State Key project with grant number 2010CB731505 and the National Center for Mathematics and Interdisciplinary Sciences, CAS.

REFERENCES

- [1] A. Tarantola, "Inversion of Seismic Reflection Data in the Acoustic Approximation", *Geophysics*, Vol. 49, No.8, 1984, pp. 1259-1266.
- [2] C. Bunks, F. Saleck, S. Zaleski, and G. Chavent "Multiscale Seismic Waveform Inversion", *Geophysics*, Vol.60, No.5, 1995, pp.1457-1473.
- [3] R. G. Pratt, C. Shin, and G. J. Hicks, "Gauss-Newton and Full Newton Methods in Frequency-space Seismic Waveform Inversion", *Geophysical Journal International*, Vol.133, No.2, 1998, pp.341-362.
- [4] C. Burs and O. Ghattas, "Algorithmic Strategies for Full Waveform Inversion: 1D experiments", *Geophysics*, Vol.74, No.6, 2009, pp.WC37-WC46
- [5] R. G. Pratt, "Seismic Waveform Inversion in Frequency Domain Part I: Theory and Verification in a Physical Scale Model", *Geophysics*, Vol. 64, No.3, 1999, pp.888-901.
- [6] Z. M. Song, P. R. Williamson, and R. G. Pratt, "Frequency-domain Acoustic-wave Modeling and Inversion of Cross Hole data: Part 2—Inversion Method, Synthetic Experiments, and Real-data Results", *Geophysics*, Vol.60, No.3, 1995, pp.796-809.
- [7] R. G. Pratt and M. H. Worthington, "The Application of Diffraction Tomography to Cross Hole Data", *Geophysics*, Vol.53, No.10, 1988, pp.1284-1294.
- [8] C. Shin and W. Ha, "A Comparison between the Behavior of Objective Functions for Waveform Inversion in the Frequency and Laplace Domains", *Geophysics*, Vol.73, No.5, 2008, pp.VE119-VE133.
- [9] W. Hu, A. Abubakar and T. M. Habashy, "Simultaneous Multifrequency Inversion of Full-waveform Seismic Data", *Geophysics*, Vol.74, No.2, 2009, pp.R1-R14.
- [10] L. Sirgue and R. G. Pratt, "Efficient Waveform Inversion and Imaging: A Strategy for Selecting Temporal Frequency", *Geophysics*, Vol.69, No.1, 2004, pp.231-248.
- [11] R. Fletcher and C. Reeves, "Function Minimization by Conjugate Gradients", *Computer Journal*, Vol.7, No.2, 1964, pp.149-154.
- [12] C. G. Broyden, "The Convergence of a Class of Double-Rank Minimization Algorithms. 2. The New Algorithm", *J. of the Institute of Math. And its Appl.*, Vol. 6, 1970, pp.222-231.
- [13] R. Fletcher, "A New Approach to Variable Metric Algorithms", *Computer Journal*, Vol.13, No.3, 1970, pp.317-322.
- [14] D. Goldfarb, "A Family of Variable Metric Methods Derived by Variational Means", *Mathematics of Computation.*, Vol. 24, No. 109, 1970, pp. 23-26.
- [15] D. F. Shanno, "Conditioning of Quasi-Newton Methods for Function Minimization", *Math. Comput.*, Vol.24, No.111, 1970, pp.647-656 .
- [16] R. Clayton and B. Engquist, "Absorbing Boundary Conditions for Acoustic and Elastic Wave Equations", *Bulletin of the Seismological Society of America*, Vol.67, No.6, 1977, pp.1529-1540.
- [17] J. P. Berenger, "A Perfectly Matched Layer for Absorbing of Electromagnetic Waves", *J. Comput. Phys.*, Vol. 114, No.2, 1994, pp.185-200.
- [18] M. R. Hestenes and E. L. Stiefel, "Methods of Conjugate Gradients for Scaling Linear Systems", *J. Res. National Bureau Standards*, Vol. 49, No.6, 1952, pp.409-436.
- [19] E. Polak and Ribière, "Note Sur la Convergence de Directions Conjugate", *Rev. Francaise Informat Recherche Opertionelle*, 3e Année, 16, 1969, pp.35-43.
- [20] B.T. Polyak, "The Conjugate Gradient Method in Extreme Problems", *USSR Comp. Math. and Math. Phys.*, Vol.9, No.4, 1969, pp. 94-112.
- [21] L. Armijo, "Minimization of Functions Having Lipschitz Continuous First Partial Derivatives", *Pacific Journal of Mathematics*, Vol.16, No.1, 1966, pp.1-3.
- [22] P. Wolfe, "Convergence Conditions for Ascent Methods",

- SIAM Rev., Vol.11, No.2, 1969, pp.226-235.
- [23] P. Wolfe, "Convergence Conditions for Ascent Methods II: Some Corrections", SIAM Rev., Vol.13, No.2, 1971, pp.185-188.
- [24] J. Nocedal, Y. Yuan, "Analysis of a Self-scaling Quasi-Newton Method", *Math. Program*, Vol.61, No.1-3, 1993, pp.19-37.
- [25] W. Zhang, "Imaging Methods and Computations Based on the Wave Equation", Beijing, Science Press, 2009.