Exact line search method for using the L1-norm misfit function in full waveform inversion

XIAONA MA[1,2], GUANGHE LIANG[1,2], SHANHUI XU[3], ZHIYUAN LI[1,2]✉ AND HAIXIN FENG[4]

1 Key Laboratory of Mineral Resources, Institute of Geology and Geophysics, Chinese Academy of Sciences, Beijing 100029, China (lizhiyuan_lzy@163.com)

2 Innovation Academy for Earth Science, Chinese Academy of Sciences, Beijing 100029, China

3 Institute of Geophysics, China Earthquake Administration, Beijing 100081, China

4 The Third Institute of China Electronics Technology Corporation, Beijing 100015, China

Received: January 15, 2020; Revised: July 7, 2020; Accepted: September 4, 2020

ABSTRACT

Full waveform inversion (FWI) is a non-linear inverse problem that can be sensitive to noise. The tolerance of the noise-interference characteristics depends on the types of misfit functions. To date, different misfit functions, such as the least-squares norm (L2), the least-absolute-value norm (L1), and combinations of the two (e.g., the Huber and hybrid criteria), have been applied to FWI. The L2 norm is highly sensitive to non-Gaussian errors in the data and gives rise to high-amplitude artifacts in reconstructed models. For non-Gaussian noise data, the L1 norm and the Huber and hybrid criteria always reliably reconstruct models. However, the Huber and hybrid criteria require tedious error investigations to estimate their threshold criterion. Thus, the L1 norm is adopted here to improve the anti-noise ability of the FWI. The step length is closely related to the misfit function, and an optimal step-length estimation method can rapidly make the FWI algorithm reach the global minimum, with a reduced number of iterations and fewer extra forward modeling simulations during each iteration. The step length can usually be obtained using the exact or inexact line search method. Generally, the exact line search method is faster than the inexact one. Therefore, we derived an exact line search method for the L1 norm in the FWI process. Its effectiveness was tested using noise-free data from Overthrust and the SEG/EAGE salt models. The results demonstrate that this method can recover high-resolution velocity models with low computational costs. Numerical tests using the synthetic Overthrust model contaminated by strong noise were used to further validate the robustness of this exact line search method.

Keywords: full waveform inversion, L1 misfit function, exact line search method

1. INTRODUCTION

At present, geophysicists are interested in applying full waveform inversion (FWI) to describe subsurface geological structures, evaluate soil properties, and characterize energy

reservoirs. FWI makes full use of the dynamic and kinematic information in the data, so it can yield a high-resolution geometric picture of the target structures and preferable model parameters (e.g., velocity, density, and quality factor). *Lailly (1983)* and *Tarantola (1984)* solved the seismic waveform inversion using reverse time migration in the time domain. Ever since, FWI has been extensively studied using similar back-propagation techniques (*Mora, 1987*; *Sigure and Pratt, 2004*; *Virieux and Operto, 2009*; *Plessix, 2006*; *Liu and Liu, 2018*).

Despite its great potential, the FWI is a non-linear and ill-posed inverse problem, with a high computational cost (*Wang and Rao, 2006*; *Vigh and Starr, 2008*; *Virieux and Operto 2009*). Moreover, its solution is non-unique, and its misfit function is multimodal, which may fall within the local minima (*Symes, 2008*; *Guitton, 2012*). Once the low-frequency data are acquired, and one does not consider the computational limit, FWI can provide a subsurface velocity model with a higher-precision than travel-time tomography or migration-based velocity analysis (*Bednar, 1999*; *Virieux and Operto, 2009*). FWI updates the model parameters by backward propagating and iteratively minimizing the misfit function between the observed and synthetic data (*Tarantola, 1984*; *Ren et al., 2014*). However, there are many obstacles to waveform inversion, e.g., the absence of low-frequency data, two-dimensional (2D) acoustic approximations of 3D real-earth wave propagation, poor coupling between the source-receiver and the Earth, inaccurate source-wavelet estimations, and the existence of many kinds of noise.

Noise is one of the most important issues because ambient background vibrations always contaminate real seismic data (*Ha et al., 2009*). The tolerance of the noise-interference characteristic is significant in inversion processes, and it depends on the type of misfit function (*Choi and Alkhalifah, 2012*; *Ren et al., 2014*). To date, many different misfit functions have been used, such as the least-squares norm (L2), the least-absolute-values norm (L1) (*Tarantola, 1987*; *Crase et al., 1990*), the Huber criterion (*Huber, 1973*; *Guitton and Symes, 2003*), and the Hybrid criterion (*Bube and Langan, 1997*; *Claerbout and Muir, 1973*). The L2 norm is widely used in FWI, but it may not be robust in the presence of strong noise or an amplitude mismatch between the modeled synthetic data and the recorded field data. Alternative functions, such as the Huber criterion and the hybrid criterion, which combine the L2 norm (for small residuals) and L1 norm (for large residuals), are more robust to noise. However, they are sensitive to a threshold criterion that controls the transition between the criteria, and they require tedious trial-and-error deviation investigations to obtain a reliable estimation.

Theoretically, the L1 norm is more robust to noise than the L2 norm, especially when strong outliers are present (*Claerbout and Muir, 1973*; *Bube and Nemeth, 2007*), but the L1 norm has a singularity when the residual becomes zero (*Crase et al., 1990*; *Ha et al., 2009*). *Brossier et al. (2010)* claimed that they never encountered this case in all of the tests they performed. The L1 norm was introduced by *Tarantola (1987)* and *Crase et al. (1990*) for the time-domain FWI. As a measure of the data misfit, it exhibits a weak sensitivity to large measurement errors. The statistical interpretation of this insensitivity to large noise is that the robust measurements are related to the long-tailed density functions in the same way that L2 is related to the short-tailed Gaussian density function (*Tarantola, 1987*; *Guitton and Symes, 2003*). *Djikpesse and Tarantola (1999)* tried using the L1 norm to invert field data from the Gulf of Mexico area with time-domain FWI. Moreover, several studies have also discussed FWI using the L1 norm to invert field data

successfully (*Shin and Ha, 2008*; *Brossier et al., 2009a, 2010*; *Pyun et al., 2009*; *Guitton, 2012*). Therefore, we applied the L1 norm in the time-domain FWI to improve its anti-noise ability.

The step length is a key factor to successful FWI. It is necessary to choose an Optimal step length to reach a global minimum along the searching direction with a fast convergence rate, which can typically be calculated using an inexact or exact line search method (*Hu et al., 2014*; *Liu et al., 2016*). *Nocedal and Wright (2006)* proposed the inexact line search algorithm to calculate the step length appropriately using the backtracking line search method, in which an initial-guess value combined with an adaptive evaluation condition can achieve a sufficient decrease in a certain number of iterations (*Dos Santos and Pestana, 2015*), and they proposed many different methods of performing the initial estimation. *More and Thuente (1994)* set an interpolation interval for the initial step -length, used the quadratic or cubic interpolation method, and combined the sufficient decrease and the Wolfe conditions in the inexact line search algorithm to calculate the step -length, which also performed extremely well.

The exact line search method is usually performed by solving the misfit function in the current iteration. The parabolic fitting search method is performed by solving the L2 misfit function, which is a commonly used exact line search method (*Operto et al., 2007*; *Vigh et al., 2009*; *Xu and McMechan, 2014*). This method may be accurate and converge fast sometimes, but it usually requires two extra forward modeling steps in each iteration, leading to the computational costs of the forward modeling being high for the time domain. *Pica et al. (1990)* proposed an exact calculation method for obtaining the optimal step length for the L2 norm misfit function, which is directly derived and only requires extra forward modeling. The accuracy and computational efficiency of this optimal step length have been demonstrated (*Liu et al., 2017b*; *Ma et al., 2019*). *Bube and Nemeth (2007)* also used the corresponding exact step-length estimation method for the hybrid norm. We believe that an exact step-length formula for the L1 misfit function also needs to be derived. Thus, inspired by Pica's step-length estimation method for the L2 norm (*Pica et al., 1990*), we derived an exact line search method to calculate the optimal step length for the L1 norm, which can obtain a high-resolution inverted velocity with a low computational cost even if the data are contaminated by non-Gaussian noise.

The rest of this paper is organized as follows. In Section 2, the theory of FWI in the time domain is presented, and the exact line search method of the L1 norm is derived. In Section 3, the synthetic data of the Overthrust model and the SEG/EAGE salt dome AA′ line are used to demonstrate the effectiveness of our method. Then, we test this exact line search method on synthetic seismic data with three different types of noise: spiky-type noise, strong ground-motion noise, and strong ground-motion noise combined with spiky-type noise. In addition, we compare the results obtained using the L1 norm with those obtained using the L2 norm. Finally, the conclusions are presented in Section 4.

2. THEORY

Gradient methods, such as the steepest-descent method and the conjugate gradient method; and quasi-Newton methods, such as the limited memory Broyden–Fletcher–Goldfarb–Shanno (L-BFGS) method; and the Newton method are usually adopted to solve

the FWI problem (*Liu and Nocedal, 1989*; *Pratt et al., 1998*; *Brossier et al.,* 2009b; *Warner et al., 2013*; *Hager and Zhang, 2006*). The steepest-descent method does not usually guarantee convergence to the global minimum. The Newton and L-BFGS methods generally possess better convergence properties. However, the computation of the Newton method is always expensive. To balance the computation costs and the inversion accuracy, we use the L-BFGS inversion algorithm for all of the numerical examples in this paper (*Nocedal and Wright, 1999*). So the velocity model can be updated by

$$\boldsymbol{m}_{k+1} = \boldsymbol{m}_k + \alpha_k \boldsymbol{d}_k , \tag{1}$$

where $\boldsymbol{m}_{k+1}$ and $\boldsymbol{m}_k$ are the model parameters (e.g., velocity, density, and quality factor) at the (k+1)-th and k-th iteration, respectively; $\boldsymbol{d}_k = -\boldsymbol{H}_k \nabla f_k$ is the search direction, $\boldsymbol{H}_k$ is the inverse Hessian approximation matrix at the k*th* iteration, ∇f_k is the gradient at the k-th iteration, and α_k is the step length at the k-th iteration, which is a positive scalar. Next, we introduce our exact line search method based on the L1 norm.

2.1. Time-domain inversion based on the L1 norm

FWI is defined as an optimization problem that seeks to minimize the misfit function f, which measures the difference between the observed data $p^{obs}(x,z,t)$ and the calculated data $p^{cal}(x,z,t)$, where $t \in [0,T]$ represents the running time, T is the total recording time, x and z are the coordinates of the position, $(x,z) \in R$, R being the spatial domain, $p(x,z,t)$ is the pressure field. *Tarantola (1987)* and *Crase et al. (1990)* proposed the L1 norm for the misfit function

$$f = \sum_s \sum_g \left| p^{cal}(x,z,t) - p^{obs}(x,z,t) \right| , \tag{2}$$

where s represents the source (shot) and g represents the receiver. Here the calculated data are generated by a 2D constant-density acoustic-wave equation

$$\left(\frac{\partial^2 p(x,z,t)}{\partial x^2} + \frac{\partial^2 p(x,z,t)}{\partial z^2} + s(x,z,t) \right) = \frac{1}{c^2} \frac{\partial^2 p(x,z,t)}{\partial t^2} , \tag{3}$$

where c is the acoustic velocity, which is one of the model parameters; and $s(x,z,t)$ is a source term.

Implementation of FWI requires the calculation of the gradient of a misfit function of the observed data $p^{obs}(x,z,t)$ and the calculated data $p^{cal}(x,z,t)$ with respect to the velocity c. There are two main steps to calculate the gradient of the L1 norm misfit function: forward propagation and backward propagation. The forward-propagation source wavefields are estimated using Eq. (3) with the initial conditions (t = 0)

$$p(x, z, 0) = 0 , \quad \frac{\partial p(x, z, 0)}{\partial t} = 0 . \tag{4}$$

The error between the calculated and observed data in the L1 norm sense is

$$f = \iint\limits_{(x,z)\in R} dx\, dz \int\limits_0^T func\left(p^{cal}(x, z, t) - p^{obs}(x, z, t) \right) \mathrm{d}t , \tag{5}$$

where

$$func(r) = |r| = \sum_s \sum_g \left| p^{cal}(x, z, t) - p^{obs}(x, z, t) \right| . \tag{6}$$

Then, the Lagrange multiplier method is used:

$$J = f + \iint\limits_{(x,z)\in R} \int\limits_0^T \lambda \left[\left(\frac{\partial^2 p}{\partial x^2} + \frac{\partial^2 p}{\partial z^2} + s \right) - \frac{1}{c^2} \frac{\partial^2 p}{\partial t^2} \right] \mathrm{d}t\, dx\, dz , \tag{7}$$

where $\lambda(x, z, t)$ is the Lagrange multiplier function. By integrating Eq. (7) in parts, we obtain

$$J = f + \iint\limits_{(x,z)\in R} \int\limits_0^T \left\{ p \left[\left(\frac{\partial^2 \lambda}{\partial x^2} + \frac{\partial^2 \lambda}{\partial z^2} \right) - \frac{1}{c^2} \frac{\partial^2 \lambda}{\partial t^2} \right] + \lambda s \right\} \mathrm{d}t\, dx\, dz . \tag{8}$$

By using $\partial J / \partial p = 0$, we obtain the adjoint wave equation with respect to λ:

$$\frac{1}{c^2} \frac{\partial^2 \lambda}{\partial t^2} = \left(\frac{\partial^2 \lambda}{\partial x^2} + \frac{\partial^2 \lambda}{\partial z^2} \right) + backc\left(p^{cal}(x, z, t) - p^{obs}(x, z, t) \right) . \tag{9}$$

The back-propagation residual of the L1 norm is expressed as a sign function:

$$backc(r) = \mathrm{sgn}(r) = \begin{cases} +1 , & r > 0 \\ 0 , & r = 0 , \\ -1 & r < 0 . \end{cases} \tag{10}$$

For the final condition ($t = T$) we set

$$\lambda(x, z, T) = 0 , \quad \frac{\partial \lambda(x, z, T)}{\partial t} = 0 . \tag{11}$$

Finally, we obtain the gradient of the misfit function:

$$\nabla f = \sum_s \sum_g \frac{2}{c^3} \int\limits_0^T \lambda \frac{\partial^2 p}{\partial t^2} \mathrm{d}t . \tag{12}$$

The change in misfit function only changes the adjoint source, not the entire gradient expression. We can also obtain a similar gradient expression for the other norms, such as the Huber criterion.

2.2. Exact line search method

For the local optimization method, we always need a step length to scale the gradient of the misfit function, which ensures that the misfit function value will be reduced iteratively. The step-length estimation is closely related to the misfit function. As we discussed in the previous section, we derived an exact line search method to calculate the step-length formula for the L1 norm in the FWI. The L1 norm is a least-absolute-values norm, so first, we must express Eq. (6) as

$$f = \sum_s \sum_g \sqrt{\left(p^{cal}(x,z,t) - p^{obs}(x,z,t) \right)^2} \,. \tag{13}$$

In the current state, by applying the Taylor-series expansion of the step length α to the first order, the calculated data can be approximated as follows:

$$p^{cal}(\boldsymbol{c}_k + \alpha_k \boldsymbol{d}_k) \approx p^{cal}(\boldsymbol{c}_k) + \alpha_k \left[\nabla p^{cal}(\boldsymbol{c}_k) \right]^{\mathrm{T}} \boldsymbol{d}_k \,, \tag{14}$$

where $\boldsymbol{c}_k$ is the velocity at the k-th iteration; $\boldsymbol{d}_k$ is the search direction at the k-th iteration; and α_k is the step length at the k-th iteration. Therefore, we can obtain the misfit function as a new formula:

$$f(\boldsymbol{c}_k + \alpha_k \boldsymbol{d}_k) \approx \sum_s \sum_g \sqrt{\left[p^{cal}(\boldsymbol{c}_k) + \alpha \left(\nabla p^{cal}(\boldsymbol{c}_k) \right)^{\mathrm{T}} \boldsymbol{d}_k - p^{obs}(\boldsymbol{c}_k) \right]^2} \,. \tag{15}$$

If the value of the misfit function approaches a local or global minimum, Eq. (15) satisfies the following condition:

$$\frac{\partial f(\boldsymbol{c}_k + \alpha_k \boldsymbol{d}_k)}{\partial \alpha} = 0 \,. \tag{16}$$

The optimal step-length α satisfies Eq. (16), which is the partial derivative of misfit function with respect to the step-length. Equation (16) contains the step-length information we can derive from Eq. (15). So we expand and arrange Eq. (16) and obtain the general form of the optimal step-length as follows:

$$\alpha \approx -\alpha_t \frac{\sum_s \sum_g \dfrac{\tilde{\delta} p \, \tilde{\delta} p_t}{\left| \tilde{\delta} p + \mu \tilde{\delta} p_t \right|}}{\sum_s \sum_g \dfrac{\tilde{\delta} p_t \, \tilde{\delta} p_t}{\left| \tilde{\delta} p + \mu \tilde{\delta} p_t \right|}} \,, \tag{17}$$

where α_t is the test step length and $\mu = [0, 1]$ is a proportion coefficient,

$$\tilde{\delta} p = p^{cal}(x, z, t) - p^{obs}(x, z, t), \tag{18}$$

$$\tilde{\delta} p_t = p^{cal}(\boldsymbol{c}_k + \alpha_t \boldsymbol{d}_k) - p^{cal}(\boldsymbol{c}_k). \tag{19}$$

The step-length calculation expression in Eq. (17) mainly consists of four parts: α_t, $\tilde{\delta} p$, $\tilde{\delta} p_t$ and $\left|\tilde{\delta} p + c_1 \tilde{\delta} p_t\right|$. In this exact line search scheme, the presence of test step length α_t indicates that extra forward modeling is required.

3. NUMERICAL EXAMPLES

In Subsection 3.1, we demonstrate the effectiveness of the exact line search method in the L1 norm for the FWI. The Overthrust (*Ravaut et al., 2004*) and Sigabee2A velocity models (*Aminzadeh et al., 1994*) are used to validate this method. In the second subsection, to further investigate the robustness of this method, we add strong noise to the original synthetic data of the Overthrust model, such as spiky-type noise. Under these more realistic conditions, we also compare the efficiency of the L1 and L2 norms.

In all of the numerical experiments, we use the finite-difference method to achieve the forward source propagation and back propagation of the residuals, and the perfectly-matched-layer (PML) conditions are used at the boundaries (*Berenger, 1994*). We also adopt the multi-scale method in the FWI scheme in the time domain, which allows us to implement the method successfully (*Bunks et al., 1995*; *Boonyasiriwat et al., 2009*). The Wiener filter function is used to filter the data and to obtain the datasets for the different frequency bands for the multi-scale inversion (*Liu et al., 2017b*). In addition, in all of the numerical examples, we adopt the following stopping conditions in the inversion process. First, the iteration number reaches the maximum value of 60 at each stage. Second, the misfit is within an acceptable value. Third, the relative rate of change of the two successive function misfit values is less than a small constant (*Ma et al., 2019*).

3.1. Efficiency test

The Overthrust (*Ravaut et al., 2004*) and Sigabee2A velocity models (*Aminzadeh et al., 1994*) are used to validate our exact line search method in the L1 norm for the FWI in this subsection. We also compare our exact step-length estimation method with the parabolic search method and backtracking line search method in the L1 norm. Although parabolic search method is an exact line search method that is derived from the L2 norm, it has potential for other misfit functions in the FWI process (*Liu et al., 2017a*). So we try the parabolic search method for L1 norm in this case, and compare it with our method. In order to distinguish it from the parabolic search method, our method is temporarily called the *direct line search method* in the following figures. We also compare our method with the backtracking line search method, which uses a sufficient decrease condition and an initial step length to calculate the inexact step length (*Nocedal and Wright, 2006*). Back-tracking line search method is an inexact line search method, which is applicable for both

L1 and L2 norm misfit functions. All of the parameter and inversion sets are the same, and the only difference is the line search method.

3.1.1. Overthrust model

Here, we use the overthrust velocity model to test the efficiency of the exact line search method in the L1 norm. Figure 1a shows the overthrust velocity model, and Fig. 1b shows the smooth velocity model, which is used as the initial model for the inversion. The velocity model is resized to 601×150 grid samples in the horizontal and vertical directions, respectively, with a grid spacing of 12 m. To simulate the synthetic data of the overthrust model, the acquisition geometry consists of 60 shots separated by an interval of 120 m, and the source signal is a Ricker wavelet with a peak frequency of 18 Hz. In this test, the time sampling interval is 1 ms, and the total recording time is 3 s. The receiver spacing is 12 m, and 601 receivers are evenly distributed along the surface.

Figure 2 depicts the inverted results obtained using FWI with the different line search methods. By comparing Fig. 2a with Fig. 1a, it can be seen that the inversion result in Fig. 2a is very close to the true overthrust model (Fig. 1a). This exact line search method can recover nearly all of the structures in the shallow and deep layers, and its resolution is comparable with those of the other line search methods. To provide more details, we show the depth profiles of velocities at the horizontal positions of 1.8, 4.8, and 6.0 km in Fig. 3. The inverted velocity profiles for our method are in slightly better agreement with the real velocities, especially in the sections indicated by arrows. Figure 4a shows the normalized root-mean-square (*rms*) error of the iterations. Our method has a higher convergence rate, and its *rms* is lower than that of the other methods (Fig. 4a). We analyzed the computational efficiency from the point of the modeling simulation numbers (*Ma et al., 2019*; *Anagaw and Sacchi, 2014*). The number of iteration in our method is 102, the number of extra calculations of the corresponding misfit function needed for the step-length computation is also 102, and the total computational cost of the modeling is the lowest (Fig. 4b).

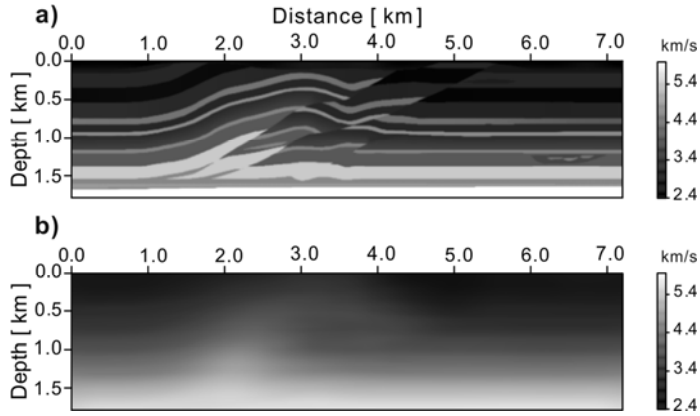

Fig. 1. **a)** The true Overthrust velocity model; **b)** the initial smooth velocity model.

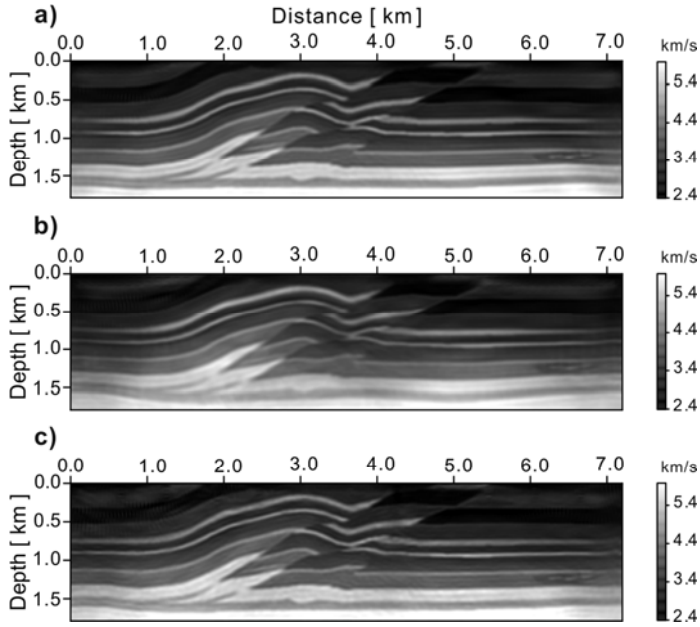

Fig. 2. Full waveform inversion results of the Overthrust velocity model obtained by: **a)** proposed direct line search method; **b)** back-tracking line search method; **c)** parabolic search method.

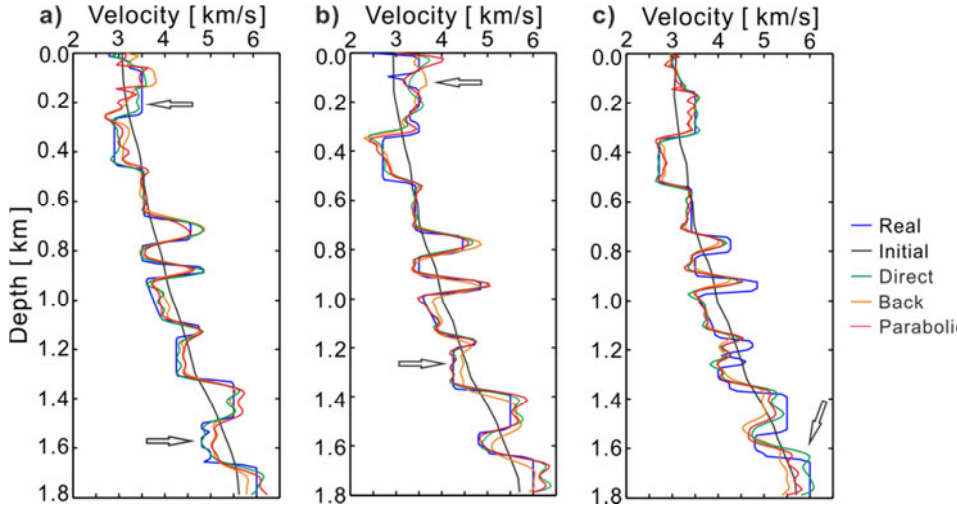

Fig. 3. Depth profiles of the real Overthrust velocity model, initial smooth velocity model, and inverted Overthrust velocity models obtained using the direct line, back-tracking line, and parabolic search methods at horizontal positions of **a)** 1.8, **b)** 4.8, and **c)** 6.0 km. Arrows point to sections with good agreement of the proposed direct line method with the real velocity.

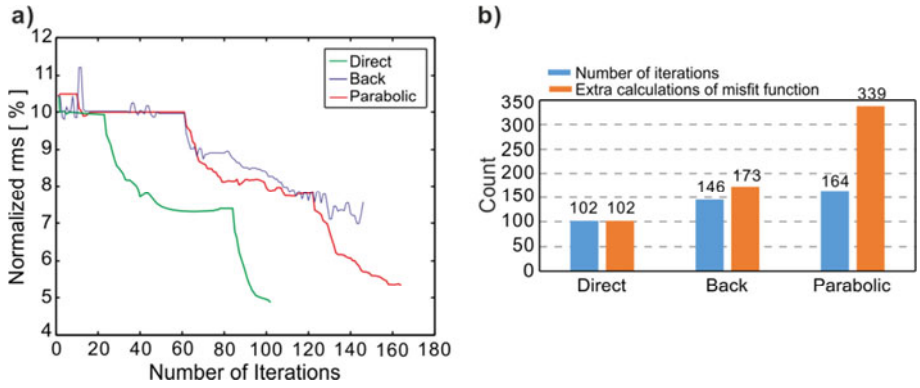

Fig. 4. **a)** Normalized root mean square (*rms*) errors of the three line search methods used with the Overthrust velocity model; **b)** total number of iterations and number of extra calculations of the misfit function for the three line search methods used with the Overthrust velocity model.

3.1.2. SEG/EAGE salt dome AA′ line

To further validate our method, we performed the inversion for the SEG/EAGE salt-dome AA′ line, which was extracted from the original SEG/EAGE salt dome (*Alkhalifah, 1998*) (Fig. 5a). The size of the modified salt-dome model is 7.8 × 2.29 km with a grid spacing of 10 m. The acquisition system includes 78 sources, evenly spaced at 100 m on the surface of the model. There are 780 receivers on the surface of the model, with an interval of 10 m. The source function is an 18-Hz Ricker wavelet (*Ma et al., 2019*). Figure 5b shows the initial model for the inversion.

Figure 6 shows the inversion results, from which we can see that the model inverted using our method (Fig. 6a) is very close to the true one (Fig. 5a). The layered structures

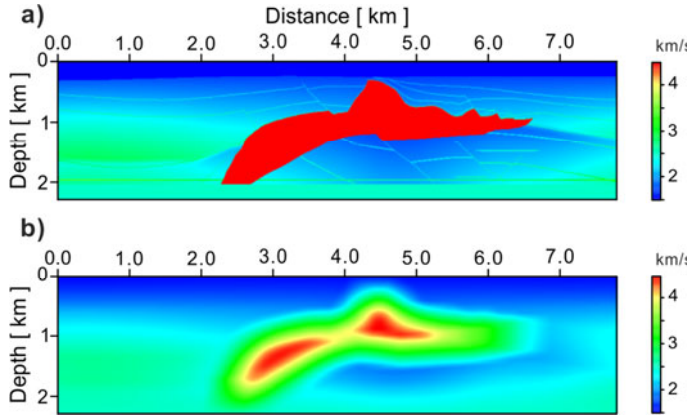

Fig. 5. **a)** The true velocity model of the SEG/EAGE salt dome (*Alkhalifah, 1998*) AA′ line; and **b)** the initial smooth velocity model.

and the top of the recovered salt dome are well reconstructed (as indicated by the ellipse and the arrows), and they are better than those obtained using the other methods. However, the parabolic search method can recover the deep parts better than the other methods, such as the deep parts at the left-hand side. To give more details, Fig. 7 shows the depth-velocity profiles at horizontal positions of 3.0 and 6. 0 km. We find that our method matches the true velocities well. The amplitude of the parabolic search methodis too small at 6.0 km (indicated by arrows). Our method provided the best inversion results with a satisfactory convergence rate for the velocity model (Fig. 8a). Moreover, the total forward modeling costs of our method are the lowest, and the efficiency of the parabolic search method is low because it entails more extra computational costs for the forward simulations (Fig. 8b).

The above two examples both demonstrate that our method is efficient and exhibits an outstanding performance for the L1 norm in terms of the computational cost and convergence rate. The parabolic search method may not be suitable for the L1 norm in the FWI because it consumes a significant amount of computational resources and assumes that the misfit function is a quadratic function.

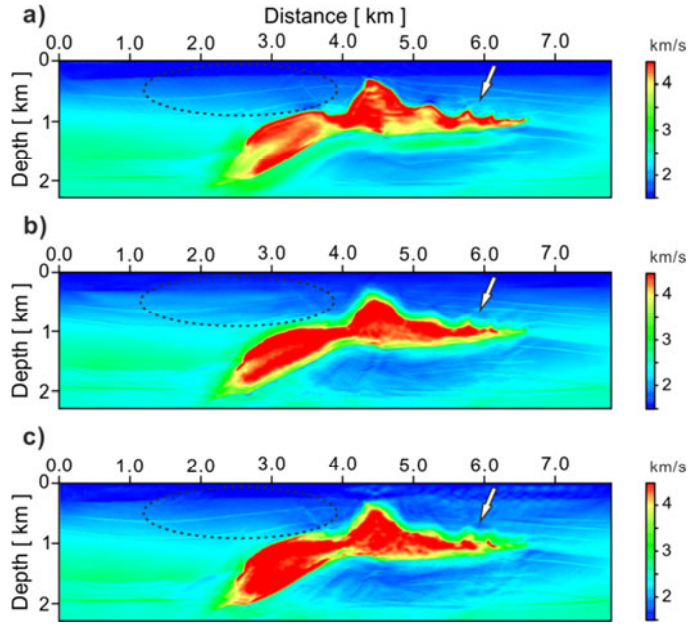

Fig. 6. Full waveform inversion of the salt velocity model obtained using **a)** direct line search method, **b)** back-tracking line search method, and **c)** parabolic search method. The ellipses and arrows indicate where the layered structures and the top of the salt dome are well reconstructed by the proposed direct line method and are better than the results of the other methods.

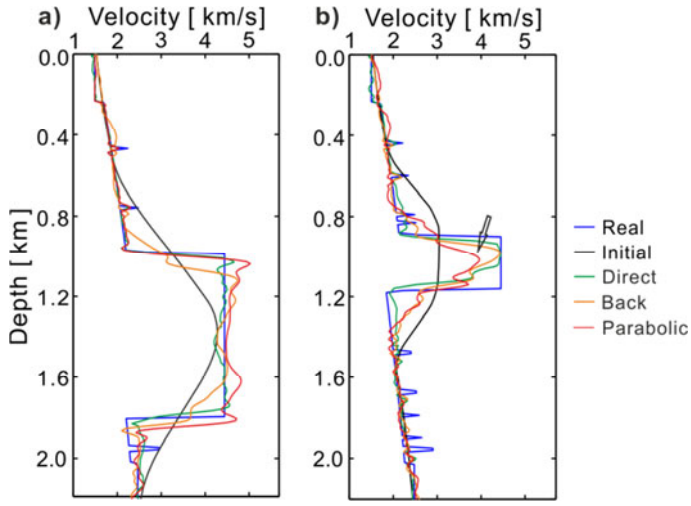

Fig. 7. Depth-velocity profiles of the real salt velocity model, initial smooth velocity model, and inverted salt velocity models obtained using the direct line, the back-tracking line, and the parabolic search methods at horizontal positions of **a)** 3.0 and **b)** 6.0 km. The arrow indicates where the amplitude of the results obtained using the parabolic search method is small.

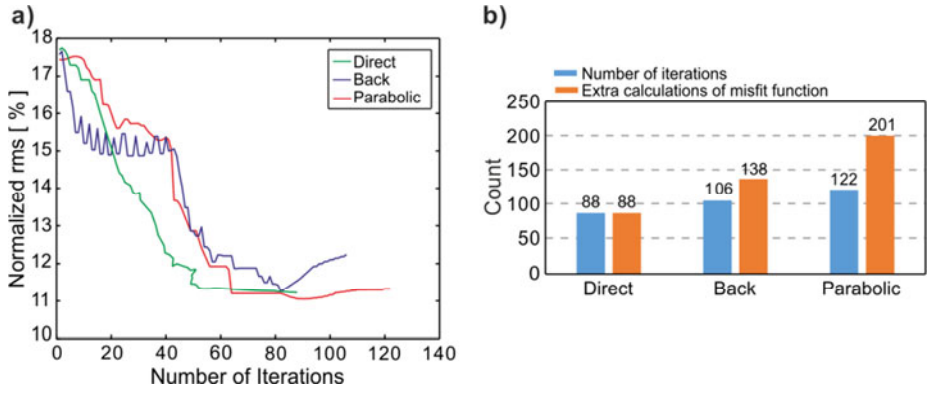

Fig. 8. **a)** Normalized root-mean-square (*rms*) errors of the three line search methods for the salt velocity model. **b)** Total number of iterations and number of extra calculations for the misfit function of the three line search methods for the salt velocity model.

3.2. The impactor of noise

To test the robustness of our proposed method under more realistic conditions, we added strong noise to the original synthetic data of the Overthrust velocity model. various types of noise were added, including spiky-type noise, strong ground-motion noise, and strong ground-motion noise combined with spiky-type noise, in order to simulate a poorly

pre-processed dataset (Fig. 9). We also compare the robustness of the L1 norm with that of the L2 norm for noisy datasets.

3.2.1. Synthetic data with spiky-type noise

We added six spikes to simulate spiky-type noise, in which the synthetic data are first replaced by their average value within the time window, and then, they are rescaled by a factor of 15. The sizes and positions of the six rectangles are fixed for all 60 shots, as shown in Fig. 9b. The inversion results are shown in Fig. 10. Compared with the true Overthrust velocity model shown in Fig. 1a, the L1 norm combined with our method performs better in terms of resolution and quality. Nearly all of the structures and features of the Overthrust velocity model are properly reconstructed, and the thrust fault and thin beds at the two sides are recovered satisfactorily (Fig. 10a). In contrast, the L2 norm produces a poorer reconstructed velocity model. A dislocation phenomenon occurs in the model that is polluted by strong artifacts, particularly in the thrust area, and in the deep

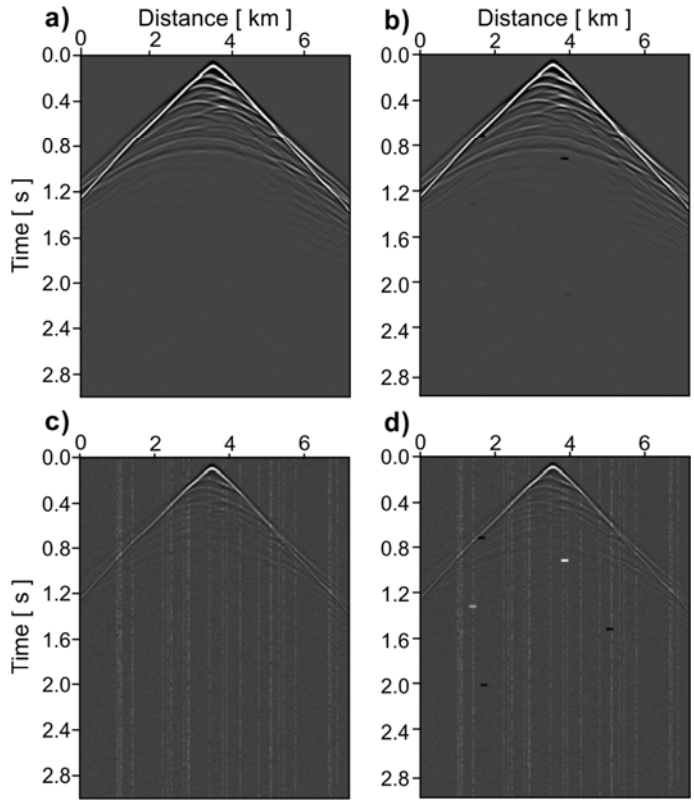

Fig. 9. **a)** Seismograms obtained from the true Overthrust model of the 30-th shot; **b)** data contaminated with spiky-type noise; **c)** data contaminated with strong ground-motion noise; **d)** data contaminated with both strong ground-motion and spiky-type noise. Small dashes indicate the strong ground-motion noise.

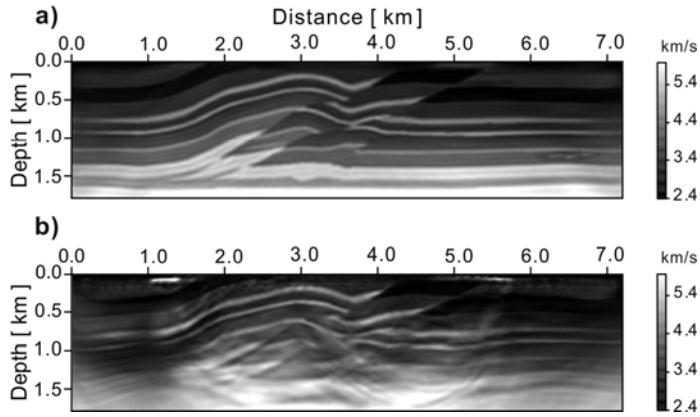

Fig. 10. Full waveform inversion of the Overthrust velocity model with the noisy data containing spiky-type noise obtained using **a)** L1, and **b)** L2 norm.

structure (Fig. 10b). Through the depth-velocity profile analysis shown in Fig. 11, we found that our L1 method can still recover the structures and features of the velocity model, and the results are as acceptable as the results obtained from the noise-free data. However, the velocity profiles of the L2 norm show some destabilization, even opposite the spatial position (indicated by arrows). In terms of the convergence rates, the L1 norm obviously converges more steadily and faster than the L2 norm in this case (Fig. 12).

3.2.2. Strong ground motion

Here, we randomly rescale some of the seismic traces contaminated by Gaussian white noise with signal-to-noise ratio (S/N) of 20 dB to simulate a poorly pre-processed strong ground-motion dataset (Fig. 9c). The inversion results of the two norms are shown in Fig. 13. As it can be seen from Fig. 13a, the Overthrust velocity model is recovered using our method. The resolution is lower than the result in Fig. 2a, because strong ground-motion noise was added to the data. The L2 norm failed to recover the true velocity model (Fig. 13b). The Depth-velocity profiles shown in Fig. 14 demonstrate that the velocity profiles of our method match the true profile better. Fig. 15 shows the normalized *rms* errors of the inversion, with our method converging after 86 iterations.

3.2.3. Ground-motion plus spiky-type noise

To further test the robustness of our method under more complex non-Gaussian noise conditions, based on the strong ground-motion dataset, we added spiky-type noise (Fig. 9d, *Liu et al., 2017a*). For this strong noise, we just show the inversion results of the L1 norm in Fig. 16. Compared with the true Overthrust velocity model shown in Fig. 1a, our method is able to recover the true velocity model. The structures and features of the model are reconstructed properly, although the resolution of the inverted image is low compared to that in Fig. 2a. Detailed analysis of the depth-velocity profiles revealed that the velocity profiles of our method match the true profile well (Fig. 17). The normalized

rms errors of the velocity model obtained from the inversion are shown in Fig. 18, from which it can be seen that the L1 norm converges after 66 iterations.

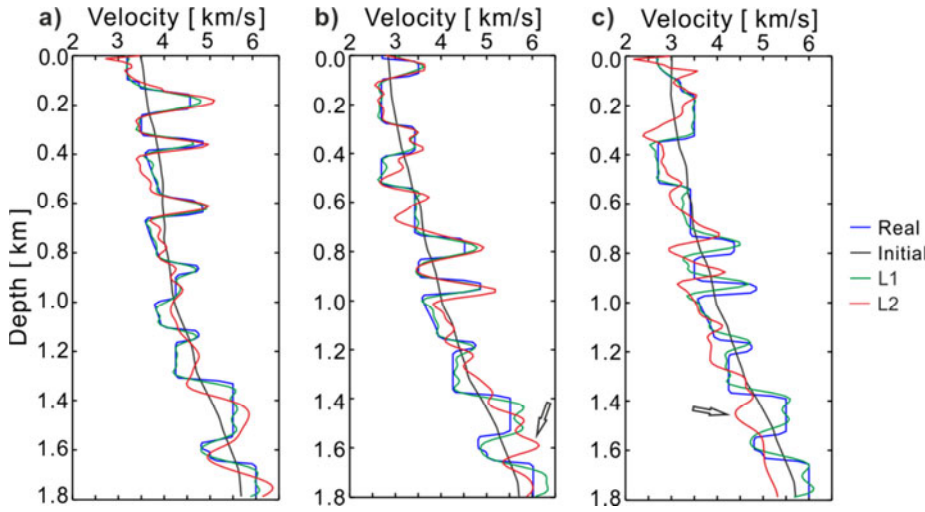

Fig. 11. Depth-velocity profiles of the Overthrust velocity model (contaminated with spiky-type noise) obtained using the L1 and L2 norms at horizontal positions of **a)** 3.0, **b)** 4.8, and **c)** 6.0 km. The arrows indicate where the inversion results of the L2 norm show some destabilization, even opposite to the spatial position.

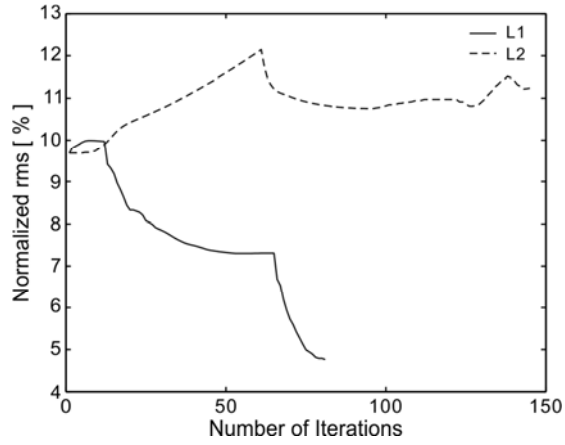

Fig. 12. Normalized root-mean-square (*rms*) errors of the L1 and L2 norms for the Overthrust velocity model (contaminated with spiky-type noise).

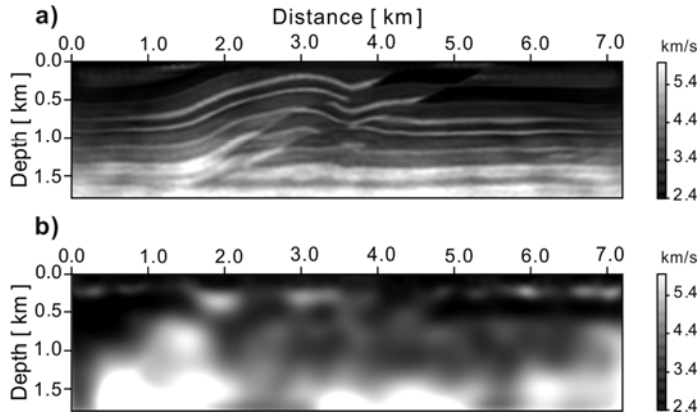

Fig. 13. Full waveform inversion results of the Overthrust velocity model with the noisy data containing ground-motion noise obtained using **a)** L1, and **b)** L2 norm.

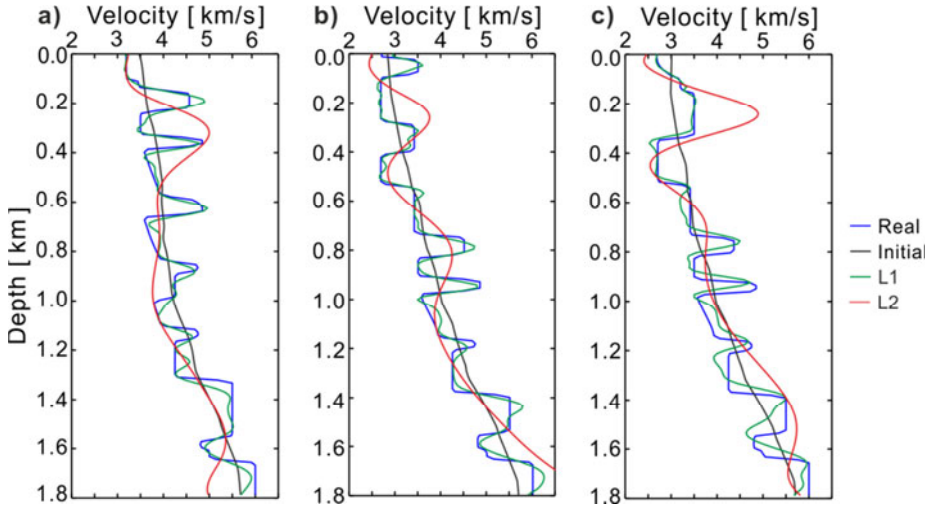

Fig. 14. Depth-velocity profiles of the Overthrust velocity model (contaminated with ground-motion noise) obtained using L1 and L2 norm at horizontal positions of **a)** 3.0, **b)** 4.8, and **c)** 6.0 km.

The above three examples demonstrate the robustness of the proposed exact line search method when using data contaminated by spiky-type and strong ground-motion noise. As it was expected, because of the non-Gaussian noise added, the inversion results for the L2 norm were not good, and the method could not handle the large sparse errors in the noisy data. The L1 norm effectively suppressed the negative impact of the strong noise during the inversion process, and thus, it provided more reliable results.

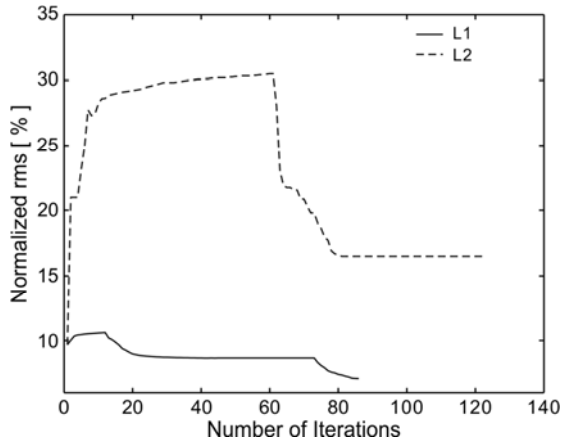

Fig. 15. Normalized root-mean-square (*rms*) error of the L1 and L2 norms for the Overthrust velocity model (contaminated with ground-motion noise).

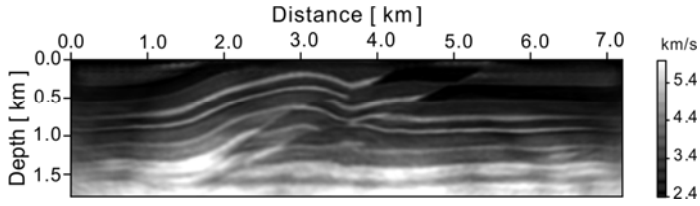

Fig. 16. Full waveform inversion of the Overthrust velocity model with the noisy data containing ground-motion plus spiky-type noise obtained using L1 norm.

4. CONCLUSIONS

We derived an exact line search method for the L1 norm and compared it with other line search methods. The noise-free data used in the two numerical tests, i.e., the Overthrust model and the SEG/EAGE salt dome AA′ line, verify the effectiveness of our method, which performs extremely well for the L1 norm in terms of the computational costs. We also validated the robustness of our method under more realistic conditions. The Overthrust velocity model with noisy data containing strong outliers demonstrates that our method can obtain high resolution results with a fast convergence rate. In addition, we compared the abilities of the L1 and L2 norms to deal with the noisy datasets and found that the robust L1 norm can handle the strong noise very well, whereas the L2 norm failed to successfully invert the velocity models.

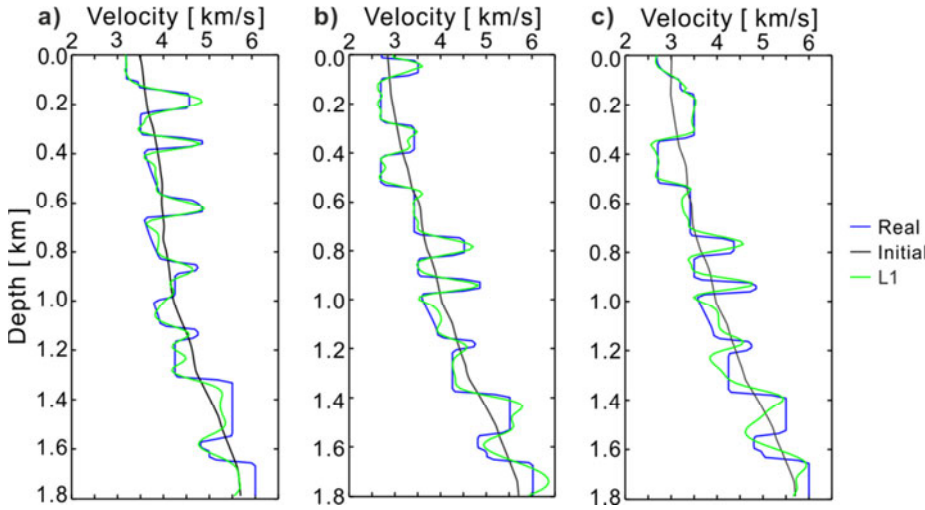

Fig. 17. Depth-velocity profiles of the Overthrust velocity model (contaminated with ground-motion plus spiky-type noise) obtained using the L1 norm at horizontal positions of **a)** 3.0, **b)** 4.8, and **c)** 6.0 km.

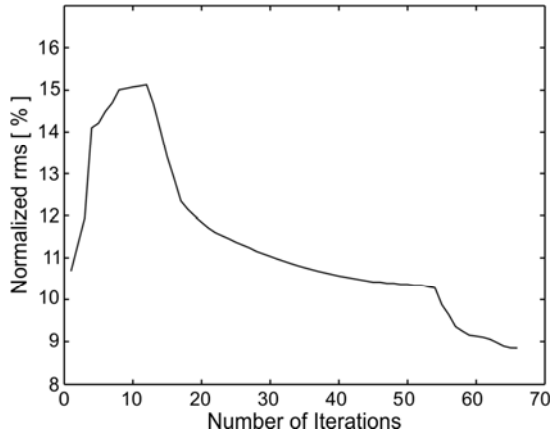

Fig. 18. Normalized root-mean-square (*rms*) error of the L1 norm for the Overthrust velocity model (contaminated with ground-motion plus spiky-type noise).

Acknowledgements: We thank Liu Youshan for his insightful discussions. This research was financially supported by the Key Research Program of the Institute of Geology and Geophysics, Chinese Academy of Sciences (Grant No. IGGCAS-201902), Open Issue of Key Laboratory of Mineral Resources, Chinese Academy of Sciences (Grant No. KLMR2017-17), and a second class general financial grant from the China Postdoctoral Science Foundation (Grant No. 2018M641470).

References

Alkhalifah T., 1998. The fast marching method in spherical coordinates: SEG/EAGE salt dome model. *Stanford Exploration Project*, Report 97, 251–264.

Aminzadeh F., Burkhard N., Nicoletis L., Rocca F. and Wyatt K., 1994. SEG/EAGE 3-D modeling project: 2nd update. *The Leading Edge*, **13**, 949–952, DOI: 10.1190/1.1437054.

Anagaw A.Y. and Sacchi M.D., 2014. Comparison of multifrequency selection strategies for simultaneous-source full-waveform inversion. *Geophysics*, **79**, R165–R181, DOI: 10.1190/GEO2013-0263.1.

Berenger J., 1994. A perfectly matched layer for the absorption of electromagnetic waves. *J. Comput. Phys.*, **114**, 185–200, DOI: 10.1006/jcph.1994.1159.

Bednar J., 1999. Theoretical comparison of equivalent-offset migration and dip moveout pre-stack imaging. *Geophysics*, **64**, 191–196, DOI: 10.1190/1.1444515.

Boonyasiriwat C., Valasek P., Routh P., Cao W., Schuster G.T. and Macy B., 2009. An efficient multiscale method for time-domain waveform tomography. *Geophysics*, **74**, WCC59–WCC68, DOI: 10.1190/1.3151869.

Brossier R., Operto S. and Virieux J., 2009a. Robust elastic frequency domain full waveform inversion using the L1 norm. *Geophys. Res. Lett.*, **36**, L20310, DOI: 10.1029/2009GL039458.

Brossier R., Operto S. and Virieux J., 2009b. Seismic imaging of complex onshore structures by 2D elastic frequency-domain full waveform inversion. *Geophysics*, **74**, WCC105–WCC118, DOI: 10.1190/1.3215771.

Brossier R., Operto S. and Virieux J., 2010.Which data residual norm for robust elastic frequency-domain full waveform inversion? *Geophysics*, **75**, R37–R46, DOI: 10.1190/1.3379323.

Bube K.P. and Langan R.T., 1997. Hybrid minimization with applications to tomography. *Geophysics*, **62**, 1183–1195, DOI: 10.1190/1.1444219.

Bube K.P. and Nemeth T., 2007. Fast line searches for the robust solution of linear systems in the hybrid and Huber norms. *Geophysics*, **72**, A13–A17, DOI: 10.1190/1.243163.

Bunks C., Saleck F.M., Zaleski S. and Chavent G., 1995. Multiscale seismic waveform inversion. *Geophysics*, **60**, 1457–1473, DOI: 10.1190/1.1443880.

Choi Y. and Alkhalifah T., 2012. Application of multi-source waveform inversion to marine streamer data using the global correlation norm. *Geophys. Prospect.*, **60**, 748–758, DOI: 10.1111/j.1365-2478.2012.01079.x.

Claerbout J. and Muir F., 1973. Robust modeling with erratic data. *Geophysics*, **38**, 826–844, DOI: 10.1190/1.1440378.

Crase E., Pica A., Noble M., McDonald J. and Tarantola A., 1990. Robust elastic nonlinear waveform inversion: Application to real data. *Geophysics*, **55**, 527–538, DOI: 10.1190/1.1442864.

Djikpesse H.A. and Tarantola A., 1999. Multiparameter L1 norm waveform fitting: Interpretation of Gulf of Mexico reflection seismograms. *Geophysics*, **64**, 1023–1035, DOI: 10.1190/1.1444611.

Dos Santos A.W.G. and Pestana R.C., 2015. Time-domain multiscale full-waveform inversion using the rapid expansion method and efficient step-length estimation. *Geophysics*, **80**, R203–R216, DOI: 10.1190/GEO2014-0338.1.

Guitton A., 2012. Blocky regularization schemes for full-waveform inversion. *Geophys. Prospect.*, **60**, 870–884, DOI: 10.1111/j.1365-2478.2012.01025.x.

Guitton A. and Symes W.W., 2003. Robust inversion of seismic data using the Huber norm. *Geophysics*, **68**, 1310–1319, DOI: 10.1190/1.1598124.

Ha T., Chung W. and Shin C., 2009. Waveform inversion using a backpropagation algorithm and a Huber function norm. *Geophysics*, **74**, R15–R24, DOI: 10.1190/1.3112572.

Hager W.W. and Zhang H.C., 2006. A survey of nonlinear conjugate gradient methods. *Pac. J. Optim.*, **2**, 35–58.

Hu G.H., Wang L.X. and Fang W.B., 2014. Full waveform inversion method and application. China Petroleum Industry Press, Beijing, China (in Chinese).

Huber P.J., 1973. Robust regression: asymptotics, conjectures, and Monte Carlo. *Ann. Stat.*, **1**, 799–821, DOI: 10.1214/aos/1176342503.

Lailly P., 1983. The seismic inverse problems as a sequence of before stack migration. In: Bednar J.B. and Weglein A. (Eds), *Conference on Inverse Scattering - Theory and Application*. SIAM, Philadelphia, PA, 206–220.

Liu D. and Nocedal J., 1989. On the limited memory BFGS method for large scale optimization. *Math. Program.*, **45**, 503–528, DOI: 10.1007/BF01589116.

Liu X.J. and Liu Y.K., 2018. Plane-wave domain least-squares reverse time migration with free-surface multiples. *Geophysics*, **83**,477–487.

Liu Y.S., Teng J.W., Xu T., Bai Z.M., Lan H.Q. and Badal J., 2016. An efficient step-length formula for correlative least squares reverse time migration. *Geophysics*, **81**, S221–S228, DOI: 10.1190/GEO2015-0529.1.

Liu Y.S., Teng J.W., Xu T., Wang Y.H., Liu Q.Y. and Badal J., 2017a. Robust time-domain full waveform inversion with normalized zero-lag cross-correlation objective function. *Geophys. J. Int.*, **209**, 106–122, DOI: 10.1093/gji/ggw485.

Liu Y.S., Teng J.W., Xu T., Badal J., Liu Q.Y. and Zhou B., 2017b. Effects of conjugate gradient methods and step-length formulas on the multiscale full waveform inversion in time domain: Numerical experiments. *Pure Appl. Geophys.*, **174**, 1983–2206.

Ma X.N., Li Z.Y., Ke P., Xu S.H., Liang G.H. and Wu X.Q., 2019. Research of step-length estimation methods for full waveform inversion in time domain. *Explor. Geophys.*, **50**, 583–599, DOI: 10.1080/08123985.2019.1641266.

More J.J. and Thuente D.J., 1994. Line search algorithms with guaranteed sufficient decrease. *ACM Trans. Math. Softw.*, **20**, 286–307.

Nocedal J. and Wright S.J., 2006. *Numerical Optimization*. Springer Nature, Cham, Switzerland.

Pica A., Diet J. and Tarantola A., 1990. Nonlinear inversion of seismic reflection data in a laterally invariant medium. *Geophysics*, **55**, 284–292, DOI: 10.1190/1.1442836.

Plessix R.E., 2006. A review of the adjoint-state method for computing the gradient of a functional with geophysical applications. *Geophys. J. Int.*, **167**, 495–503, DOI: 10.1111 /gji.2006.167.issue-2.

Pratt R.G. and Worthington M.H., 1990. Inverse theory applied to multisource cross-hole tomography. Part I: Acoustic wave-equation method. *Geophys. Prospect.*, **38**, 287–310.

Pratt R.G., Shin C. and Hicks G.J., 1998. Gauss-Newton and full Newton methods in frequency-space seismic waveform inversion. *Geophys. J. Int.*, **133**, 341–362, DOI: 10.1046/j.1365-246X.1998.00498.x.

Pyun S., Son W. and Shin C., 2009. Frequency-domain waveform inversion using an L1-norm objective function. *Explor. Geophys.*, **40**, 227–232, DOI: 10.1071/EG08103.

Ravaut C., Operto S., Improta L., Virieux J., Herrero A. and Dell'Aversana P., 2004. Multiscale imaging of complex structures from multifold wide-aperture seismic data by frequency-domain full-waveform tomography: application to a thrust belt. *Geophys. J. Int.*, **159**, 1032–1056, DOI: 10.1111/j.1365-246X.2004.02442.x.

Ren Z.M., LiuY. and Zhang Q.S., 2014. Multiscale viscoacoustic waveform inversion with the second generation wavelet transform and adaptive time–space domain finite-difference method. *Geophys. J. Int.*, **197**, 948–974, DOI: 10.1093/gji/ggu024.

Shin C. and Cha Y.H., 2008. Waveform inversion in the Laplace domain. *Geophys. J. Int.*, **173**, 922–931, DOI: 10.1111/j.1365-246X.2008.03768.x.

Tarantola A., 1984. Inversion of seismic reflection data in the acoustic approximation. *Geophysics*, **49**, 1259–1266, DOI: 10.1190/1.1441754.

Tarantola A., 1987. *Inverse Problem Theory*. 1st Edition. Elsevier, Amsterdam, The Netherlands.

Vigh D. and Starr E.W., 2008. 3D prestack plane-wave, full-waveform inversion. *Geophysics*, **73**, VE135–VE144, DOI: 10.1190/1.2952623.

Vigh D., Starr E.W. and Kapoor J., 2009. Developing earth models with full waveform inversion. *The Leading Edge*, **28**, 432–435, DOI: 10.1190/1.3112760.

Virieux J. and Operto S., 2009. An overview of full-waveform inversion in exploration geophysics. *Geophysics*, **74**, WCC1–WCC26. DOI: 10.1190/1.3238367.

Wang Y. and Rao Y., 2006. Crosshole seismic waveform tomography I: Strategy for real data application. *Geophys. J. Int.*, **166**, 1224–1236, DOI: 10.1111/j.1365-246X.2006.03030.x.

Warner M., Ratcliffe A., Nangoo T., Morgan J., Umpleby A., Shah N., Vinje V., Tekl I., Guasch L., Win C., Conroy G. and Bertrand A., 2013. Anisotropic 3D full-waveform inversion. *Geophysics*, **78**, R59–R80, DOI: 10.1190/geo2012-0338.1.

Xu K. and McMechan G.A., 2014. 2D frequency-domain elastic full-waveform inversion using time-domain modeling and a multistep-length gradient approach. *Geophysics*, **79**, DOI: 10.1190/GEO2013-0134.1.