Pre-stack Seismic Inversion with L1-2-norm Regularization Via a Proximal DC Algorithm and Adaptive Strategy

Shuangquan Chen (chensq@cup.edu.cn)
China University of Petroleum Beijing https://orcid.org/0000-0002-4069-9595
Guoquan Wang
China University of Petroleum Beijing

Research Article

Keywords: Pre-stack seismic inversion, norm regularization, Sparse constraint, Proximal-DC algorithm, Adaptive parameter selection

Posted Date: February 18th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1361049/v1

Pre-stack Seismic Inversion with L_{1-2}-norm Regularization Via a Proximal DC Algorithm and Adaptive Strategy

Guoquan Wang [1,2], Shuangquan Chen[1,2,*]

[1] State Key Laboratory of Petroleum Resources and Prospecting, China University of Petroleum (Beijing), Beijing, 102249, China

[2] CNPC Key Laboratory of Geophysical Prospecting, China University of Petroleum (Beijing), Beijing, 102249, China

[*]corresponding author: chensq@cup.edu.cn

Abstract

Seismic inversion in geophysics is an effective way to obtain underground rock properties from seismic survey data on the Earth's surface. Especially, we can obtain much more information to characterize subsurface geological structure and lithology via pre-stack seismic inversion due to offset information added to the inversion than by post-stack seismic inversion. However, pre-stack seismic inversion is usually a nonlinear and complicated process. In this article, we adopt a L_{1-2}-norm as a constraint on pre-stack seismic inversion, promoting the generation of a sparse solution. We also propose a novel pre-stack seismic inversion method that reduces the complexity of the solving method by utilizing an objective function decomposition scheme. Comparison of calculation time, accuracy and sparsity of the inversion solutions indicates that the proposed algorithm has better accuracy and robustness. Moreover, considering the difficulty of regularization parameter selection, we develop an adaptive parameter selection strategy based on Generalized Stein unbiased risk estimation (G-SURE) and incorporate it into the solving algorithm. The adaptive approach finds an appropriate regularization parameter in each iteration, and obtains the optimal solution directly, which is beneficial for improving computational efficiency. A synthetic data test verifies that the adaptive method can converge to the optimal solution iteratively in the case of arbitrary initial regularization parameters. Finally, in application to real field data, we explain why the adaptive method is the better choice even though adaptive and non-adaptive methods can obtain solutions with similar accuracy.

Keywords Pre-stack seismic inversion · L_{1-2} -norm regularization · Sparse constraint · Proximal-DC algorithm · Adaptive parameter selection

Article Highlights

- A novel pre-stack seismic inversion method is proposed based on a proximal difference-of-convex algorithm (pDCA).
- A new adaptive regularization parameter selection strategy is proposed based on Generalized Stein unbiased risk estimation (G-SURE).
- It was verified that one of the regularization parameters has a limited effect in L_{1-2}-norm.

Statements and Declarations

The authors have no competing interests to declare that are relevant to the content of this article.

Acknowledgements

This work is sponsored by National Natural Science Foundation of China (grant no. 42174130), the National Key R & D Program of China (grant no. 2018YFA0702501) and the Strategic Cooperation Technology Projects of CNPC and CUPB (grant no. ZLZX2020-03). It is published with the permission of the State Key Laboratory of Petroleum Resources and Prospecting. We also would like to express our sincere thanks to Dr. David Booth for proofreading.

1 Introduction

It is necessary to obtain velocity and density information representing the characteristics of an underground medium based on the pre-stack seismic gather with the help of optimization theory. This process is called seismic inversion, which is essentially an optimization problem in mathematics (Wang 2014). Seismic inversion methods generally include linear and nonlinear methods. Nonlinear inversion methods are easier for solving the global optimal solution because they continuously search in the solution space, but linear inversion significantly reduces the computational cost of nonlinear methods, and is more suitable for fast inversion of large-scale seismic data (Maurya et al. 2020). However, seismic inversion is generally an underdetermined problem, making the linear inversion algorithm more dependent on the initial model. To make linear inversion converge at the global extreme value and reduce the instability and multiplicity of the inversion solutions, a regularization constraint is generally used to solve the inverse problem in geophysics (Tarantola 2005).

The sparse constraint is common in signal processing; due to its ability in noise suppression, it has been applied in geophysics for a long time. Claerbout and Muir (1973) proved that L_1-norm could obtain much better results than L_2-norm in most cases for seismic numerical modelling. Mainly, with compressed sensing (CS) technology in signal processing (Tao et al. 2010), more researchers have focused widely on deconvolution, seismic inversion and seismic data reconstruction (Kazemi and Sacchi 2014; She et al. 2019; Ma 2013). In seismic inversion, Zhang and Castagna (2011) applied the basis pursuit (BP) algorithm to post-stack seismic data to solve the L_1-norm constrained objective function. This practice shows that the inverted impedance has a higher resolution than the traditional method. Total variation (TV) regularization is a special form of sparse constraint, which is not the L_1-norm of the parameters calculated directly but the L_1-norm of the calculated parameters after difference calculation. In this way, sparse edges can be controlled, and the parameters obtained by inversion show characteristics of periodic variation. In seismic inversion, TV regularization can highlight vertical variation characteristics of strata, and the elastic parameters of discontinuous variation can be obtained by inversion (Mozayan et al. 2018). Scholars have applied various methods, such as Iteratively Reweighted Least Squares (IRLS), the split-Bregman and the Alternating Direction Method of Multipliers (ADMM), to retrieve elastic parameters from post-stack data (Zhang et al. 2014; Liu and Yin 2015; Pan et al. 2017). Compared with post-stack data, pre-stack data contains more lithology and fluid information, and a relatively stable solution can also be obtained after extending the L_1-norm regularization method to the pre-stack AVA inversion (Li and Zhang 2017; Zhi et al. 2016).

L_{1-2}-norm is a recently proposed sparse constrained regularization term. It was first addressed by Lou et al. (2015) in the context of nonnegative least squares problems and group sparsity with applications to spectroscopic imaging. Some efforts have been made in an application and solving algorithm which has made L_{1-2}-norm gradually applied to geophysics. Wang et al. (2018) make use of the sparsity of L_{1-2}-norm to compensate for the attenuation of seismic data and effectively improve signal resolution. Wang et al. (2019) carried out pre-stack seismic inversion under the constraint of L_{1-2}-norm regularization, and improved the lateral continuity of inversion results by using *f-x* prediction filtering and obtained a high-quality elastic parameter inversion result. Huang et al. (2021) used L_{1-2}-norm regularization

to carry out AVA joint inversion based on time domain matching of PP- and PS-waves by using the DTW algorithm, and the algorithm showed good stability.

In fact, a more complex iterative algorithm is needed to solve the sub-problems in each iteration for pre-stack seismic inversion, a large-scale problem, by using L_{1-2}-norm, because the traditional difference-of-convex algorithm (DCA) may lead to high cost in large-scale inverse problems (Gotoh et al. 2018). Therefore, in this article, we propose a novel pre-stack seismic inversion scheme by using L_{1-2}-norm. The objective function is composed of a misfit function and a constraint term of L_{1-2}-norm. Based on the proximal DCA, we develop an optimization algorithm by reformulating the objective function as the difference between two convex functions. In each iteration of DCA, we extrapolate the last solution to obtain the starting point of the new iteration and then use the soft thresholding algorithm to calculate the optimal solution of the current iteration. Moreover, we introduce an adaptive regularization parameter selection method into the new algorithm, and propose a strategy to solve the problem that the amplitude of the inversion parameters cannot be well recovered in some cases. To verify the effectiveness of the algorithm and the adaptive parameter selection method, we use synthetic data and actual data to test, respectively. The inversion results verify that the new method is effective for the inverse problem constrained by L_{1-2}-norm in the pre-stack seismic inversion, and the adaptive parameter selection method is appropriate.

2 Theory and Methodology

2.1 Pre-stack Seismic Forward Model

According to the seismic convolution model (Robinson 1967), the pre-stack seismic gather can be represented as reflectivity coefficients for different angles convoluted with the seismic wavelet. A noise term should also be added for real field seismic records. Therefore, for an *N*-trace pre-stack angle gather, the forward modeling can be expressed as

$$\mathbf{s}(\theta_i) = \mathbf{w}(\theta_i) * \mathbf{r}(\theta_i) + \mathbf{n}(\theta_i), \quad i = 1, 2, \mathrm{L}\ , N \tag{1}$$

where $\mathbf{s}(\theta_i)$, $\mathbf{w}(\theta_i)$, $\mathbf{r}(\theta_i)$ and $\mathbf{n}(\theta_i)$ are the pre-stack seismic trace data, the seismic wavelet source, the reflectivity coefficient series and the noise at the *i*th angle of incidence θ_i, respectively. $\langle * \rangle$ represents the convolution operation. In Eq. (1), the reflection coefficient controls the amplitude of the seismic reflection wave and directly reflects the difference of impedance between the upper and lower layers of the reflection interface.

Mathematically, the pre-stack seismic forward modeling can be written as, omitting the noise term for simplicity

$$\mathbf{d} = \mathbf{G}\mathbf{m}, \tag{2}$$

where $\mathbf{d}$ represents the pre-stack seismic gather, $\mathbf{G}$ is the forward modeling operator, and $\mathbf{m}$ represents the model parameters vector.

2.2 L_{1-2}-norm in Seismic Inversion

The inverse problem of seismic inversion is highly ill-posed which leads to multiple solutions, especially for pre-stack seismic. By adding prior knowledge into the inverse problem, the regularization method in mathematics can solve the problem and obtain a consistent solution with the prior information. Because seismic inversion is mainly for the reflection coefficient sequence in the time domain, and the reflection coefficient of underground media has

obvious sparsity, we can use sparsity and add it into the inversion process. That is the sparse regularization in the optimization of inverse problems in mathematics.

At present, there are three commonly used regularization operators: L_0-norm, L_1-norm and L_2-norm, while L_0-norm and L_1-norm are referred to as sparse regularization constraints. As an optimal convex approximation of L_0-norm, L_1-norm can guarantee sparsity and it has better solving characteristics. However, the optimization problem containing L_0-norm belongs to an NP-hard problem, which is challenging to solve. Therefore, L_1-norm is generally used to construct the objective function, which aims to solve the sparse solution in image processing or geophysical inverse problems (Oldenburg et al. 1983; Yin 2015; Hamid and Pidlisecky 2015). As a recently emerging sparse constraint method, $L_{1\text{-}2}$-norm is getting more attention (Wang et al. 2018; Wang et al. 2019; Huang et al. 2021).

By comparing the four different regularization terms (L_0, $L_{1\text{-}2}$, L_1 and L_2), the advantage of $L_{1\text{-}2}$ norm in sparse constraints can be illustrated. The three-dimensional surface comparison of the four different norms (L_0, $L_{1\text{-}2}$, L_1 and L_2) function values is shown in Fig. 1. The function values are calculated by different norm terms using a set of two-dimensional data, and the projection of different surfaces on the two-dimensional plane is a contour map. Under the assumption of the misfit function being unchanged, the process of minimizing the objective function can be regarded as finding the solution of minimizing the regularization term in three-dimensional space. The lowest point of the surface is the optimal point to be searched. After the 3D surface is projected to a 2D plane, it can be found that the lowest point is close to the *x*-axis and *y*-axis on the 2D contours. An obvious conclusion is that the closer the norm curve is to the *x*-axis, the more possibility there is of obtaining the sparse solution by inversion. Therefore, the solution of L_0-norm regularization is the sparsest among these regularization terms. Although the solution of L_2-norm regularization is not sparse, it will make the solving process fast and stable. Another important conclusion is that L_{1-2}-norm is better than the traditional L_1-norm in sparsity as a regularization penalty term, and it is more likely to get a sparse solution.

According to the pre-stack seismic forward equation, Eq. (2), the pre-stack seismic inversion estimates the model parameters $\mathbf{m}$ by using the pre-stack seismic data $\mathbf{d}$. Moreover, Eq. (2) represents an underdetermined linear equation. Considering the layered distribution of underground media, a sparse regularization constraint is helpful to obtain sparse reflection coefficient sequences. As in the case of the seismic deconvolution method (Oldenburg et al. 1983), a stable and sparse solution is inverted by using the L_2-norm misfit function with L_1-norm regularization

$$J(\mathbf{m}) = \|\mathbf{Gm}-\mathbf{d}\|_2^2 + \lambda\|\mathbf{m}\|_1 . \tag{3}$$

In Eq. (3), λ is a trade-off parameter to balance the first error, or misfit, term and the second term.

By adopting $L_{1\text{-}2}$-norm to regularize the pre-stack seismic inverse problem, the constructed objective function includes a general misfit function and a sparse constraint regularization term

$$J(\mathbf{m}) = f(\mathbf{m}) + H(\mathbf{m}), \tag{4}$$

where $f(\mathbf{m}) = \|\mathbf{Gm}-\mathbf{d}\|_2^2$, $H(\mathbf{m}) = \lambda\left(\|\mathbf{m}\|_1 - \alpha\|\mathbf{m}\|_2\right)$. $H(\mathbf{m})$ is the $L_{1\text{-}2}$-norm regulariza-

tion penalty term, and $\alpha \in (0,1]$ is a constant, which can promote the generation of sparse solution (Lou et al. 2015).

2.3 Inversion Algorithm

For the objective function of the pre-stack seismic inversion in Eq. (4), we can use the difference-of-convex algorithm (DCA) to solve. In general, we need to set $H(\mathbf{m})$ in the form of the difference between two scalar norms

$$H(\mathbf{m}) = H_1(\mathbf{m}) - H_2(\mathbf{m}) = \lambda \|\mathbf{m}\|_1 - \lambda\alpha \|\mathbf{m}\|_2. \tag{5}$$

Therefore, the algorithm then transforms the solution of Eq. (4) into alternate iterations of two variables

$$\begin{cases} \mathbf{b}^k \in \partial H_2(\mathbf{m}^k) \\ \mathbf{m}^{k+1} = \arg\min\left(f(\mathbf{m}) + H_1(\mathbf{m})\right) - \left(H_2(\mathbf{m}^k) + \left\langle \mathbf{b}^k, \mathbf{m} - \mathbf{m}^k \right\rangle\right) \end{cases}, \tag{6}$$

where $\mathbf{b}_k$ is a subgradient of $H_2(\mathbf{m}_k)$ and $\langle \cdot \rangle$ denotes the inner product of two vectors. The key to Eq. (6) is solving $\mathbf{m}^{k+1}$ in a new convex optimization problem. A common solution is to introduce additional variables and use the Lagrange multiplier method to construct a form, which can be solved with the help of the ADMM algorithm (Yin et al. 2015; Wang et al. 2019).

The mathematical derivation has shown that the stability and convergence of DCA depend on the concrete decomposition form of $H(\mathbf{m})$ (Tao and An 1998). To get the closed-form solution of each subproblem, Gotoh et al. (2018) designed a new type of DCA, which is called proximal DCA (pDCA)

$$H(\mathbf{m}) = \left[\frac{L}{2}\|\mathbf{m}\|_2^2 + H_1(\mathbf{m})\right] - \left[\frac{L}{2}\|\mathbf{m}\|_2^2 + H_2(\mathbf{m})\right]. \tag{7}$$

where L is the Lipschitz constant and it can be obtained by calculating $\max\left(\mathrm{svd}\left(\mathbf{G}^{\mathrm{T}} * \mathbf{G}\right)\right)$. Through this algorithm, each component of DC decomposition will be strongly convex. Eq. (4) can be written as

$$J(\mathbf{m}) = \left[\frac{L}{2}\|\mathbf{m}\|_2^2 + H_1(\mathbf{m})\right] - \left[\frac{L}{2}\|\mathbf{m}\|_2^2 - f(\mathbf{m}) + H_2(\mathbf{m})\right]. \tag{8}$$

By taking the derivative with respect to $\mathbf{m}$ in Eq. (8) and making it equal to zero, we can solve the extremum problem of the objective function $J(\mathbf{m})$ to obtain the inversion solution in each iteration. The solution of the Eq. (8) is given by Wen et al. (2018)

$$\begin{aligned} \mathbf{m}^{k+1} &= \arg\min\left\{\left\langle \nabla f(\mathbf{m}^k) - \mathbf{b}^k, \mathbf{m}\right\rangle + \frac{L}{2}\|\mathbf{m} - \mathbf{m}^k\|_2^2 + H_1(\mathbf{m})\right\} \\ &= \arg\min\left\{\frac{L}{2}\left\|\mathbf{m} - \left(\mathbf{m}^k - \frac{1}{L}\left[\nabla f(\mathbf{m}^k) - \mathbf{b}^k\right]\right)\right\|_2^2 + H_1(\mathbf{m})\right\}. \end{aligned} \tag{9}$$

In order to accelerate the convergence speed, we can extrapolate a point $\mathbf{y}^k$ in the direction of gradient descent on the line between $\mathbf{m}^{k-1}$ and $\mathbf{m}^k$ according to the FISTA (Aster et al. 2012)

$$\mathbf{y}^k = \mathbf{m}^k + \omega(\mathbf{m}^k - \mathbf{m}^{k-1}), \tag{10}$$

where $\omega = (\theta_{t-1} - 1)/\theta_t$ with $\theta_t = \left(1 + \sqrt{1 + 4\theta_{t-1}^2}\right)/2$. FISTA is an upgraded version of ISTA,

which is often used to solve the L_1-norm regularized problem. Let $\mathbf{y}=\mathbf{m}^{k+1}$,

$$\mathbf{m}^{k+1}=\arg\min\left\{\left\langle\nabla f(\mathbf{y}^k)-\mathbf{b}^k,\mathbf{y}\right\rangle+\frac{L}{2}\left\|\mathbf{y}-\mathbf{y}^k\right\|_2^2+H_1(\mathbf{y})\right\}. \tag{11}$$

We use a variable $\mathbf{h}^k$ to represent the constant term. Eq. (11) can be rewritten as

$$\mathbf{m}^{k+1}=\arg\min\left\{\frac{L}{2}\left\|\mathbf{y}-\mathbf{h}^k\right\|_2^2+\lambda\left\|\mathbf{y}\right\|_1\right\}. \tag{12}$$

The variable $\mathbf{h}^k$ is the algebraic result of the solution $\mathbf{y}^k$ in the last iteration

$$\mathbf{h}^k=\mathbf{y}^k-\frac{1}{L}\left[\nabla f\left(\mathbf{y}^k\right)-\mathbf{b}^k\right]. \tag{13}$$

Taking the derivative with respect to $\mathbf{y}$ in Eq. (12),

$$L(\mathbf{y}-\mathbf{h}^k)+\lambda\mathrm{sgn}(\mathbf{y})=0. \tag{14}$$

Thus, the solution is given by

$$\mathbf{y}=\mathbf{h}^k-\frac{\lambda}{L}\mathrm{sgn}\left(\mathbf{y}\right). \tag{15}$$

Since $\mathbf{h}^k$ and λ/L are constants in Eq. (15), we can get the solution by the soft thresholding algorithm (Aster et al. 2012). When $\mathbf{h}^k>\lambda/L$, if $\mathbf{y}<0$, then $\mathrm{sgn}(\mathbf{y})=-1$ and $\mathbf{y}=\mathbf{h}^k+\lambda/L>0$, and there is a contradiction. But if $\mathbf{y}>0$, then $\mathrm{sgn}(\mathbf{y})=1$ and $\mathbf{y}=\mathbf{h}^k-\lambda/L>0$, which is the correct solution in accordance with the conditions. In the same way, we can also derive the solution when $\mathbf{h}^k<\lambda/L$. In summary, the solution can be expressed as

$$\mathbf{y}=\begin{cases}\mathbf{h}^k-\dfrac{\lambda}{L}, & \mathbf{h}^k>\dfrac{\lambda}{L}\\ \mathbf{h}^k+\dfrac{\lambda}{L}, & \mathbf{h}^k<-\dfrac{\lambda}{L}\\ 0, & otherwise\end{cases}. \tag{16}$$

According to the above statements, the procedure for our inversion method is as follows:

1) Set the initial solution $\mathbf{m}^0$, tolerance value ε and the number of iterations N.

2) Calculate the Lipschitz constant L.

3) For each iteration k = 1, 2, 3, ⋯, N:

a) Update ω and $\mathbf{b}^k$

$$\mathbf{b}^k=\lambda\alpha\,\mathrm{sgn}\left(\mathbf{m}^k\right) \tag{17}$$

b) Compute $\mathbf{y}^k$

$$\mathbf{y}^k=\mathbf{m}^k+\alpha(\mathbf{m}^k-\mathbf{m}^{k-1}), \tag{18}$$

and $\mathbf{h}^k$

$$\mathbf{h}^k=\mathbf{y}^k-\frac{1}{L}\left[\mathbf{G}^{\mathrm{T}}\mathbf{G}y^k-\mathbf{G}^{\mathrm{T}}\mathbf{d}-\mathbf{b}^k\right]. \tag{19}$$

c) Update $\mathbf{m}^{k+1}$ by the soft thresholding algorithm

$$\mathbf{m}^{k+1}=sthresh(\mathbf{h}^k,\frac{\lambda}{L}) \tag{20}$$

d) Check the convergence condition

$$\frac{\left\|\mathbf{m}^{k+1}-\mathbf{m}^{k}\right\|_2}{1+\left\|\mathbf{m}^{k+1}\right\|_2}<\varepsilon \tag{21}$$

3 Synthetic Data Test and Analysis

3.1 Strata Model

A multilayer geological model (Fig. 2) is created and used to generate a pre-stack seismic gather dataset to elucidate the robustness, convergence speed and stability of the proposed inversion method. The detailed model parameters are shown in Table 1, including the P-wave velocity V_{P}, S-wave velocity V_{S} and bulk density ρ of each layer. The model consists of 21 layers with different thicknesses, which generates 20 spike reflection coefficients at the interface between two adjacent strata. According to the vertical variation of the model parameters, there will be some small reflection coefficients in the spike sequence (e.g. 2nd, 9th, 11th). The pre-stack reflection coefficient series $r(t,\theta)$ at different incidence angles are generated by the Aki-Richards approximate equation (Aki and Richards 1980). We assume that the wavelet is a zero-phase Ricker wavelet whose peak frequency is 30Hz. The synthetic gathers with different incidence angles (5°, 15° and 25°) can be obtained by convolution. Moreover, we add Gaussian random noise with different signal-to-noise ratio (SNR, are 15, 10 and 5, respectively) to verify the robustness of the inversion method in the following section. Noise-free and noisy records are shown in Fig. 3.

3.2 Analysis of $L_{1\text{-}2}$-norm Sparse Constraint

To illustrate the sparsity advantage of the L_{1-2}-norm over the L_1-norm, inversion methods based on different regularization terms are applied to the synthetic data. First, we adopted the FISTA (Pérez et al. 2012) method to do the pre-stack inversion, which utilizes the L_1-norm regularization term to solve the objective function. Furthermore, our proposed algorithm is implemented, in which L_{1-2}-norm regularization is adopted. Both algorithms stop the iteration process under the same convergence condition, that is, Eq. (21). The tolerance value is set to 1×10^{-5}. The input data are synthetic data with SNR = 5 in the inversion. The inversion results are shown in Fig. 4, including the reflection coefficient inversion results for three angles. The gray bars represent the true reflectivity series, and the black bars are the inversion results.

In FISTA, the regularization parameter is assigned to be 5×10^{-2}, which was selected by the L-curve. In our proposed method, there are two regularization parameters that need to be determined, of which $\lambda=8\times10^{-3}$ and $\alpha=1\times10^{-3}$ in this example. The iteration number of the new algorithm and FISTA are 163 and 59, respectively. For the environment we tested (i7-10700F processor, 32G memory, 64-bit processing system), it took about 3.3520 seconds and 0.4292 seconds, respectively. The comparison shows that FISTA has a faster speed when reaching the same convergence condition. At the same time, we carried out error analysis of the different inversion results. The correlation coefficient (CC) is 0.9762, and the normalized root-mean-square error (NRMSe) is 1.46% in FISTA. Correspondingly, our proposed method adopted L_{1-2}-norm regularization can significantly reduce the error of the solution, the CC is increased to 0.9971, and the NRMSe is reduced to 0.55%. Since most of the elements

in the sparse solution are zero, the difference in error analysis may not be as significant as in Fig. 4. Obviously, the inversion results based on L_1-norm regularization are not consistent with the amplitude of the real reflection coefficient at some time points. L_{1-2}-norm regularization can make up for this deficiency under the condition of sparsity. The analysis of the first numerical experiment shows that the L_{1-2}-norm regularization can significantly improve the accuracy of the inversion solution compared with the common L_1-norm regularization.

3.3 Analysis of Inversion Solving Method

In the methodology part, we briefly introduced common methods for solving the L_{1-2}-norm regularization, one of which is DCA-ADMM. For the detailed derivation of DCA-ADMM, please refer to Yin et al. (2015). However, it is quite different from our new algorithm, whether the construction of DCA, or the subsequent iterative solution method. Compared with the DCA-ADMM method, our proposed algorithm has a more straightforward solving process in L_{1-2}-norm regularization since there is no need to update multiple variables alternately.

We also use synthetic data with SNR = 5 to compare the advantages and disadvantages of two different solving methods. The inversion results are shown in Fig. 5. Similarly, the gray and black rods represent the real reflection coefficient and the inversion result. Different to the previous comparison, the solutions based on the same regularization have no noticeable difference in Fig. 5. Combined with the results of error analysis, there is little difference between the DCA-ADMM and new algorithm in CC (0.9963 and 0.9971), NRMSe (0.57% and 0.55%) and program running time (3.1973 seconds and 3.6673 seconds) in this example. The main difference between the two algorithms is sparsity. The previous section introduced the model with 20 reflection interfaces, so in the case of three incidence angles, the L_0-norm of the real reflection coefficient is 60, which is called sparsity. By calculating the norm of the solution obtained by the two methods, we can determine the ability of the two algorithms to obtain sparse solutions. The sparsity of the solution obtained by using the proposed algorithm in this paper is 80. However, the sparsity of the solution obtained by using DCA-ADMM reaches 1497, which means that the solution is not strictly sparse. There are a lot of tiny values in many places that are supposed to be zero. Even if values below 1.9×10^{-4} are ignored, which is the calculation result of multiplying the minimum real reflection coefficient amplitude 1.9×10^{-3} by 0.1, and the L_0-norm is reduced to 158, the sparsity is still lower than the solution obtained by the new method. It can be observed in Fig. 5a-c, where there are fewer small spikes than in Fig. 5d-f. Therefore, our proposed new method is more suitable for sparse constraint inversion.

3.4 Analysis of Noise Effects

Different amounts of noise are generally included in pre-stack synthetic data. It is a crucial problem to ensure the stability of the inversion method without the influence of noise. The SNRs of the synthetic data used for inversion analysis are 5, 10 and 15. The corresponding results are shown in Fig. 6 by using the new method. Regardless of the noise level, a suitable regularization parameter can be determined to make the inversion results match well with the actual model. The CC of the inversion result reaches 0.9996 when the SNR is 15. Even if the SNR decreases to 5 due to a noise increase in the gather, the CC is still as high as 0.9971. The analysis shows that the proposed algorithm is robust on noisy seismic data.

4 Adaptive Regularization Parameter Selection Method

4.1 Analysis of Regularization Parameters

The last problem that needs to be considered is the selection of regularization parameter in inversion processing. In our proposed algorithm, there are two regularization parameters λ and α that need to be determined. The effect of λ is like the regularization parameter in Tikhonov regularization, which is used to adjust the weight of the misfit function and regularization constraint. As a weight parameter within the L_{1-2} norm, the influence of α is rarely analyzed. Here, we update the sparsity of the model parameters by adding or combining thin layers based on the original multilayer model. The sparsity of the two new models is 30 and 90. By changing α in a wide range, we analyze the influence of weight parameter on the inversion results by using three models, which are different in sparsity. We implemented the algorithm on the synthetic records related to these models and calculated the NRMSe of the solutions. Fig. 7 shows the variation of NRMSe with regularization parameter α. We find that the influence of parameter α on the inversion results is closely related to its value. According to the analysis, the error of the solution will decrease rapidly when α is less than 1, and then settles into a stable state. Likewise, the trend of the variation curve is like the L-curve, where there is an inflection point. More importantly, the influence of α on the solution tends to be stable when it is less than 0.1. Therefore, to reduce the difficulty of solving the inverse problem, we set the parameter α to a fixed value of 0.01 in the following test.

4.2 Adaptive Selection Method

Since the algorithm only updates the solution in a critical step and does not need to update multiple variables alternately like ADMM, it is possible to use other methods to obtain regularization parameters adaptively. By reasonably changing the regularization parameters to adjust the weight of the regularization term, L_{1-2}-norm regularization can effectively solve the ill-posedness of the inverse problem and help us obtain a sparse solution, as proved in the previous section. Generally, the larger the regularization parameter, the greater the dependence of the inversion solution on the initial model, and the less sensitive to small perturbations (Thore 2015). The strategy to select an appropriate regularization parameter quickly and effectively is an important problem to be solved. Common methods include L-curve (Hansen 1992) and GCV (Golub et al. 1979) and have been widely used in seismic inversion (Gholami 2016; Huang et al. 2017). To adjust the regularization parameters adaptively during inversion iteration, we introduce a parameter selection method based on generalized Stein unbiased risk estimation (G-SURE) (Eldar 2008).

In order to find a function $F(\mathbf{d})=\hat{\mathbf{m}}$ with noisy data $\mathrm{d}(\theta_i)$ to minimize the mean square error (MSE), where $\hat{\mathbf{m}}$ represents an arbitrary estimate of $\mathbf{m}$, Stein (1981) proposed Stein's unbiased risk estimate (SURE), which is proved to be better than the common maximum likelihood estimation. To further extend the applicability of the SURE to a broad class of problems, Eldar (2008) proposed a generalized SURE by adding a penalty term to the expression. For example, a sufficient statistic $\mathbf{u}$ for estimating $\mathbf{m}$ in linear Gaussian model is given by

$$\mathbf{u}=\mathbf{G}^{\mathrm{T}}\mathbf{d}. \tag{22}$$

In any case, $F(\mathbf{d})$ can be expressed as $F(\mathbf{u})$ based on sufficient statistic $\mathbf{u}$. We can express the MSE of $\hat{\mathbf{m}}$ as

$$E\left\{\left\|\hat{\mathbf{m}}-\mathbf{m}\right\|_2^2\right\}=\left\|\mathbf{m}\right\|_2^2+E\left\{\left\|F(\mathbf{u})\right\|_2^2\right\}-2E\left\{F^T(\mathbf{u})\mathbf{m}\right\}, \tag{23}$$

where $\|\mathbf{m}\|_2^2$ is a constant. The purpose of G-SURE is to minimize MSE, so we define

$$v(F,\mathbf{m})=E\left\{\left\|F(\mathbf{u})\right\|_2^2\right\}-2E\left\{F^T(\mathbf{u})\mathbf{m}\right\}, \tag{24}$$

where $v(F,\mathbf{m})$ is a function that estimates the parameters of the model, and it can help us estimate MSE accurately. In Eq. (24), the size of MSE depends on the actual model parameter $\mathbf{m}$ and $F(\mathbf{u})$, while $\mathbf{m}$ is unknown. Specifically, Eldar (2009) solved the problem by constructing a function $g(F(\mathbf{u}))$ which satisfies

$$E\left\{g(F(\mathbf{u}))\right\}=E\left\{F^T(\mathbf{u})\mathbf{m}\right\}. \tag{25}$$

Then, unbiased estimation of $v(F,\mathbf{m})$ can be expressed as

$$\hat{v}(F)=\left\|F(\mathbf{u})\right\|_2^2-2g(F(\mathbf{u})). \tag{26}$$

The details of Eq. (26) are provided in Appendix A. The unbiased risk estimation based on $F(\mathbf{u})$ can be expressed as

$$\hat{v}(F)=\left\|F(\mathbf{u})\right\|_2^2+2Tr\left[\frac{\partial F(\mathbf{u})}{\partial \mathbf{u}}\right]-2F^T(\mathbf{u})\left(\mathbf{G}^{\mathrm{T}}\mathbf{G}\right)^{-1}\mathbf{u}. \tag{27}$$

In each solving algorithm, once the critical steps of the iterative solution are determined, the detailed form of Eq. (27) can be obtained. By deriving $\hat{v}(F)$ for the regularization parameter λ and making it equal to 0, the regularization parameters in each iteration can be calculated.

Now we extend the approach to the algorithm proposed in the paper. Substituting Eq. (19) and Eq. (22) into Eq. (16),

$$F(\mathbf{u})=\begin{cases}\mathbf{y}^k-\dfrac{1}{L}\left[\mathbf{G}^{\mathrm{T}}\mathbf{G}\mathbf{y}^k-\mathbf{u}-\mathbf{b}^k\right]-\dfrac{\lambda}{L}, & h^k>\dfrac{\lambda}{L}\\ \mathbf{y}^k-\dfrac{1}{L}\left[\mathbf{G}^{\mathrm{T}}\mathbf{G}\mathbf{y}^k-\mathbf{u}-\mathbf{b}^k\right]+\dfrac{\lambda}{L}, & h^k<-\dfrac{\lambda}{L}\\ 0, & otherwise\end{cases}. \tag{28}$$

We suppose that there are p components greater than λ/L and q components less than $-\lambda/L$ in the k_{th} iteration. By arranging them together, we can form two vectors with dimensions p and q, marked as $\mathbf{a}_1$ and $\mathbf{a}_2$. In the same iteration, the components of $(\mathbf{G}^{\mathrm{T}}\mathbf{G})^{-1}\mathbf{u}$ corresponding to the indexes $\mathbf{a}_1$ and $\mathbf{a}_2$ in $\hat{v}(F)$ are recorded as $\mathbf{w}_1$ and $\mathbf{w}_2$, respectively. Therefore, we can get the expression of each term in Eq. (27)

$$\left\|F(\mathbf{u})\right\|_2^2=\sum_{i=1}^{p}\left(\mathbf{a}_1^i-\frac{\lambda}{L}\right)^2+\sum_{j=1}^{q}\left(\mathbf{a}_2^j+\frac{\lambda}{L}\right)^2, \tag{29}$$

$$Tr\left[\frac{\partial F(\mathbf{u})}{\partial \mathbf{u}}\right]=\frac{p+q}{L}, \tag{30}$$

and

$$F^T(\mathbf{u})\left(\mathbf{G}^{\mathrm{T}}\mathbf{G}\right)^{-1}\mathbf{u}=\sum_{i=1}^{p}\left(\mathbf{a}_1^i-\frac{\lambda}{L}\right)\mathbf{w}_1^i+\sum_{j=1}^{q}\left(\mathbf{a}_2^j+\frac{\lambda}{L}\right)\mathbf{w}_2^j. \tag{31}$$

We can get that the specific form of Eq. (27) under the framework of the new algorithm

$$\hat{v}(F)=\sum_{i=1}^{p}\left(\mathbf{a}_1^i-\frac{\lambda}{L}\right)^2+\sum_{j=1}^{q}\left(\mathbf{a}_2^j+\frac{\lambda}{L}\right)^2+\frac{2(p+q)}{L}-2\left[\sum_{i=1}^{p}\left(\mathbf{a}_1^i-\frac{\lambda}{L}\right)\mathbf{w}_1^i-\sum_{j=1}^{q}\left(\mathbf{a}_2^j+\frac{\lambda}{L}\right)\mathbf{w}_2^j\right]$$
$$=\sum_{i=1}^{p}\left[\left(\mathbf{a}_1^i-\frac{\lambda}{L}\right)^2-2\left(\mathbf{a}_1^i-\frac{\lambda}{L}\right)\mathbf{w}_1^i\right]+\sum_{j=1}^{q}\left[\left(\mathbf{a}_2^j+\frac{\lambda}{L}\right)^2-2\left(\mathbf{a}_2^j+\frac{\lambda}{L}\right)\mathbf{w}_2^j\right]+\frac{2(p+q)}{L}. \tag{32}$$

To obtain the optimal regularization parameter λ^k in k_{th} iteration, we have that

$$\frac{d\hat{v}}{d\lambda}=\frac{2\lambda(p+q)}{L^2}-\frac{2}{L}\sum_{i=1}^{p}(\mathbf{a}_1-\mathbf{w}_1)+\frac{2}{L}\sum_{j=1}^{p}(\mathbf{a}_2-\mathbf{w}_2). \tag{33}$$

We can conclude that

$$\lambda^k=\frac{L\left[\sum_{i=1}^{p}(\mathbf{a}_1-\mathbf{w}_1)-\sum_{j=1}^{q}(\mathbf{a}_2-\mathbf{w}_2)\right]}{p+q}. \tag{34}$$

It is worth noting that the solution is unstable if $(\mathbf{G}^{\mathrm{T}}\mathbf{G})^{-1}\mathbf{u}$ is calculated directly in Eq. (31); we can solve the problem by calculating the generalized inverse $\mathbf{G}^{\dagger}$ (Infante-Pacheco et al. 2020). Both singular value decomposition (SVD) and truncated singular value decomposition (TSVD) are alternative methods. If the latter method is selected, the GCV method can determine the truncation parameter.

4.3 Applicability of $L_{1\text{-}2}$-norm Regularization

To study the applicability of this adaptive strategy in L_{1-2}-norm regularization, we use the noisy synthetic data (SNR = 5) in numerical examples as the input data for verification. Before we apply the adaptive strategy to seismic data, a significant problem is the determination of iterative convergence conditions. Once the condition is met, the inversion process is suspended to avoid meaningless computation. In the numerical experiments, we found that the sparsity of the solution will be continuously reduced if Eq. (21) is taken as the convergence condition in the adaptive process, which will cause the components of the solution to disappear. To illustrate this process, we tested the adaptive parameter selection method with three different regularization parameters as initial values independently and recorded the regularization parameters λ in the adaptive iterative process.

Fig. 8 shows the recorded results in this process, and depicts the variation of solution and regularization parameters. The change of the solution is represented by the L_{1-2} norm, which reflects sparsity. The three initial parameters are 1×10^{-3}, 1×10^{-5} and 1×10^{-7}, corresponding to the light gray, black and dark gray lines respectively in Fig. 8. During the initial iteration, there is a significant increase in the regularization term. The reason for this change is that the initial solution of the inversion is zero, and the solution is rapidly searching in the direction of descending gradient, while the solution is not sparse. After several times of jitter reduction, the value of the regularization term will decrease sharply, and the regularization parameter λ will move away from the low value range and increase significantly in a few successive iterations. In this case, the solution is optimal in the whole iterative process. If the iterative process does not stop, then a following adaptive iterative process will continue to increase the sparsity of the solution until almost all the elements in the inverse solution are zero, which will obviously obtain a solution that is different from the actual model. Eq. (21)

is not satisfied at the optimal points, which shows that it is not suitable as an iterative convergence condition.

Combined with the above analysis results, we choose the following conditions as the convergence condition during iteration

$$\left(\|\mathbf{m}\|_1^k - \alpha\|\mathbf{m}\|_2^k\right) - \left(\|\mathbf{m}\|_1^{k-1} - \alpha\|\mathbf{m}\|_2^{k-1}\right) < 0 \text{ and } \lambda^k - \lambda^{k-1} > \varepsilon. \quad (35)$$

The establishment of a new convergence condition combines the rules of parametric variation in Fig. 8. Obviously, Eq. (35) shows that the slope of the regularization term curve is positive, which is opposite to that of the regularization parameter variation curve, and the difference of the regularization parameters between the k and $k+1$ iteration is greater than the tolerance ε.

In the synthetic data test, we set the tolerance ε to 0.03 and the parameter α to 0.01. Fig. 9a-c shows the inversion results using these three regularization parameters directly, and Fig. 9d-f shows the inversion results using the adaptive adjustment strategy, where they are only used as initial values. It is worth mentioning that we only compare the situation when the incidence angle is 5° which is like that of other incidence angles. Obviously, a wrong solution can be obtained by directly applying the inversion method in the case of inappropriate regularization parameters. After applying the adaptive strategy, the sparsity of the solution is greatly improved compared with Fig. 9a-c. Nevertheless, the solution is still not optimal because the amplitude of the reflection coefficient has not been well recovered, even if the reflection coefficient appears at the correct time. Moreover, we can adjust the amplitude of the spike to the appropriate value with the help of a hybrid FISTA least-squares strategy (Pérez et al. 2013).

The main idea of the hybrid strategy is to obtain the optimized solution $\mathbf{m}_{hybrid}$ by rewriting the forward matrix $\mathbf{G}$ to $\mathbf{B}$, which is a dimension reduction form. The optimized solution $\mathbf{m}_{hybrid}$ is computed by the least square method

$$\mathbf{m}_{hybrid} = \mathbf{m}_{spike} + \left(\mathbf{B}^{\mathrm{T}}\mathbf{B}\right)^{-1}\left(\mathbf{d} - \mathbf{B}\mathbf{m}_{spike}\right), \quad (36)$$

where $\mathbf{m}_{spike}$ is the rearranged solution by removing the zero element. The results are shown in Fig. 9g-i. After applying the hybrid strategy, the CC of the inversion results is 0.9977, 0.9973 and 0.9967, respectively. Therefore, the application of the hybrid strategy can effectively improve the quality of inversion. Considering that the selection of fixed regularization parameters in the actual data may lead to uncertainty at the non-well location, the adaptive method can be applied to the inversion of real data as an optional method.

5 Real Data Application

Finally, we apply the inversion method into real field data to verify the applicability of our proposed new algorithm and adaptive parameters selection strategy. To enhance the SNR of pre-stack gathers, we utilize partially stacked gathers to do the pre-stack inversion from the original gathers. The near-, mid- and far-angles partially stacked gathers are obtained, and the corresponding incidence angles are 8°, 16° and 24°, respectively. The time interval of the target layer ranges from 3100 to 4400ms on the seismic profile (Fig. 10a-c), in which there are apparent strong amplitude anomalies and fault development, as it is a typical sandstone reservoir. The frequency range of seismic data is about 4~80 Hz, and the location of one drilled oil well is at CDP number 292.

Before applying the inversion method, it is necessary to accurately obtain the time-depth relationship (TDR) by well-seismic calibration. The corresponding range of well logging data on the seismic profile is 3660 to 4100ms. Three angle-dependent wavelets are estimated from the partially stacked seismic sections, respectively. To show the accuracy of the TDR and estimated wavelet, we use the velocity and density logging data in the well to make synthetic records, which were compared with the field data at the well location. The result is shown in Fig. 11, where the x-axis represents incidence angle, and the y-axis represents time. The synthetic data is shown as the black curve and the field data as the red curve. We analyzed the similarity of seismic traces at different incidence angles. The CC is 0.8813, 0.8751 and 0.7300, which represent incidence angles of 8^{o}, 16^{o} and 24^{o}, respectively. The analysis results show that the synthetic records are in good agreement with the field data in the time domain. And the SNR of the far-angle stack data is not as high as the near-angle stack.

In the case of applying the general inversion method to the actual data, we first need to determine the appropriate regularization parameters according to the seismic data at the well location and then extend the parameters to the seismic data volume. The regularization parameter selection method proposed in this paper can adaptively select the appropriate regularization parameters according to the inversion data and the extracted wavelet. It is no longer necessary to extend the regularization parameters determined by synthetic data to the whole dataset.

As an effective parameter to identify reservoir properties, seismic impedance can be obtained via the following equation according to the convolution model

$$EI(t) = EI(t_0) * \exp\left[2 * C * r(t)\right], \tag{37}$$

where EI is elastic impedance in time domain t, C is an integral matrix and $\exp[\cdot]$ denotes exponential operation (Zhang et al. 2014). Here, we adopt Eq. (37) to calculate the seismic impedance from the inverted reflectivity, which is relative impedance. Moreover, for real field data, we can compare it with the result calculated by the logging data. Considering the attenuation in seismic wave propagation, the logging data should be processed by a low-pass filter before calculating relative impedance. At the same time, in the process of calculating seismic impedance by using Eq. (37), if there is an error in the reflection coefficient at a specific time, there will be a low frequency cumulative error in the whole integration process. To avoid cumulative errors, the usual method is to filter the calculation results, which will cause the results to be band-limited, so the calculation results of logging data need to be processed by the same filter (Wang et al. 2019). Fig. 12 compares the actual well-logging relative impedance trace and the inverted relative impedance traces by using two inversion methods. The black solid line represents the relative impedance by filtering the well-logging curves. The gray dashed line and gray dotted line represent the inversion results of the adaptive and non-adaptive method, with CC of 0.7585 and 0.7521, respectively. Although the CC between two inversion methods is equivalent, significant errors may occur if the optimal regularization parameters obtained from the trace at the well location are extended to other traces.

Through the comparison of inversion results at the well location, the accuracy and effectiveness of our proposed inversion method are verified. Therefore, we inverted the whole seismic section data (shown in Fig. 10). Three angle reflectivity series sections are obtained by using the proposed inversion method and the adaptive parameters selection strategy for real field data processing. Moreover, we can calculate the relative seismic impedance using Eq. 37, which is based on the convolution model of reflectivity coefficient. The inversion

results for three angle stack seismic sections are shown in Fig. 13. Comparing the near- and mid-angles relative impedance sections, at 3700ms, there is a bright spot, and its impedance is significantly higher than that of the surrounding rocks, so it can be considered as a favorable exploration target. The relative impedance section of the far-angle stacked seismic section (Fig. 13c) is not consistent with the inversion results of the near- and mid-angles stacked seismic sections, but it is consistent with the pre-stack seismic gathers as shown in Fig. 11 and the stacked seismic section as shown in Fig. 10c.

6 Conclusions

We have combined the sparse regularization of $L_{1\text{-}2}$-norm and the proximal difference-of-convex algorithm (pDCA) to implement a pre-stack seismic inversion technique to obtain reflectivity coefficients from pre-stack seismic gather data. The analysis of synthetic data inversion results verified that L_{1-2}-norm is better than the traditional L_1-norm in sparsity as a regularization penalty term. At the same time, the analysis of the inversion solving method and noise effects indicated that our implemented inversion method by using the pDCA is also more suitable for sparse constraint inversion than the DCA-ADMM, which is a valuable inversion algorithm used in seismic inversion.

We presented an adaptive hybrid strategy, discussing the influence of regularization parameter α in $L_{1\text{-}2}$-norm and deriving the adaptive selection method of regularization parameter λ based on G-SURE, which could reduce the computation time to obtain the appropriate parameter. The analysis of different sparsity inversion of synthetic data inversion indicated that our adaptive parameter section method could effectively improve the quality of inversion results. Considering that the selection of fixed regularization parameters in real data may lead to uncertainty at the non-well location, the adaptive method can be applied to the inversion of real data as an optional method.

We presented an application to real field data by using the proposed pre-stack seismic inversion and adaptive hybrid strategy for selection of regularization parameters. The results showed that our inversion method is effective and stable for real field seismic data, which is generally contaminated by ambient noise. In this paper, we calculated the relative impedance from the inverted reflectivity coefficient series in the time domain according to the convolution model. The relative impedance is band-limited because the seismic data is band-limited, and filtering is required to eliminate the low frequency accumulated error in the integration process. More importantly, the elastic parameters cannot be directly inverted from the relative impedance; this requires wide-band impedance data based on a low-frequency model, but this is not the focus of this paper. Considering the wider application of sparse constraints in geophysics as reviewed in the introduction, our proposed inversion method based on L_{1-2}-norm can be applied to many real problems, such as improving seismic resolution as a deconvolution method, which is another important application direction.

Appendix

To get a concrete form of function $\hat{v}(F)$, firstly we need to determine $g(F(\mathbf{u}))$. In mathematics, the MSE can be written in the form of an integral

$$E\{g(F(\mathbf{u}))\} = \int F^T(\mathbf{u})\mathbf{m} f(\mathbf{u},\mathbf{m})\,d\mathbf{u}. \tag{A-1}$$

Under the assumption of zero-mean Gaussian distribution, the distribution function of u can be expressed as

$$\begin{aligned} f(\mathbf{u};\mathbf{m}) &= K\exp\left\{-\frac{1}{2}\mathbf{u}^{\mathrm{T}}\left(\mathbf{G}^{\mathrm{T}}\mathbf{G}\right)^{-1}\mathbf{u}\right\}\exp\left\{\mathbf{m}^{\mathrm{T}}\mathbf{u}-\frac{1}{2}\mathbf{m}^{\mathrm{T}}\mathbf{G}^{\mathrm{T}}\mathbf{G}\mathbf{m}\right\} \\ &= \varphi(\mathbf{u})\exp\left\{\mathbf{m}^{\mathrm{T}}\mathbf{u}-\frac{1}{2}\mathbf{m}^{\mathrm{T}}\mathbf{G}^{\mathrm{T}}\mathbf{G}\mathbf{m}\right\} \end{aligned}, \quad \text{(A-2)}$$

where K is a constant only related to Gaussian distribution and $\varphi(\mathbf{u}) = K\exp\left\{-\frac{1}{2}\mathbf{u}^{\mathrm{T}}\left(\mathbf{G}^{\mathrm{T}}\mathbf{G}\right)^{-1}\mathbf{u}\right\}$. Eq. (A-1) could be rewritten as

$$E\left\{g\left(F(\mathbf{u})\right)\right\} = \sum_i \int F_i(\mathbf{u}) m_i \varphi(\mathbf{u}) \exp\left\{\mathbf{m}^{\mathrm{T}}\mathbf{u}-\frac{1}{2}\mathbf{m}^{\mathrm{T}}\mathbf{G}^{\mathrm{T}}\mathbf{G}\mathbf{m}\right\} d\mathbf{u} \quad \text{(A-3)}$$

According to the derivative property of the exponential function and the partial integration method, we have that

$$\int F_i(\mathbf{u}) m_i \varphi(\mathbf{u}) \exp\left\{\mathbf{m}^{\mathrm{T}}\mathbf{u}-\frac{1}{2}\mathbf{m}^{\mathrm{T}}\mathbf{G}^{\mathrm{T}}\mathbf{G}\mathbf{m}\right\} d\mathbf{u} = -\int \frac{\partial F_i(\mathbf{u})\varphi(\mathbf{u})}{\partial u_i} \exp\left\{\mathbf{m}^{\mathrm{T}}\mathbf{u}-\frac{1}{2}\mathbf{m}^{\mathrm{T}}\mathbf{G}^{\mathrm{T}}\mathbf{G}\mathbf{m}\right\} du_i \quad \text{(A-4)}$$

Substituting Eq. (A-4) into Eq. (A-1),

$$E\left\{g\left(F(\mathbf{u})\right)\right\} = -E\left\{Tr\left(\frac{\partial F(\mathbf{u})}{\partial \mathbf{u}}\right)\right\} - E\left\{F^T(\mathbf{u})\frac{\partial \ln \varphi(\mathbf{u})}{\partial \mathbf{u}}\right\}. \quad \text{(A-5)}$$

According to the above derivation, the specific form of $g\left(F(u)\right)$ can be obtained:

$$\begin{aligned} g\left(F(\mathbf{u})\right) &= -Tr\left(\frac{\partial F(\mathbf{u})}{\partial \mathbf{u}}\right) - F^T(\mathbf{u})\frac{\partial \ln \varphi(\mathbf{u})}{\partial \mathbf{u}} \\ &= -Tr\left(\frac{\partial F(\mathbf{u})}{\partial \mathbf{u}}\right) - F^T(\mathbf{u})\left(\mathbf{G}^{\mathrm{T}}\mathbf{G}\right)^{-1}\mathbf{u}. \end{aligned} \quad \text{(A-6)}$$

Now,

$$\begin{aligned} \hat{v}(F) &= \left\|F(\mathbf{u})\right\|_2^2 - 2g\left(F(\mathbf{u})\right) \\ &= \left\|F(\mathbf{u})\right\|_2^2 + 2Tr\left[\frac{\partial F(\mathbf{u})}{\partial \mathbf{u}}\right] - 2F^T(\mathbf{u})\left(\mathbf{G}^{\mathrm{T}}\mathbf{G}\right)^{-1}\mathbf{u}, \end{aligned} \quad \text{(A-7)}$$

where $Tr[\cdot]$ denotes the trace of a matrix.

References

Aster RC, Thurber CH, Borchers B (2012) Parameter Estimation and Inverse Problems. Elsevier, London

Bauschke HH, Combettes PL (2011) Convex analysis and monotone operator theory in Hilbert spaces. Springer, London

Candès EJ, Romberg J, Tao T (2006) Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. IEEE Transactions on information theory 52(2):489-509

Claerbout JF, Muir F (1973) Robust modeling with erratic data. Geophysics 38(5):826–844

Eldar YC (2008) Generalized SURE for exponential families: Applications to regularization. IEEE Transactions on Signal Processing 57(2):471–481

Gholami A (2016) A fast automatic multichannel blind seismic inversion for high-resolution impedance recovery. Geophysics 81(5):V357–V364

Golub GH, Heath M, Wahba G (1979) Generalized cross-validation as a method for choosing a good ridge parameter. Technometrics 21(2):215–223

Gotoh J, Takeda A, Tono K (2018) DC formulations and algorithms for sparse optimization problems. Mathematical Programming 169(1):141–176

Hamid H, Pidlisecky A (2015) Multitrace impedance inversion with lateral constraints. Geophysics 80(6):M101--M111

Hansen PC (1992) Analysis of discrete ill-posed problems by means of the L-curve. SIAM Review 34(4):561–580

Huang G, Chen X, Chen Y (2021) P-P and Dynamic Time Warped P-SV Wave AVA Joint-Inversion with L1-2 Regularization. IEEE Transactions on Geoscience and Remote Sensing 59(7):5535–5548

Huang G, Li J, Luo C, Chen X (2017) Regularization parameter adaptive selection and its application in the Pre-stack AVO inversion. Exploration Geophysics 49(3):323–335

Infante-Pacheco VE, Montalvo-Arrieta JC, De León IN, Velasco-Tapia F (2020) Improvement of Shallow Seismic Characterization Using the Singular Value Decomposition (SVD) Method in Seismic Data Inversion: A Case Study of a Site in Northeast Mexico. Journal of Environmental and Engineering Geophysics 25(4):447–462

Kazemi N, Sacchi MD (2014) Sparse multichannel blind deconvolution. Geophysics 79(5):V143–V152

Lanning EN, Johnson DM (1983) Automated identification of rock boundaries: An application of the Walsh transform to geophysical well-log analysis. Geophysics 48(2):197–205

Li C, Zhang F (2017) Amplitude-versus-angle inversion based on the l1-norm-based likelihood function and the total variation regularization constraint. Geophysics 82(3):R173–R182

Liu X, Yin X (2015) Blocky inversion with total variation regularization and bounds constraint. SEG Technical Program Expanded Abstracts

Lou Y, Osher S, Xin J (2015) Computational aspects of constrained L1-L2 minimization for compressive sensing. Modelling, computation and optimization in information systems and management sciences. Springer, Cham

Ma J (2013) Three-dimensional irregular seismic data reconstruction via low-rank matrix completion. Geophysics 78(5):V181–V192

Maurya SP, Singh NP, Singh KH (2020) Seismic Inversion Methods: A Practical Approach. Springer, Cham

Mozayan DK, Gholami A, Siahkoohi HR (2018) Blocky inversion of multichannel elastic impedance

for elastic parameters. Journal of Applied Geophysics 151:166–174
Oldenburg DW, Scheuer T, Levy S (1983) Recovery of the acoustic impedance from reflection seismograms. Geophysics 48(10):1318–1337
Pan X, Zhang G, Song J et al (2017) Robust elastic impedance inversion using L1-norm misfit function and constraint regularization. Journal of Central South University 24(1):227–235
Pérez DO, Velis DR, Sacchi MD (2012) Inversion of Pre-stack seismic data using FISTA. Mecánica Computacional 31(20):3255–3263
Pérez DO, Velis DR, Sacchi MD (2013) High-resolution Pre-stack seismic inversion using a hybrid FISTA least-squares strategy. Geophysics 78(5): R185-R195
Robinson EA (1967) Predictive decomposition of time series with applications to seismic exploration. Geophysics 32(3):418-484
She B, Wang Y, Liang J et al (2019) Data-driven simultaneous seismic inversion of multiparameters via collaborative sparse representation. Geophysical Journal International 218(1):313–332
Stein CM (1981) Estimation of the mean of a multivariate normal distribution. The Annals of Statistics 1135–1151
Tao PD, An LTH (1998) A DC optimization algorithm for solving the trust-region subproblem. SIAM Journal on Optimization 8(2):476–505
Tarantola A (2005) Inverse Problem Theory. Elsevier, Amsterdam
Thore P (2015) Uncertainty in seismic inversion: What really matters? The Leading Edge 34(9):1000–1004
Wang L, Zhou H, Wang Y et al (2019) Three-parameter Pre-stack seismic inversion based on L1-2 minimization. Geophysics 84(5):R753–R766
Wang Y, Ma X, Zhou H et al (2018) L1−2 minimization for exact and stable seismic attenuation compensation. Geophysical Journal International 213(3):1629–1646
Wen B, Chen X, Pong TK (2018) A proximal difference-of-convex algorithm with extrapolation. Computational Optimization and Applications 69(2):297–324
Yin P, Lou Y, He Q et al (2015) Minimization of L1-2 for compressed sensing. SIAM Journal on Scientific Computing 37(1):A536–A563
Yin XY, Liu XJ, Zong ZY (2015) Pre-stack basis pursuit seismic inversion for brittleness of shale. Petroleum Science 12(4):618–627
Zhang F, Dai R, Liu H (2014) Seismic inversion based on L1-norm misfit function and total variation regularization. Journal of Applied Geophysics 109:111–118
Zhang R, Castagna J (2011) Seismic sparse-layer reflectivity inversion using basis pursuit decomposition. Geophysics 76(6):R147-R158.
Zhi LX, Chen SQ, Li XY (2016) Amplitude variation with angle inversion using the exact Zoeppritz equations - Theory and methodology. Geophysics 81(2):1–15

FIGURES

Figure 1

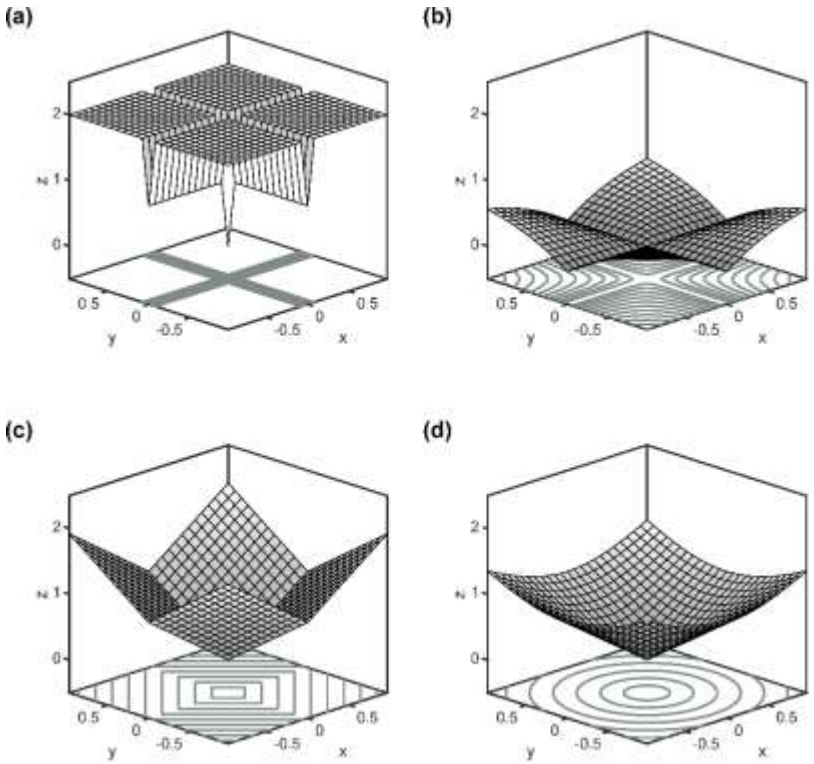

Figure 1: Comparison of 3D surface and 2D contour of the values for the different norms. **a** L_0-norm. **b** L_1-norm. **c** L_2-norm. **d** $L_{1\text{-}2}$-norm

Figure 2

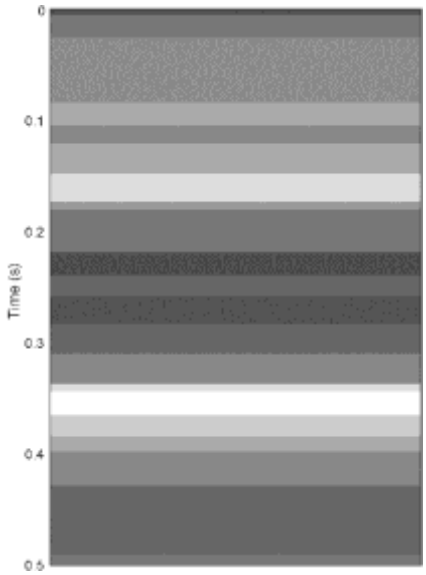

Figure 2: Strata model with different velocities and bulk density in 21 layers

Figure 3

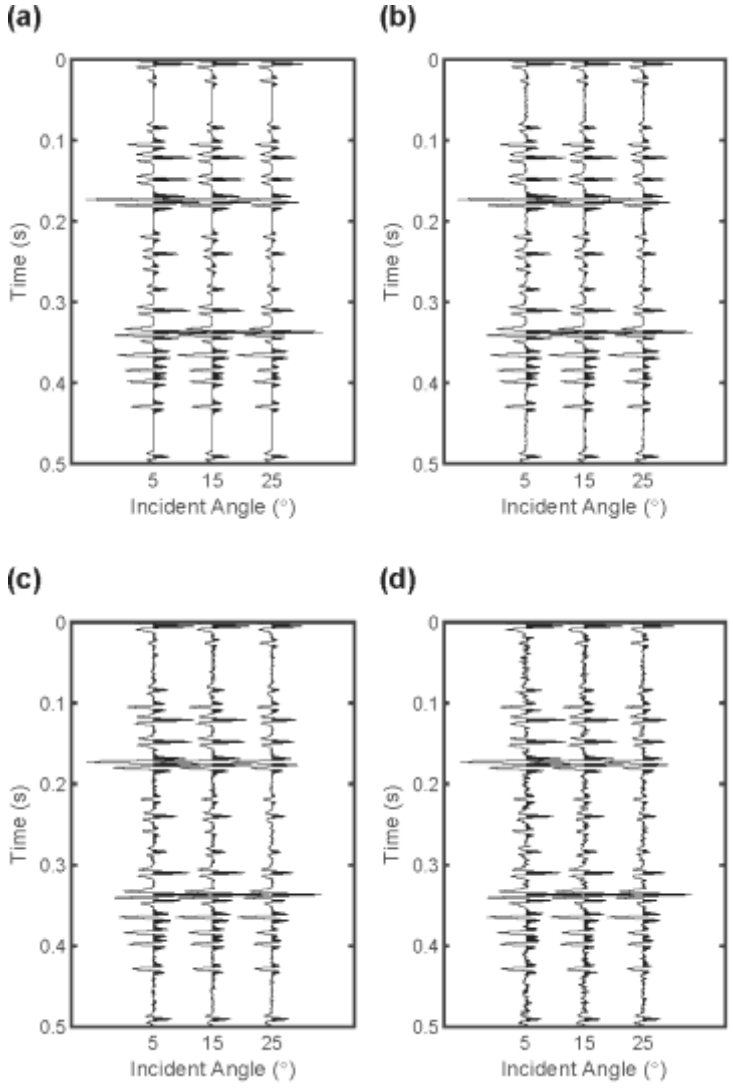

Figure 3: Synthetic pre-stack gathers with different signal-to-noise (SNR). Three incidence angles seismic traces are shown, for 5, 15, and 25 degrees. **a** No noise. **b** SNR=15. **c** SNR=10. **d** SNR=5

Figure 4

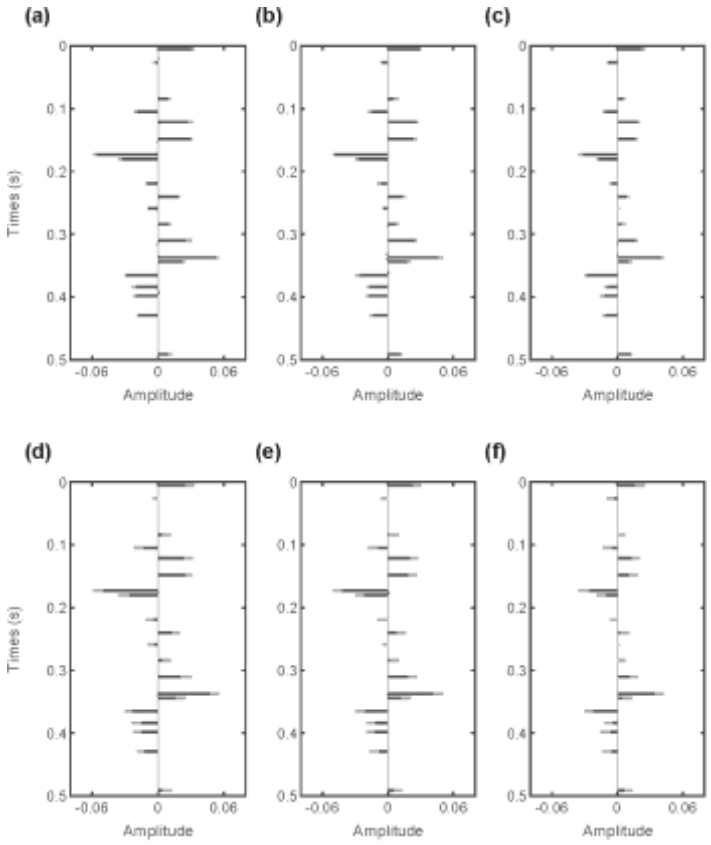

Figure 4: Comparison of inversion results with two sparse constraints for the pre-stack seismic inversion. **a-c** The inversion results with $L_{1\text{-}2}$-norm constraint. **d-f** The inversion results with L_1-norm constraint. The results are estimated with different incident angle: (**a, d**) 5 degrees, (**b, e**) 15 degrees, and (**c, f**) 25 degrees,. The gray and black bars correspond to the actual reflectivity series and the inversion results, respectively

Figure 5

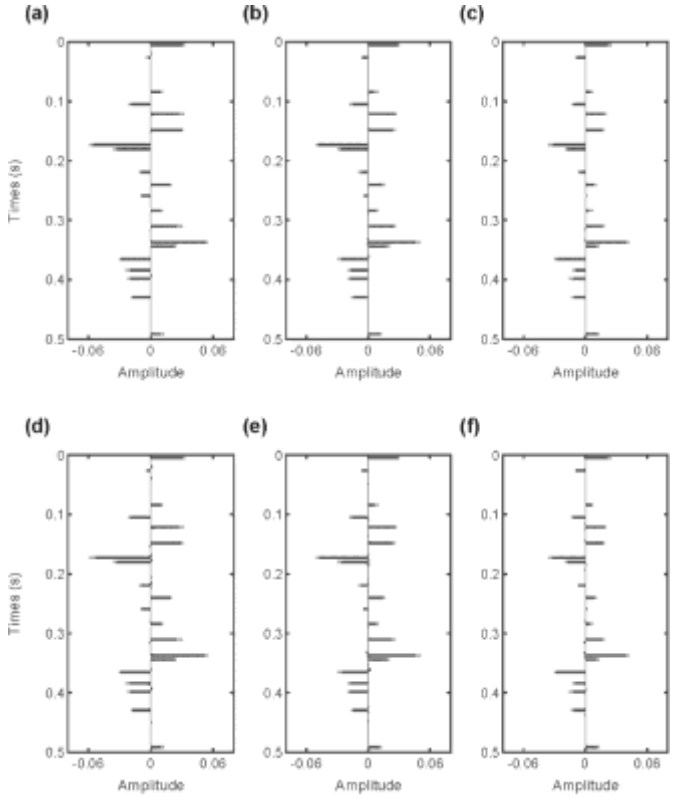

Figure 5: Comparison of inversion results with different solving methods by $L_{1\text{-}2}$-norm sparse constraint. **a-c** Inversion results by the new method. **d-f** Inversion results by DCA-ADMM. The results are estimated with different incident angles; (**a, d**) 5 degrees, (**b, e**) 15 degrees, and (**c, f**) 25 degrees. The gray and black bars correspond to the actual reflectivity series and the inversion results, respectively

Figure 6

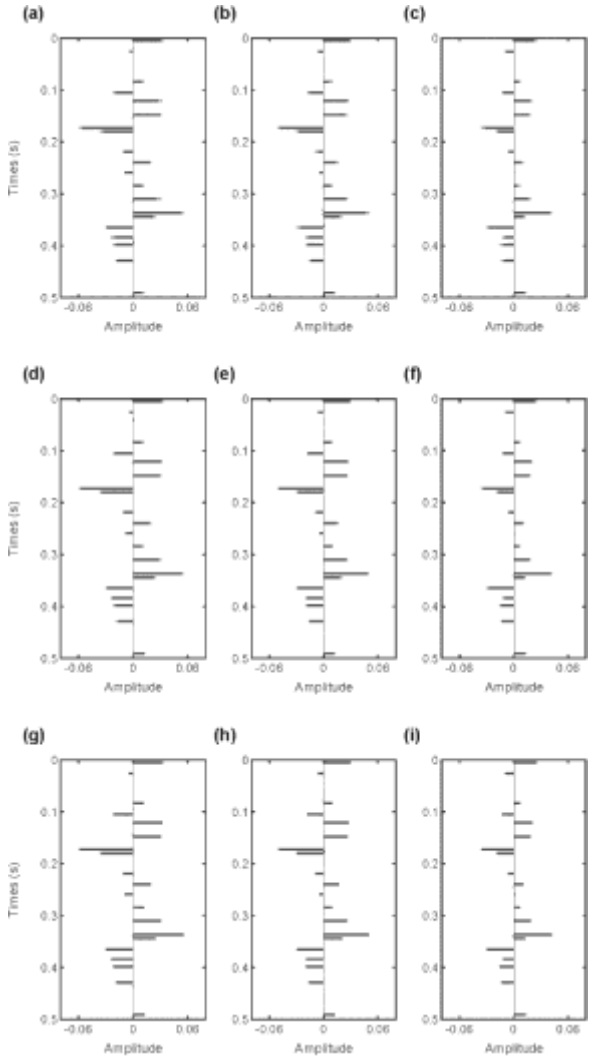

Figure 6: Comparison of inversion results with different SNR. **a-c** SNR = 5, **d-f** SNR = 10, **g-i** SNR = 15. The results are estimated with different incident angles: (**a, d, g**) 5 degrees, (**b, e, h**) 15 degrees, and (**c, f, i**) 25 degrees. The gray and black bars correspond to the actual reflectivity series and the inversion results, respectively

Figure 7

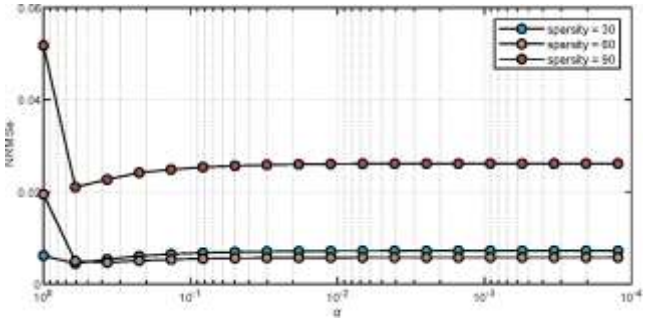

Figure 7: Variation of NRMSe with regularization parameter α in different sparsity models (sparsity = 30, 60, 90)

Figure 8

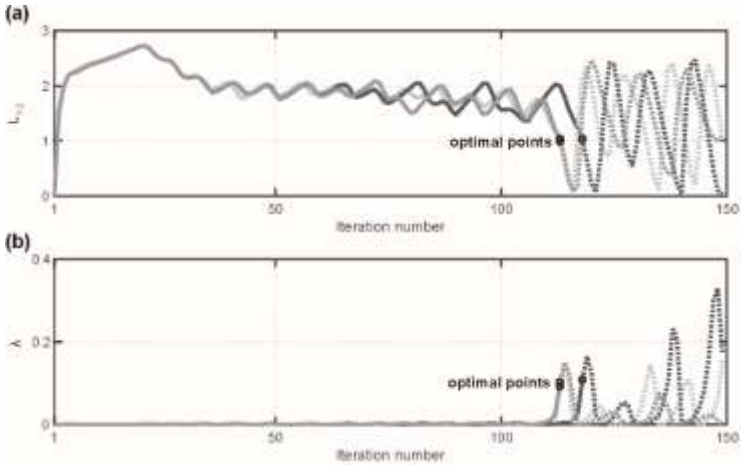

Figure 8: a Variation of $L_{1\text{-}2}$-norm in the iteration process. **b** Variation of regularization parameter λ in the iteration process. The solid lines represent the variation of properties when the convergence condition is satisfied, and the dashed line represents the variation of properties after the convergence condition is satisfied. The optimal points are marked in the case of different initial regularization parameters

Figure 9

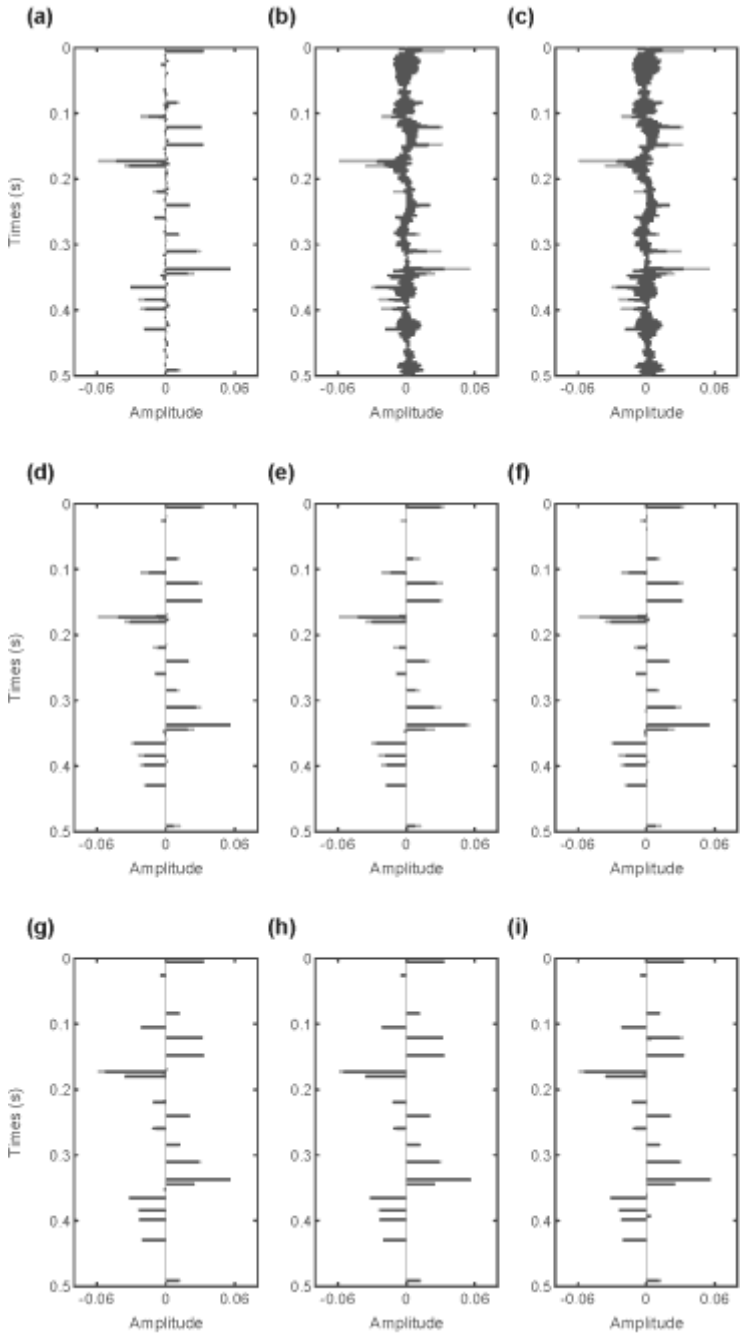

Figure 9: Inversion results of three strategies for regularization parameter. **a-c** The fixed regularization parameter strategy, **d-f** the adaptive method and **g-i** the hybrid strategy. The results are estimated with different initial regularization parameters: (**a, d, g**) 1×10^{-3}, (**b, e, h**) 1×10^{-5}, and (**c, f, i**) 1×10^{-7}. The gray and black bars correspond to the actual reflectivity series and the inversion results, respectively

Figure 10

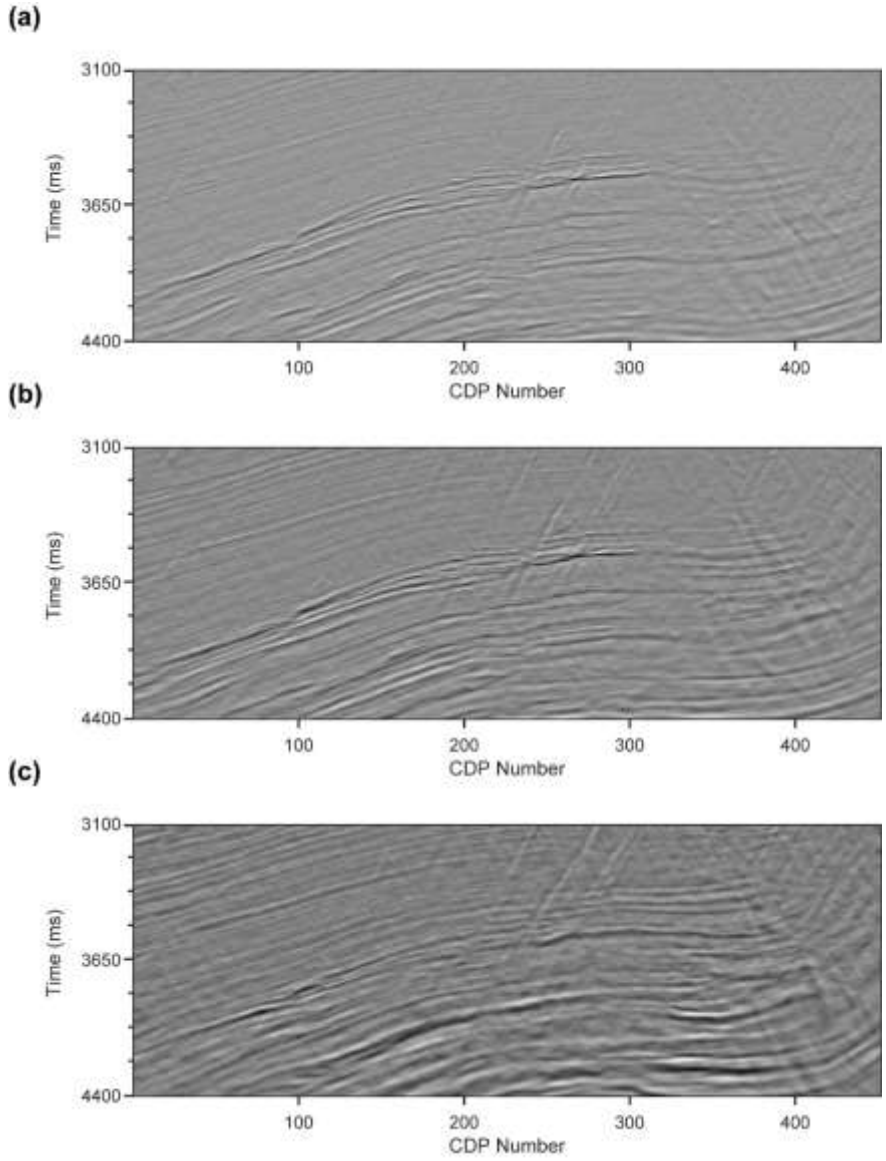

Figure 10: Three partial stacked seismic sections. **a** Near-angle stacked section, **b** Mid-angle stacked section, and **c** Far-angle stacked section

Figure 11

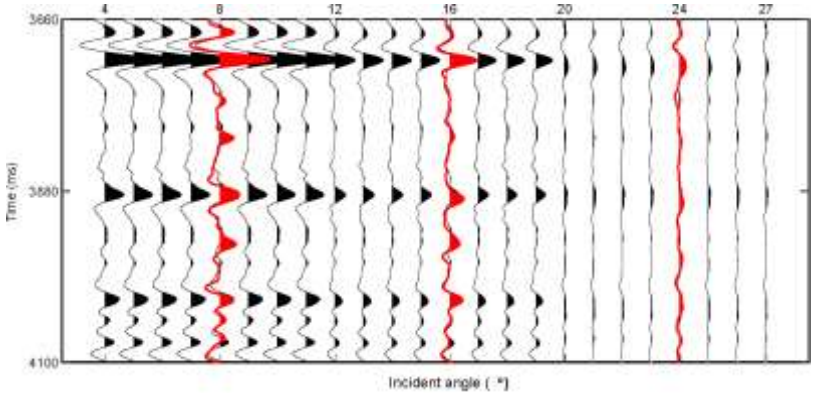

Figure 11: Comparison of field data (red) and synthetic data (black) at the well location in the time domain

Figure 12

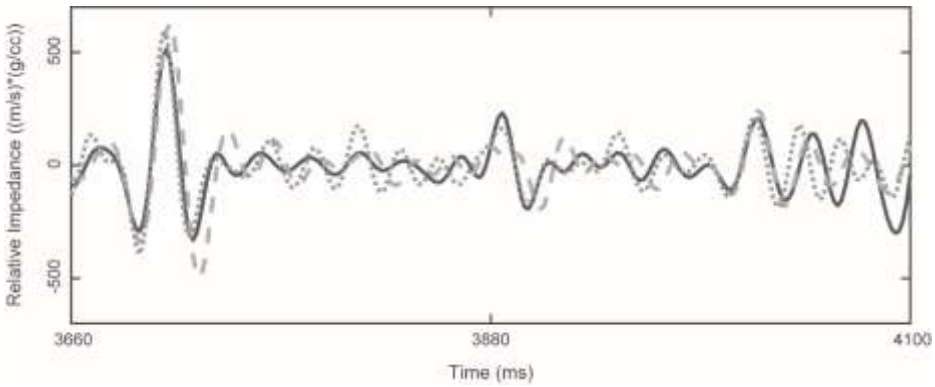

Figure 12: Comparison of the actual well-logging relative impedance trace and the inverted relative impedance traces by using two inversion methods in the time domain. The black line is relative impedance by filtering from the well-logging curves, the gray dashed line represents the result by the adaptive method, and the gray dotted line represents the result by the non-adaptive method

Figure 13

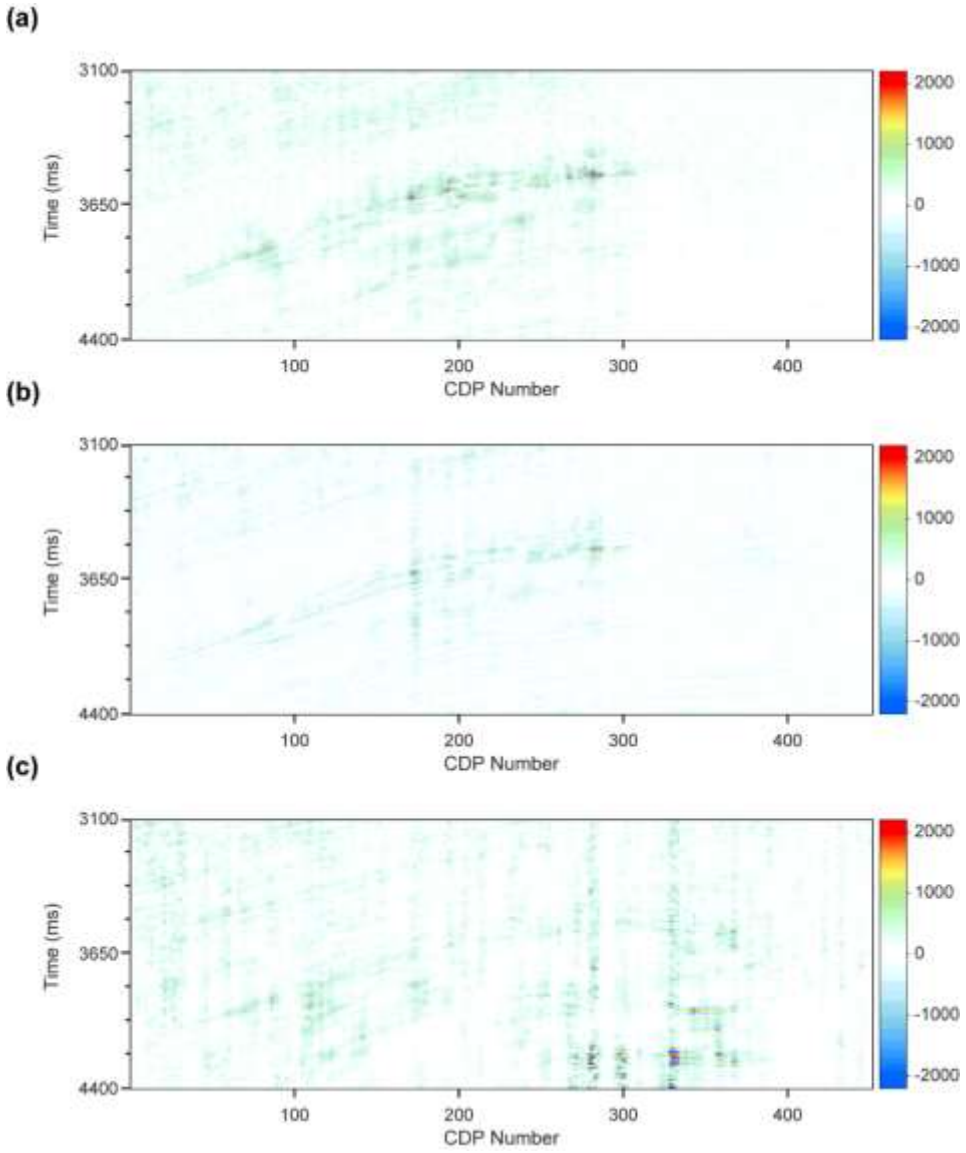

Figure 13: Inverted relative impedance section of **a** the near-angle stacked seismic section, **b** the mid-angle stacked seismic section and **c** the far-angle stacked seismic section, which corresponds to Fig. 10

TABLES

Table 1: Elastic parameters of the multilayer model in the synthetic example.

Layer	v_p (m/s)	v_s (m/s)	ρ (kg/m^3)
L1	4378.10	2330.70	2525.70
L2	4396.60	2359.70	2555.00
L3	4398.50	2395.60	2610.70
L4	4582.40	2462.00	2624.30
L5	4585.40	2543.70	2628.20
L6	4602.50	2547.40	2633.30
L7	4671.00	2593.30	2634.00
L8	4711.10	2631.00	2637.50
L9	4727.80	2646.60	2644.30
L10	4775.30	2659.30	2649.00
L11	4784.10	2686.90	2672.10
L12	4787.60	2743.40	2678.00
L13	4866.80	2747.20	2678.80
L14	4971.10	2814.70	2678.90
L15	5012.30	2855.20	2685.50
L16	5016.10	2893.90	2698.90
L17	5080.80	2909.70	2709.80
L18	5313.00	2985.00	2714.80
L19	5368.10	3133.30	2716.50
L20	5409.60	3169.90	2741.00
L21	5660.60	3239.80	2756.30