



Original Paper

Huber inversion-based reverse-time migration with de-primary imaging condition and curvelet-domain sparse constraint

Bo Wu ^{a, b}, Gang Yao ^{a, b, *}, Jing-Jie Cao ^{c, d}, Di Wu ^{a, e}, Xiang Li ^{a, b}, Neng-Chao Liu ^{a, e}

^a State Key Laboratory of Petroleum Resources and Prospecting, China University of Petroleum (Beijing), Beijing, 102249, China

^b Unconventional Petroleum Research Institute, China University of Petroleum (Beijing), Beijing, 102249, China

^c Key Laboratory of Intelligent Detection and Equipment for Underground Space of Beijing-Tianjin-Hebei Urban Agglomeration, Ministry of Natural Resources, Hebei GEO University, 050031, China

^d Hebei Key Laboratory of Strategic Critical Mineral Resources, Hebei GEO University, Shijiazhuang, Hebei, 050031, China

^e College of Geophysics, China University of Petroleum (Beijing), Beijing, 102249, China



ARTICLE INFO

Article history:

Received 7 June 2021

Accepted 25 January 2022

Available online 4 March 2022

Edited by Jie Hao

Keywords:

Least-squares reverse-time migration

Huber-norm

Sparse constraint

Curvelet transform

Iterative soft thresholding

ABSTRACT

Least-squares reverse-time migration (LSRTM) formulates reverse-time migration (RTM) in the least-squares inversion framework to obtain the optimal reflectivity image. It can generate images with more accurate amplitudes, higher resolution, and fewer artifacts than RTM. However, three problems still exist: (1) inversion can be dominated by strong events in the residual; (2) low-wavenumber artifacts in the gradient affect convergence speed and imaging results; (3) high-wavenumber noise is also amplified as iteration increases. To solve these three problems, we have improved LSRTM: firstly, we use Huber-norm as the objective function to emphasize the weak reflectors during the inversion; secondly, we adapt the de-primary imaging condition to remove the low-wavenumber artifacts above strong reflectors as well as the false high-wavenumber reflectors in the gradient; thirdly, we apply the L_1 -norm sparse constraint in the curvelet-domain as the regularization term to suppress the high-wavenumber migration noise. As the new inversion objective function contains the non-smooth L_1 -norm, we use a modified iterative soft thresholding (IST) method to update along the Polak-Ribière conjugate-gradient direction by using a preconditioned non-linear conjugate-gradient (PNCG) method. The numerical examples, especially the Sigsbee2A model, demonstrate that the Huber inversion-based RTM can generate high-quality images by mitigating migration artifacts and improving the contribution of weak reflection events.

© 2022 The Authors. Publishing services by Elsevier B.V. on behalf of KeAi Communications Co. Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

Seismic migration is a key technology in seismic data processing. It can image the subsurface structures by restoring the reflection events recorded on the surface to the true underground location. To obtain high-quality subsurface images, migration imaging technologies have evolved from ray-based Kirchhoff migration (Keho and Beydoun, 1988; Schneider, 1978) to Gaussian beam migration (Hill, 1990, 2001; Bai et al., 2019; Zhang et al., 2019), and from one-way wave-equation based migration (Claerbout, 1971; Claerbout and Doherty, 1972; Gazdag, 1978) to reverse-time migration (RTM), which is based on two-way wave equation

(Baysal et al., 1983; McMechan, 1983; Whitmore, 1983; Xu et al., 2011; Shi et al., 2019; Zhao et al., 2021). At present, RTM is one of the most popular migration methods for imaging models with complex structures. The reason is that RTM employs the two-way wave equation; therefore, it can handle steeply dipping structures, adapt to strong velocity variation, accurately migrate all frequencies and deal with complex waves, such as prism waves (Zhang et al., 2011).

However, instead of inverse operators, RTM still applies adjoint operators that only preserve the kinematics of the seismic image correctly, resulting in inaccurate amplitudes and degraded resolution (Etgen et al., 2009). Especially, in cases of significant aliasing, truncation, noise, and incomplete seismic data, the adjoint of the modeling operators approximate the inverse operators roughly (Claerbout and Karrenbach, 1992). In such cases, even with an

* Corresponding author.

E-mail address: yaogang@cup.edu.cn (G. Yao).

accurate migration velocity model, the image quality of RTM is still affected.

To further improve the quality, RTM is formulated in the least-squares inversion framework to produce the optimal reflectivity image, which is referred to as least-squares reverse-time migration (LSRTM) (Dai and Schuster, 2013; Liu et al., 2013; Zhang et al., 2015; Yao and Jakubowicz, 2016; Yang et al., 2019b; Hu et al., 2021; Wang et al., 2021). LSRTM can generate images with more balanced amplitudes, higher resolution, and fewer artifacts caused by missing data, truncation, noise, and operator mismatch than conventional RTM, thanks to the inverse operator applied by LSRTM (Yao and Wu, 2015; Wang and Wang, 2017; Yang et al., 2019a; Guo and Wang, 2020; Liu et al., 2020). LSRTM could also include anisotropy (e.g., Qu et al., 2017), attenuation (e.g., Qu and Li, 2019), elasticity (e.g., Fang et al., 2019; Gu et al., 2019), and multiples (e.g., Qu et al., 2021; Zhang et al., 2021) for honoring the physics more accurate.

However, conventional LSRTM utilizes L_2 -norm as the objective function. Consequently, the inversion is dominated by strong seismic events or other outliers in recorded data. For instance, the events originating from the boundary of salt bodies are much stronger than the events from the interfaces of sedimentary strata under the salt bodies. These strong events dominate the whole inversion process of LSRTM, resulting in an inadequate update to the region under the salt domes.

The second weakness of conventional LSRTM is that it still uses the zero-lag cross-correlation imaging condition, which generates the low-wavenumber migration artifacts, especially above strong reflectors, due to the two-way wave equations employed. In addition, the zero-lag cross-correlation imaging condition also generates false high-wavenumber reflectors in geologically complex areas (Fei et al., 2010). Although conventional LSRTM can suppress the low-wavenumber migration artifacts to some degree through the fitting process, this will affect the convergence speed of inversion in the early iterations.

The wavefield decomposition method is a popular choice to mitigate the low-wavenumber migration artifacts. Liu et al. (2011) proposed an efficient wavefield decomposition method to suppress low-wavenumber migration artifacts in the RTM images. Fei et al. (2015) proposed the de-primary imaging condition to further improve the effectiveness of imaging. Except for suppressing low-wavenumber migration artifacts, this imaging condition can also eliminate false high-wavenumber reflectors. Kim et al. (2019) used the de-primary imaging condition in LSRTM and utilized a preconditioned linear conjugate-gradient method (PLCG) for inversion. However, the migration operator is no longer the adjoint of the modeling operator as the de-primary imaging condition changes the gradient of LSRTM. This method slows down the convergence of inversion. To mitigate this problem, they used the de-primary imaging condition only at early iterations, and then switched to the zero-lag cross-correlation imaging. However, the later iterations with the zero-lag cross-correlation imaging condition still introduce low-wavenumber artifacts again.

Thirdly, the inversion process of conventional LSRTM also produces high-wavenumber noise to overfit the record because of the ill-conditioning nature of geophysics problems as well as the modeling engine that only simulates the record partially (Trad, 2020). To solve this problem, L_1 -norm sparse constraints are incorporated into LSRTM to penalize the high-wavenumber migration noise and reduce the sidelobes around the reflectors (Lin and Lian, 2015; Wu et al., 2016; Dutta, 2017; Li et al., 2017). This method suppresses the artifacts and also the weak reflectors. Li et al. (2020) further implemented the L_1 -norm sparse constraint in the wavelet domain. It can perform multi-scale and angle analysis, so it can protect weak signals better while removing noise.

In this paper, we propose a new strategy to achieve LSRTM for

solving these problems illustrated above. Firstly, we employ a Huber norm (Huber, 1964) as the objective function of the inversion. Thus, the inversion acts as L_1 -norm inversion for large residuals but as L_2 -norm inversion for small residuals. Consequently, it is strongly convex when close to the minimum and less steep for large residuals. Thus, this can mitigate the problem: large residuals dominate the inversion of LSRTM. Since the inversion is not based on L_2 -norm only, we call it Huber inversion-based RTM instead of LSRTM. Secondly, the de-primary imaging condition is employed to formulate the gradient instead of the zero-lag cross-correlation imaging condition, which causes both low-wavenumber artifacts and high-wavenumber false reflectors. Finally, we use the L_1 -norm sparse constraint in the curvelet-domain as the regularization term to remove high-wavenumber noise. The reason is that the curvelet transform can sparsely represent images at multiple scales and angles, and has a stronger ability to express edge information than other transforms, such as wavelet transform (Candès and Guo, 2002). As the proposed Huber inversion-based RTM contains a non-smooth L_1 -norm, we utilize an improved iterative soft thresholding (IST), which is updated along the Polak-Ribière conjugate-gradient direction by using a preconditioned non-linear conjugate-gradient (PNCG) method. The numerical examples, especially the Sigsbee2A model, demonstrate that our proposed method can improve the image quality and show more structural details in poorly illuminated areas.

In the following chapters, firstly, we introduce the theory of Huber inversion-based RTM with the curvelet-domain constraint. Secondly, numerical examples are presented to prove the effectiveness of the proposed method. Before the conclusion, we discuss the convergence of all these RTM methods, the effectiveness of the curvelet-domain sparse constraint for noise suppression with Huber inversion-based RTM, and the influence of velocity accuracy on inversion-based RTM methods.

2. Method

2.1. Huber inversion-based reverse-time migration with curvelet-domain constraint

Least-square reverse-time migration (LSRTM) generally uses the L_2 -norm objective function.

$$\varphi_{L_2} = \frac{1}{2} \|\mathbf{d} - \mathbf{d}_{\text{obs}}\|_2^2 \quad (1)$$

where \mathbf{d} denotes the predicted data and \mathbf{d}_{obs} represents the observed data. Note that a boldface letter in an equation represents the vector form of a variable, whereas a normal letter is for its scalar form in the whole paper. The L_2 -norm objective function is particularly efficient for reflection events that have similar amplitudes from sedimentary rocks. However, if models include large contrast in geological bodies, e.g. salt bodies, then an objective function which emphasizes more on weak events is expected. Besides, a constraint should be incorporated into the migration inversion system to suppress high-wavenumber noise generated by the inversion process. By considering the two points, we combine the Huber norm and the curvelet-domain constraint as the objective function

$$\varphi = \varphi_H + \lambda C^{-1} \|\mathbf{Cm}\|_1 \quad (2)$$

where φ_H denotes the Huber norm, λ is the regularization (trade-off) parameter, \mathbf{m} represents the reflectivity model, $\|\cdot\|_1$ denotes L_1 -norm, $C(\cdot)$ and $C^{-1}(\cdot)$ stand for the forward and inverse curvelet transforms, respectively. At present, there are two most

commonly used fast discrete curvelet transforms (Candès and Donoho, 2004): USFFT (unequally spaced fast Fourier transform) algorithm and Wrapping-based transform algorithm. In this paper, we use the USFFT algorithm, the code of which can be downloaded from the website (<http://www.curvelet.org/>), to realize the discrete curvelet transform.

In this paper, we choose a pseudo-Huber loss function (Charbonnier et al., 1997) to replace the L_2 -norm objective function

$$\varphi_H = \delta^2 \sum_i \left(\sqrt{1 + \left((d^i - d_{\text{obs}}^i) / \delta \right)^2} - 1 \right) \quad (3)$$

where d^i and d_{obs}^i represents the i th element of \mathbf{d} and \mathbf{d}_{obs} , respectively, and δ is the threshold of Huber-norm. An advantage of the pseudo-Huber norm over the pice-wise Huber norm is that its gradient is continuous for all values, which is convenient for implementing gradient-based inversion methods. We choose half the maximum value of the residual, i.e. data misfit $d^i - d_{\text{obs}}^i$, as δ in each iteration. Therefore, when the residual is large, the Huber norm acts like L_1 -norm; otherwise, it acts like L_2 -norm. Since the residual shrinks as inversion progresses, δ decreases during inversion. Reducing δ as inversion progresses is an effective strategy in optimization called “cooling”. In the early iterations, the value of δ is relatively large, therefore, strong events converge quickly. As inversion progresses, reducing δ increases the contribution of weak signals in the objective function and avoids strong events continuing to dominate the updating direction of the model. Note that δ in the pseudo-Huber norm is slightly larger than the transition point between the two norms. This is illustrated by Fig. 1.

The mapping relationship between predicted data and reflectivity is essential for minimizing the objective function. Generally, there are two simulation methods for predicted data: one is based on Born approximation, and the other is based on Kirchhoff approximation (Jaramillo and Bleistein, 1999; Bleistein et al., 2005). In this paper, we used the Kirchhoff approximation:

$$d(\mathbf{x}_r, \omega; \mathbf{x}_s) = i\omega \sum_{\mathbf{x}} G_r(\mathbf{x}_r | \mathbf{x}, \omega) m(\mathbf{x}) G_s(\mathbf{x} | \mathbf{x}_s, \omega) s(\omega) \quad (4)$$

where s is the source wavelet, $m(\mathbf{x})$ represents the stacked reflectivity image, $G_s(\mathbf{x} | \mathbf{x}_s, \omega)$ is the Green's function from the source point \mathbf{x}_s to the reflection point \mathbf{x} , $G_r(\mathbf{x}_r | \mathbf{x}, \omega)$ is the Green's function from the reflection point \mathbf{x} to the receiver point \mathbf{x}_r , and $i\omega$ gives a 90° phase shift to the simulated data, which is essential for zero-phase imaging. The zero-phase imaging with Kirchhoff approximation matches the images of conventional migration methods, such as Kirchhoff migration. By contrast, imaging with Born approximation gives 90° phase rotation compared with the images produced with conventional migration methods. A more detailed

derivation and implementation of the forward model operator in Eq. (4) can be found in Yao et al. (2018).

Eqs. (2)-(4) constitute the inversion system used in this paper. To minimize the function in Eq. (2), we employ the iterative soft thresholding (IST) method, which is an effective approach to minimizing the mixed norm problems. IST is a special type of the proximal method, which solves a general optimization problem

$$\min \varphi(\mathbf{m}) + h(\mathbf{m}) \quad (5)$$

iteratively with

$$\mathbf{m}^i = \text{prox}_{\alpha_i h} \left(\mathbf{m}^{i-1} - \alpha_i \nabla \varphi \left(\mathbf{m}^{i-1} \right) \right) \quad (6)$$

where $\varphi(\mathbf{m})$ is a smooth function, e.g. φ_H , $h(\mathbf{m})$ is a convex function but not differentiable everywhere, e.g. $\lambda C^{-1} \|\mathbf{Cm}\|_1$, α represents the step-length, i indicates the iteration number, and $\text{prox}(\cdot)$ is called the proximity operator and is the soft thresholding operator for IST (Neal and Stephen, 2014). Originally, the proximal method updates the model along the gradient direction, i.e. $\frac{\partial \varphi_H}{\partial m(\mathbf{x})}$. We found a fast convergence way for IST is to update the model along the Polak-Ribière conjugate-gradient direction preconditioned by the stack of all shots source wavefield energy.

2.2. Gradient approximated by de-primary imaging condition

Substituting Eq. (4) into Eq. (3) can yield the gradient of the objective function

$$\begin{aligned} \frac{\partial \varphi_H}{\partial m(\mathbf{x})} &= \sum_t \left(\frac{\partial d(\mathbf{x}_r, t)}{\partial m(\mathbf{x})} \right)^\dagger \left(\frac{d(\mathbf{x}_r, t) - d_{\text{obs}}(\mathbf{x}_r, t)}{\sqrt{1 + ((d(\mathbf{x}_r, t) - d_{\text{obs}}(\mathbf{x}_r, t)) / \delta)^2}} \right) \\ &= \sum_{\mathbf{x}_s} \sum_{\mathbf{x}_r} \sum_{\omega} -i\omega (G_s(\mathbf{x} | \mathbf{x}_s, \omega) s(\mathbf{x}_s, \omega))^\dagger G_r(\mathbf{x}_r | \mathbf{x}, \omega)^\dagger \Delta d(\mathbf{x}_r, \omega) \end{aligned} \quad (7)$$

where $\Delta d(\mathbf{x}_r, \omega)$ is the frequency domain counterpart of $\frac{d(\mathbf{x}_r, t) - d_{\text{obs}}(\mathbf{x}_r, t)}{\sqrt{1 + ((d(\mathbf{x}_r, t) - d_{\text{obs}}(\mathbf{x}_r, t)) / \delta)^2}}$, representing one element of the residual vector of Huber-norm $\Delta d(\mathbf{x}_r, t)$, $G_s(\mathbf{x} | \mathbf{x}_s, \omega) s(\mathbf{x}_s, \omega)$ represents the forward wavefield that is generated by propagating the seismic source from the source point \mathbf{x}_s to the reflection point \mathbf{x} , $G_r^\dagger(\mathbf{x}_r | \mathbf{x}, \omega) \Delta d(\mathbf{x}_r, \omega)$ represents the backward wavefield that is generated by propagating the residual from the reflection point \mathbf{x}_r to the receiving point \mathbf{x} , and \dagger denotes conjugate transpose. The physical meaning of Eq. (7) is the zero-lag cross-correlation of the first-time derivative of the forward wavefield and the backward wavefield of the residual. This is coincident with the zero-lag cross-

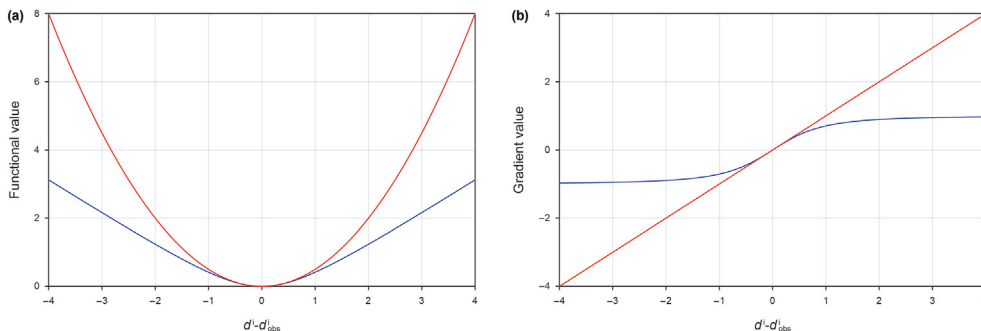


Fig. 1. A sketch of the pseudo-Huber loss function (blue, $\delta = 1$) and the L_2 -norm loss function $\frac{1}{2} \|\mathbf{d} - \mathbf{d}_{\text{obs}}\|_2^2$ (red). (a) The functional value and (b) their gradients.

$$\frac{\partial \phi_H}{\partial \mathbf{m}(\mathbf{x})} \approx \sum_{\mathbf{x}_s} \sum_{t=0}^{T_{\max}} \left(\frac{\partial S}{\partial t} \left(\sum_{\mathbf{x}_r} G(\mathbf{x}|\mathbf{x}_r, t) * \Delta d(\mathbf{x}_r, t) \right) - \mathcal{H}_z \left(\frac{\partial S}{\partial t} \right) \mathcal{H}_z \left(\sum_{\mathbf{x}_r} G(\mathbf{x}|\mathbf{x}_r, t) * \mathcal{H}_t(\Delta d(\mathbf{x}_r, t)) \right) \right) \quad (11)$$

correction imaging condition widely used in reverse-time migration (RTM). For a common shot gather, it can be expressed as

$$I(\mathbf{x}) = \sum_{\mathbf{x}_s} \sum_{t=0}^{T_{\max}} S(\mathbf{x}, t; \mathbf{x}_s) R(\mathbf{x}, t; \mathbf{x}_s) \quad (8)$$

where $I(\mathbf{x})$ represents the RTM image, $S(\mathbf{x}, t; \mathbf{x}_s)$ indicates the source wavefield (or the forward wavefield), and $R(\mathbf{x}, t; \mathbf{x}_s)$ stands for the receiver wavefield (or the backward wavefield). Eq. (8) can be expressed by wavefield decomposition as follows (Liu et al., 2011)

$$I(\mathbf{x}) = \sum_{\mathbf{x}_s} \sum_{t=0}^{T_{\max}} (S_d R_d + S_u R_u + S_d R_u + S_u R_d) \quad (9)$$

$$= \sum_{\mathbf{x}_s} \left(\sum_{t=0}^{T_{\max}} S_d R_d + \sum_{t=0}^{T_{\max}} S_u R_u + \sum_{t=0}^{T_{\max}} S_d R_u + \sum_{t=0}^{T_{\max}} S_u R_d \right)$$

where the subscripts u and d indicate the up-going and down-going wavefields. $S_d R_d$ produces the primary reflectors, $S_u R_u$ usually generates false reflectors in RTM, $S_d R_u$ and $S_u R_d$ produce low-wavenumber migration artifacts. The process of producing the false reflectors is shown in Fig. 2. Therefore, the de-primary imaging condition only retains the first term of Eq. (9). This term can be expressed by using Hilbert Transform (Fei et al., 2015)

where $G(\mathbf{x}|\mathbf{x}_r, t)$ represents the Green's function from the record

$$I(\mathbf{x}) = \sum_{\mathbf{x}_s} \sum_{t=0}^{T_{\max}} S_d R_d = \sum_{\mathbf{x}_s} \sum_{t=0}^{T_{\max}} \left\{ \begin{array}{l} SR - \mathcal{H}_z(S) \mathcal{H}_z(R) - S \mathcal{H}_z \left[\sum_{\mathbf{x}_r} G(\mathbf{x}|\mathbf{x}_r, t) * \mathcal{H}_t(d_{\text{obs}}(\mathbf{x}_r, t)) \right] \\ - \mathcal{H}_z(S) \mathcal{H}_z \left[\sum_{\mathbf{x}_r} G(\mathbf{x}|\mathbf{x}_r, t) * \mathcal{H}_t(d_{\text{obs}}(\mathbf{x}_r, t)) \right] \end{array} \right\} \quad (10)$$

point to the reflection point, \mathcal{H}_z and \mathcal{H}_t represent Hilbert Transform along with the space and time direction, respectively, and $*$ denotes convolution.

In the same way, in order to suppress the low-wavenumber artifacts and high-wavenumber false reflectors from the gradient of Huber inversion-based RTM, we apply de-primary imaging condition by decomposing the first-order time derivative of the forward wavefield and the backward wavefield into the up-going and down-going wavefields through the Hilbert Transform, and then only keeping the first term. Therefore, the gradient of Huber inversion-based RTM with the de-primary imaging condition can be written approximately as

imaging condition in Eq. (11) is not the exact gradient of the objective function, therefore, PLCG methods are not the best choice for inversion. Its convergence will be affected. This situation happens not only in the Huber inversion-based RTM but also in the traditional LSRTM with the de-primary imaging condition. To solve this problem, the PNCG method is a better choice to ensure the convergence of the inversion.

By combining all the aspects presented above, we minimize the function in Eq. (2) by using IST with the Polak-Ribière conjugate-gradient direction preconditioned by the stack of all shots source wavefield energy. The detailed algorithm is listed in Algorithm 1. Generally, the optimal value of λ is refined by trial and error. The preconditioner M is defined as

$$M(\mathbf{x}) = \sum_{\mathbf{x}_s} \sum_{t=0}^{T_{\max}} S(\mathbf{x}, t; \mathbf{x}_s) \cdot S(\mathbf{x}, t; \mathbf{x}_s) \quad (12)$$

Algorithm 1. The preconditioned non-linear conjugate-gradient method with IST for Huber inversion-based RTM

Input: initial $\mathbf{m}, \lambda, n_{\max}$;

$n = 0, \mathbf{g} = -\frac{\partial \phi_{H(i)}}{\partial \mathbf{m}}$

Compute the preconditioner: M

$\mathbf{r} = M^{-1} \mathbf{g}, \mathbf{s} = \mathbf{r}, \delta_{\text{new}} = \mathbf{g}^T \mathbf{s}$

While $i < n_{\max}$ do

$\alpha = \arg \min_{\alpha} \phi_{H(n)}(\mathbf{m} + \alpha \delta \mathbf{m})$, where α is the step length

computed by line search and $\delta \mathbf{m}$ is a small perturbation along \mathbf{s} .

$\widehat{\mathbf{m}} = C(\mathbf{m} + \alpha \delta \mathbf{m})$

$\mathbf{m} = C^{-1}(T_{\lambda}(\widehat{\mathbf{m}}))$, where $T_{\lambda}(\widehat{\mathbf{m}}_i) = \text{sgn}(\widehat{\mathbf{m}}_i) \max(0, |\widehat{\mathbf{m}}_i| - \lambda)$

$\mathbf{g} = -\frac{\partial \phi_{H(i)}}{\partial \mathbf{m}}$

$\delta_{\text{old}} = \delta_{\text{new}}, \delta_{\text{mid}} = \mathbf{g}^T \mathbf{r}, \mathbf{r} = M^{-1} \mathbf{g}, \delta_{\text{new}} = \mathbf{g}^T \mathbf{r}$

$\beta = \max((\delta_{\text{new}} - \delta_{\text{mid}}) / \delta_{\text{old}}, 0), \mathbf{s} = \mathbf{r} + \beta \mathbf{s}$

$i = i + 1$

end

It can be seen from Eq. (7) that the gradient with the de-primary

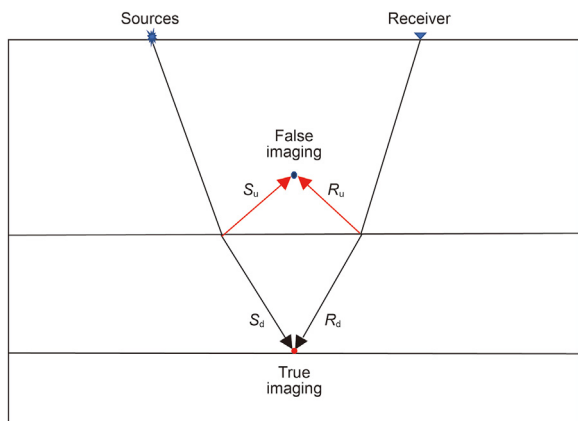


Fig. 2. Schematic illustration of the de-primary imaging condition in a three-layer model. The blue dot represents the false image produced by the up-going wavefields, while the red dot denotes the true image created by the down-going wavefields. Both of them have the same travel time.

3. Numerical examples

3.1. Marmousi2 model

In the first example, the Marmousi2 model is used to demonstrate the effectiveness of the Huber inversion-based RTM. The seismic data set is generated from the Marmousi2 model with fine stratigraphy (Fig. 3a) by solving the two-way acoustic wave equation. The data set includes 176 shots with a shot spacing of 50 m. A split-spread geometry is applied. The receiver spacing is 25 m. The minimum and maximum offsets are 0 m and 2600 m, respectively. The source wavelet is a 30 Hz Ricker wavelet. The total recording duration is 3.2 s with a sampling interval of 0.4 ms. The sources and receivers are positioned at 20 m below the surface. Two typical shot gathers are shown in Fig. 4, in which the direct arrivals are removed. A slightly smoothed velocity model shown in Fig. 3b is used for migration. In industrial applications, the migration velocity is built mainly with ray-based reflection tomography (Murphy and Gray, 1999) or reflection waveform inversion (Yao et al., 2020). Note that we deliberately smoothed the velocity model slightly to demonstrate the inversion methods' effectiveness of suppressing low-wavenumber artifacts.

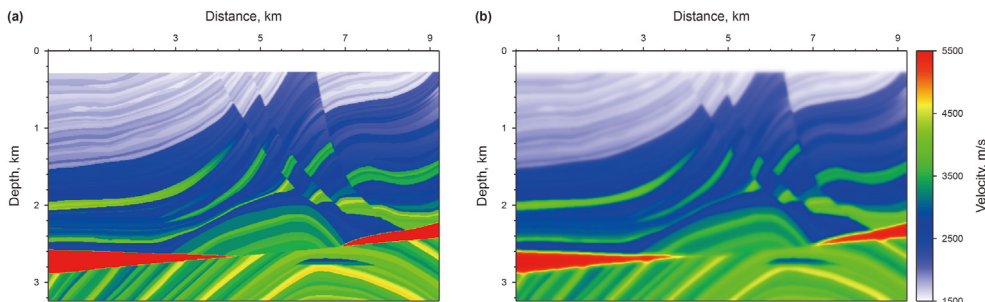


Fig. 3. The Marmousi2 model: (a) the stratigraphic velocity model; (b) a smoothed version of (a). The smooth is achieved using the Seismic Unix command “smooth2” with the parameter $r1 = r2 = 10$. The spatial sampling interval in the horizontal and vertical directions is 6.25 m and 4 m, respectively.

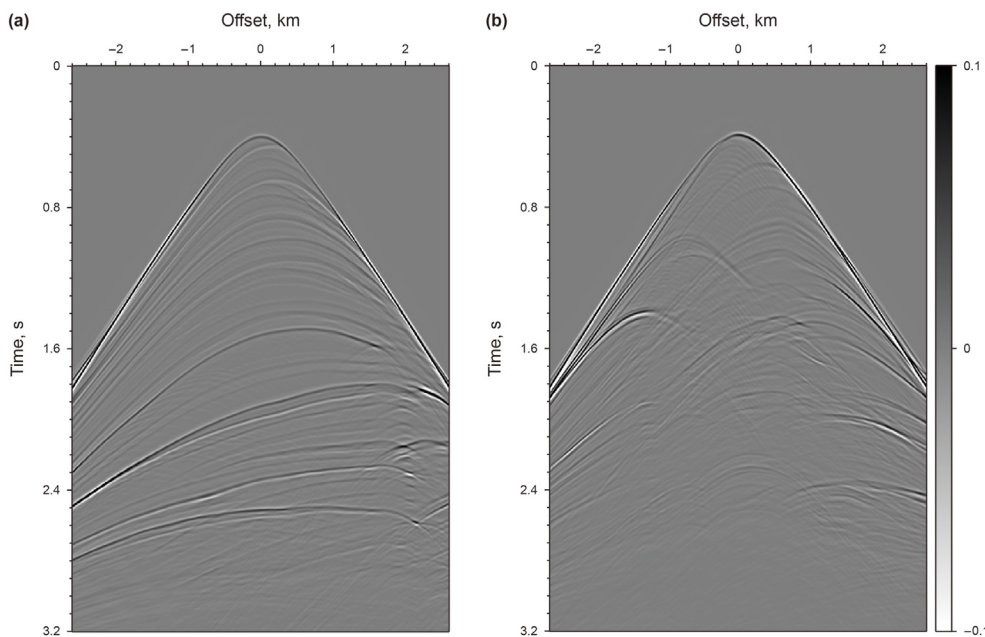


Fig. 4. Two shot gathers: the source is located at the distance of (a) 2.675 km and (b) 5.675 km. The direct arrivals are removed.

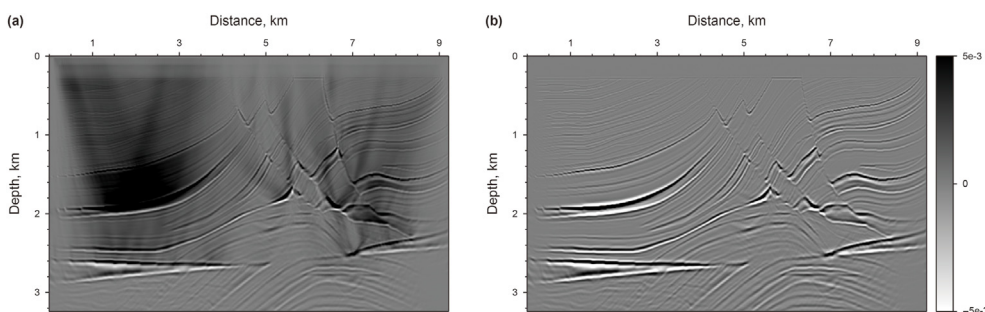


Fig. 5. RTM images: (a) the zero-lag cross-correlation imaging condition and (b) the de-primary imaging condition.

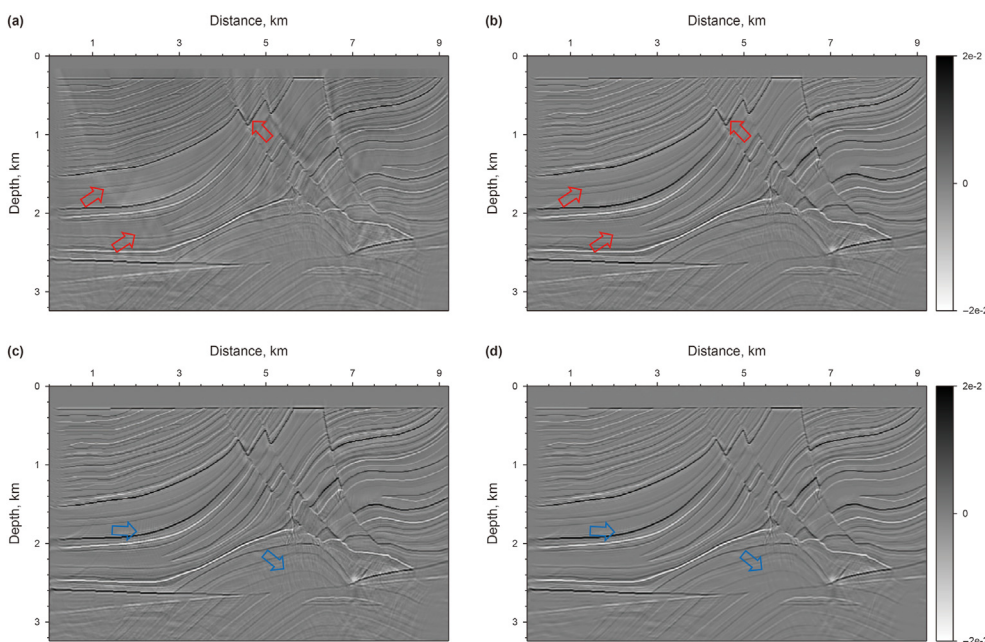


Fig. 6. Migration images. (a) LSRTM with the zero-lag cross-correlation imaging condition; (b) LSRTM with the de-primary imaging condition; (c) Huber inversion-based RTM with the de-primary imaging condition; (d) Huber inversion-based RTM with the de-primary imaging condition and the curvelet-domain sparse constraint. The initial model is shown in Fig. 5b. All the inversions contain 20 iterations. The red arrows indicate the low-wavenumber artifacts, whereas the blue arrows point out the high-wavenumber noise.

Fig. 5a and b shows the RTM images with the zero-lag cross-correlation imaging condition and the de-primary imaging condition, respectively. By comparison, low-wavenumber migration artifacts prevail because of the zero-lag cross-correlation imaging condition in Fig. 5a. These artifacts cover up some real reflectors. In Fig. 5b, the de-primary imaging condition mitigates the low-wavenumber migration artifacts effectively. To further improve the imaging quality, we carried out LSRTM using the result of RTM with the de-primary imaging condition (Fig. 5b) as the initial model. Figs. 6a and b depict the results of LSRTM with the zero-lag cross-correlation imaging condition and the de-primary imaging condition, respectively, after 20 iterations. As can be seen, although the initial model does not contain low-wavenumber migration artifacts, the result of LSRTM with the zero-lag cross-correlation imaging condition (Fig. 6a) introduces low-wavenumber artifacts back, which are indicated by the red arrows in Fig. 6. By contrast, LSRTM with the de-primary imaging condition removes the low-wavenumber migration artifacts completely. Furthermore, the reflectors are focused better by using the de-primary imaging condition than the zero-lag cross-correlation imaging condition. The reason is that removing the low-wavenumber artifacts helps inversion to focus on updating reflectors. Fig. 6c shows the imaging result of Huber inversion-based RTM with the de-primary imaging

condition. As the sedimentary strata dominates the Marmousi2 model, the events in the record share similar amplitudes. Thus, the imaging result of Huber inversion-based RTM is similar to LSRTM. However, we can still see high-wavenumber artifacts, pointed by the blue arrows in Fig. 6c. The high-wavenumber artifacts are suppressed largely by incorporating the curvelet-domain sparse constraint, which is shown in Fig. 6d.

3.2. Sigsbee2A model

In the second example, the proposed method is applied to a synthetic data set from the Sigsbee2A model, which is a very challenging model for migration due to a huge salt body in the model. A seismic data set is generated by solving the two-way acoustic wave equation with the fine stratigraphic velocity model (Fig. 7a). The data set includes 160 shots, which are excited with a shot spacing of 137.16 m. A one-side geometry is used with a receiver spacing of 45.72 m. The minimum and maximum offsets are 0 m and 3989 m, respectively. The source wavelet is a 12 Hz Ricker wavelet. The total recording duration is 12 s with a sampling interval of 1 ms. Two typical shot gathers are shown in Fig. 8. A smoothed version of the true velocity model shown in Fig. 7b is used for migration.

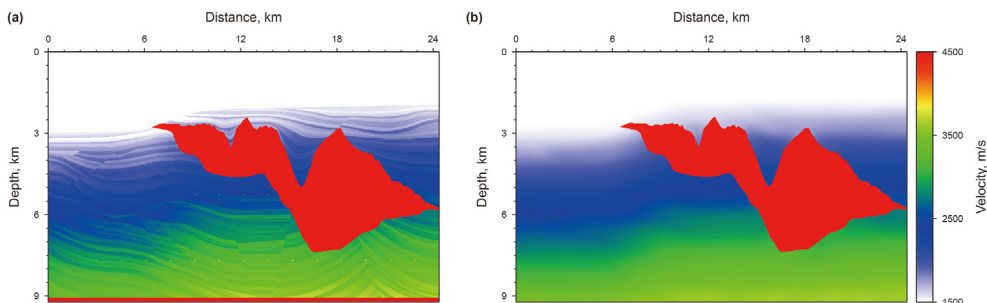


Fig. 7. The Sigsbee2A models. (a) The stratigraphic velocity model, (b) a smoothed version of (a) using a 20-point cosine-square window for migration. The spatial sampling interval along the horizontal and vertical directions is 11.43 m and 7.62 m, respectively.

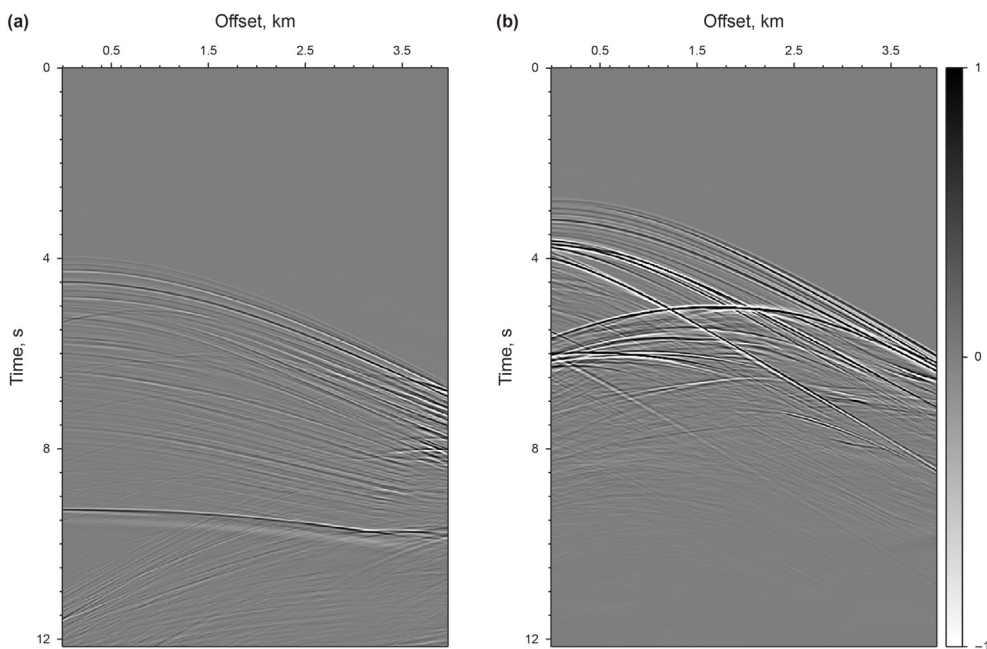


Fig. 8. Two shot gathers. The source is located at a distance of (a) 0.274 km and (b) 13.99 km. The direct arrivals are removed.

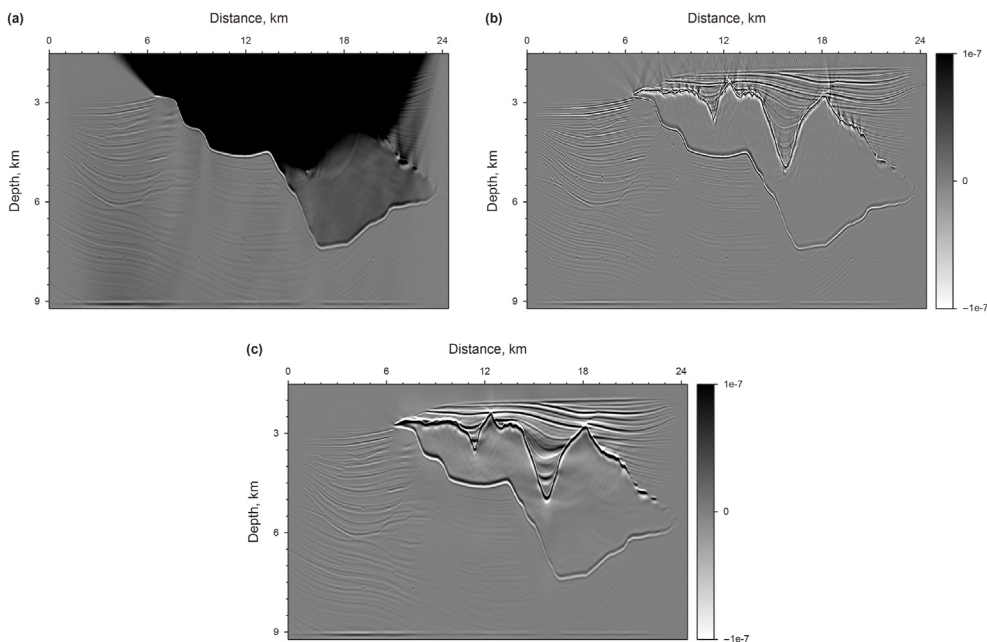


Fig. 9. RTM images. (a) Using the zero-lag cross-correlation imaging condition; (b) Laplacian filtered result of the image in (a); (c) using the de-primary imaging condition.

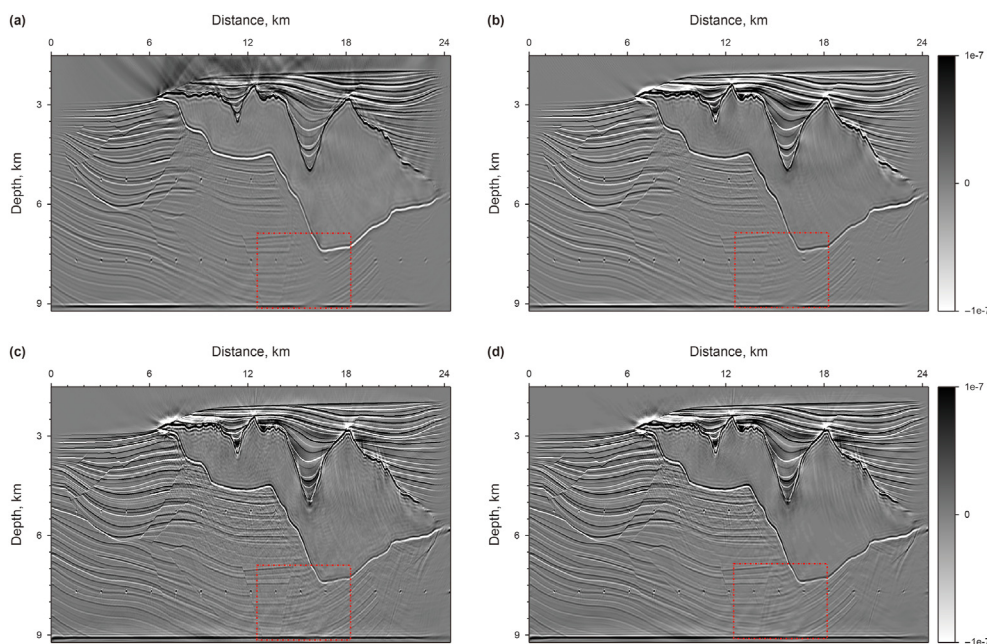


Fig. 10. Migration images: (a) LSRTM with the zero-lag cross-correlation imaging condition at the 20th iteration; (b) LSRTM with the de-primary imaging condition at the 15th iteration; (c) Huber inversion-based RTM with the de-primary imaging condition at the 15th iteration; (d) Huber inversion-based RTM with the de-primary imaging condition and the curvelet-domain sparse constraint at the 15th iteration.

Fig. 9a shows the RTM images with the zero-lag cross-correlation imaging condition. As can be seen, strong low-wavenumber migration artifacts in Fig. 9a affect the imaging result severely: it covers up the real high-wavenumber reflectors. Fig. 9b depicts the Laplacian filtered result of the image in Fig. 9a. After removing low-wavenumber artifacts through Laplacian filtering, there are still noticeable false high-wavenumber reflectors as well as artifacts. By contrast, the de-primary imaging condition gives a clean image (Fig. 9c) by effectively suppressing both low- and high-wavenumber migration artifacts. But subsalt reflectors are still unclear due to poor illumination.

We then carried out inversion-based migration starting from the result shown in Fig. 9c. Figs. 10a and b depict the results of LSRTM with the zero-lag cross-correlation imaging condition and the de-primary imaging condition, respectively. As can be seen, LSRTM with the zero-lag cross-correlation imaging condition improves sub-salt imaging. However, it introduces both low- and high-wavenumber artifacts back. Note that the result will be worse if the inversion starts from the result of the zero-lag cross-correlation imaging condition shown in Fig. 9a. By contrast, LSRTM with the de-primary imaging condition produces much fewer artifacts. However, the area indicated by the red box, the enlarged view of which is shown in Fig. 11, is still not imaged well. The main reason is that strong events generated from the salt interfaces dominate the L_2 -norm-based LSRTM inversion. Therefore, we carry out Huber inversion-based RTM with the de-primary imaging condition. The continuity of the strata (indicated by the red box in Fig. 10c) is improved. Besides, the broken section of the bottom reflector at a distance of 16 km is also fixed. This demonstrates that Huber inversion-based RTM with the de-primary condition has the capability to image complex models even with insufficient illumination. However, the inversion process also generates high-wavenumber migration noise while promoting the resolution of migration images. To suppress the noise, we added a curvelet-domain sparse constraint to the inversion. Fig. 10d shows the result, in which the high-wavenumber migration noise is suppressed largely. Note that there are still some migration artifacts

below the salt body. These artifacts are mainly generated by interbed multiples in the record. One way to remove them is to carry out de-multiples before migration (Griffiths et al., 2011; Dutta et al., 2019).

4. Discussion

4.1. Comparison of convergence speed

Conventional LSRTM is a quadratic problem, where the data linearly depend on the model. Its migration operator is the adjoint of its modeling operator. Consequently, a linear conjugate-gradient method is an effective and efficient way to solve this problem. This is demonstrated by the red curves in Fig. 12. However, when the de-primary imaging condition is applied to form an approximate gradient for reducing migration artifacts, the approximate gradient direction is no longer the exact gradient direction. As a result, the PLCG method will lose conjugacy as inversion progresses. This is illustrated by the black curve in Fig. 12. To mitigate this problem, Kim et al. (2019) carried out LSRTM with the de-primary imaging condition in the early iterations and then switched the imaging condition to the zero-lag cross-correlation imaging condition to improve the convergence, which is shown by the green curve. However, the drawback of Kim's strategy is that the switch will introduce both low- and high-wavenumber artifacts back. Since the reason for slow convergence is due to the approximation to the gradient, we chose the PNCG method described in Algorithm 1 to guarantee the convergence of LSRTM and Huber inversion-based RTM with the de-primary condition. As illustrated by the yellow and blue curves in Fig. 12, the inversions with the de-primary imaging condition converge. The convergence is also demonstrated in Fig. 13, which shows the Marmousi result of LSRTM with the de-primary imaging condition using PLCG and PNCG. As can be seen, PNCG suppresses artifacts (indicated by the red arrow) more effectively than PLCG. In the test, the Huber inversion-based RTM is even with the L_1 -norm constraint, which leads the objective function to be non-smooth. As also can be seen, the Huber inversion-based RTM converges faster

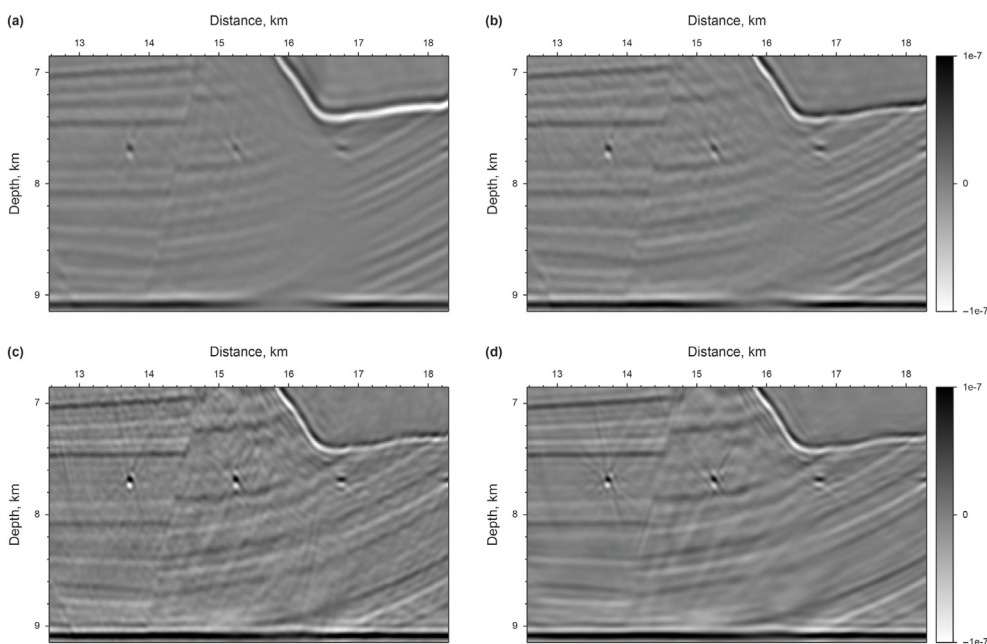


Fig. 11. Enlarged view of the migration results shown in Fig. 10. (a)-(d) are corresponding to the red box areas in (a)-(d) of Fig. 10, respectively.

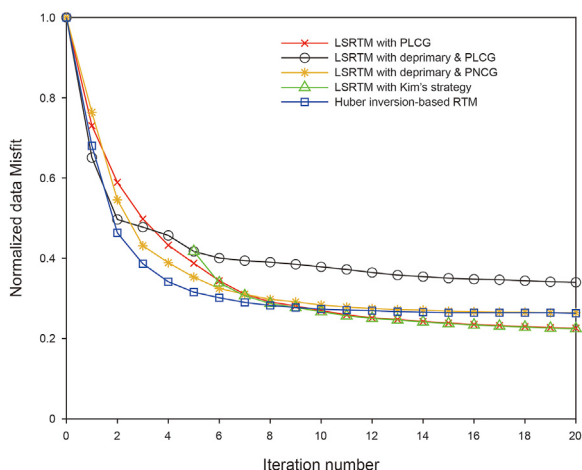


Fig. 12. Convergence curves of migration on Marmousi2 model: the red curve is for LSRTM with the zero-lag cross-correlation imaging condition and preconditioned linear conjugate gradient (PLCG); the black curve is for LSRTM with the de-primary imaging condition and PLCG; the yellow curve is for LSRTM with the de-primary imaging condition and preconditioned nonlinear conjugate gradient (PNCG); the green curve is for Kim's LSRTM using the zero-lag cross-correlation imaging condition and PLCG; the blue curve is for our proposed method.

than LSRTM in the early iteration. Note that the key benefit of the Huber norm is to promote the weak events during the inversion. We also should note that LSRTM with the zero-lag cross-correlation imaging condition gives the best convergence but this is a result of creating migration artifacts to overfit the record.

Utilizing the de-primary imaging condition in LSRTM can achieve better convergence with fewer iterations. To implement the de-primary imaging condition, it is necessary to perform the Hilbert Transform on wavefields along the time and space directions. In this paper, we implement the Hilbert Transform along the time direction by convolving the residual with the Hilbert function in the time domain while we achieve the Hilbert Transform along the space direction by multiplication after transforming into the frequency domain with FFT. Our tests show the de-primary imaging condition increases about 20% extra cost. Table 1 lists the time cost of LSRTM with and without the de-primary imaging condition. In the Marmousi2 example, we used nine nodes, each with two E5-2630 v4 CPUs.

4.2. Noise suppression with the curvelet-domain sparse constraint

The data sets of previous examples are generated by solving the full acoustic wave equation with an absorbing boundary condition. Consequently, all the data sets contain interbed multiples, which

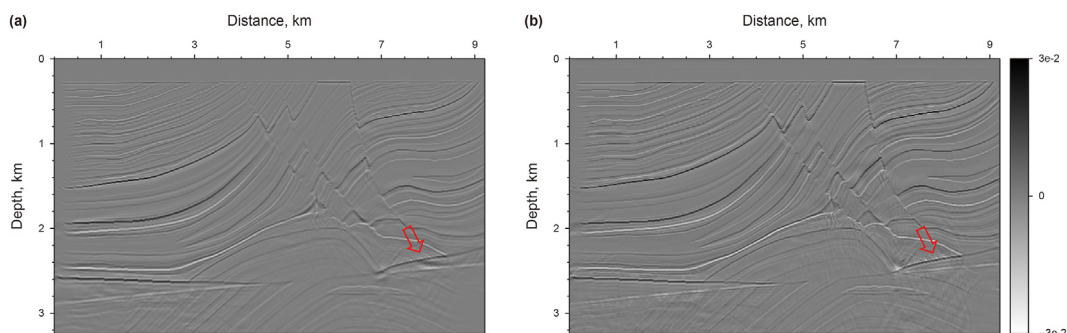


Fig. 13. The migration images of LSRTM with the de-primary imaging condition: (a) using PLCG and (b) using PNCG. The inversion takes 20 iterations.

Table 1
The time cost of LSRTM with and without the de-primary imaging condition.

	Time cost of iteration	Time cost of 20 iterations
LSRTM without de-primary imaging condition	1416 s	28735 s
LSRTM with de-primary imaging condition	1697 s	33912 s

are a type of coherent noise. The results demonstrate that the curvelet-domain sparse constraint can partially suppress the coherent noise. The fundamental mechanism is that the coherent noise does not focus very well, and its image is prone to complex artifacts, which turns into dense features in the curvelet-domain and therefore can be suppressed by the curvelet-domain sparse constraint. The partial suppression on coherent noise can be seen from Figs. 10 and 11. In terms of suppressing random noise, the curvelet-domain sparse constraint is more effective than coherent noise. We add heavy random noise ($S/N = 5$) in the observed data of the Marmousi2 model (Fig. 14b). Figs. 15a and b depict the results of LSRTM. As can be seen, the result of LSRTM without the constraint

has a lot of high-wavenumber noise as the observed data contain heavy random noise. By contrast, LSRTM with the curvelet-domain sparse constraint suppresses the heavy random noise significantly.

4.3. Accuracy of migration velocity

In the example of the Sigsbee2A model, we found that the accuracy of the migration velocity model is very important for inversion-based RTM to image the salt body properly. Fig. 16 shows the migration velocity provided from the original distribution (Paffenholz et al., 2002), which is extremely smooth and contains relatively large velocity errors, for instance, the valley area is crossed by the wave-path in Fig. 16. The migration results are shown in Figs. 17a and c. As can be seen, the lower boundary of the salt body is broken as well as weak with the original migration velocity (indicated by the blue arrows in Figs. 17a and c). By contrast, a more accurate velocity shown in Fig. 7b, which is generated by smoothing the true velocity model with a 20-point cosine-square window, gives a continuous and focused image of the lower boundary. The reason is that the inaccurate velocity in the valley (crossed by the wave-path shown in Fig. 16) leads to a travel-time error comparable to half a cycle; the inversion process then

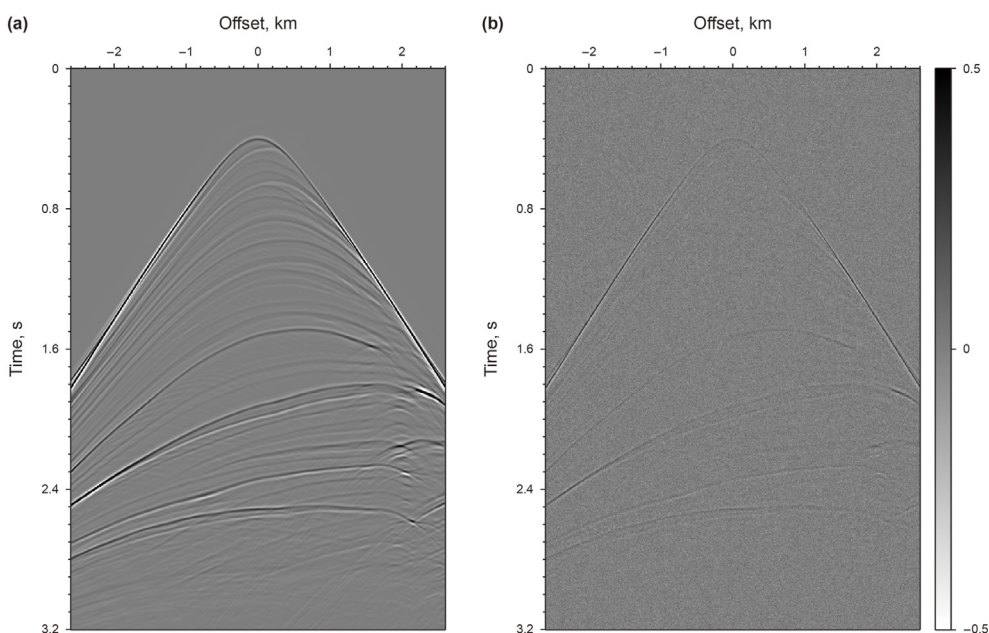


Fig. 14. One shot profile (a) without and (b) with random noise. The source is located at a distance of 2.657 km. The S/N ratio is 5, which is achieved using the Seismic Unix command “suaddnoise”. The direct arrivals are removed.

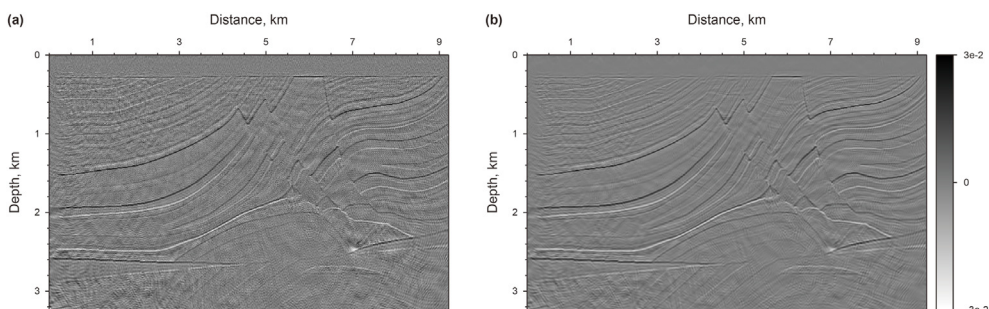


Fig. 15. Migration images with random noise: (a) LSRTM with the de-primary imaging condition at the 20th iteration; (b) LSRTM with the de-primary imaging condition and the curvelet-domain sparse constraint at the 20th iteration.

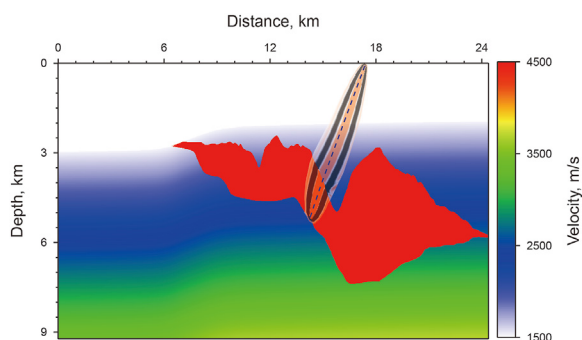


Fig. 16. The migration velocity model provided by the original distributor. A wave-path (fat ray) between a point on the surface and a point on the salt boundary is overlapped on the migration velocity model.

creates a shifted reverse polarity update resulting in the broken and weak image of the lower boundary.

4.4. Alternative loss function

In addition to the Huber-norm loss function, there are other loss functions for suppressing large residuals. One of them is based on the Student's T-distribution (Aravkin et al., 2011), which can be expressed as

$$\varphi_T = \frac{s+l}{2} \sum_i \log \left(s + (d^i - d_{obs}^i)^2 \right) \tag{13}$$

where s represents the degrees of freedom and l is the dimension of the vector \mathbf{d} , i.e. $l = 1$ for a 1D vector. Aravkin et al. (2011) suggested that s can be set as a small number, e.g. 2 or 3. If $s = 1$, then the loss function based on the Student's T-distribution is equivalent to the loss function based on the Cauchy distribution with the hyper-parameter equal to 1. For comparison, we show the two loss functions and their derivatives in Fig. 18. As can be seen, the loss function based on the Student's T-distributions suppresses the large residuals more than the Huber-norm loss function. One drawback, however, is that it is no longer strongly convex, resulting in slower convergence. On the contrary, the Huber-norm loss function can also suppress large residuals and is strictly convex. In addition, the Huber-norm loss function can adjust the value of δ to suppress large residuals adaptively. So that in the early iterations, a large δ can be used to relax the suppression of large residuals to avoid slow convergence.

5. Conclusion

Standard LSRTM can yield better images than conventional RTM. However, three problems still exist: (1) inversion can be dominated by strong events in the residual; (2) low-wavenumber artifacts in

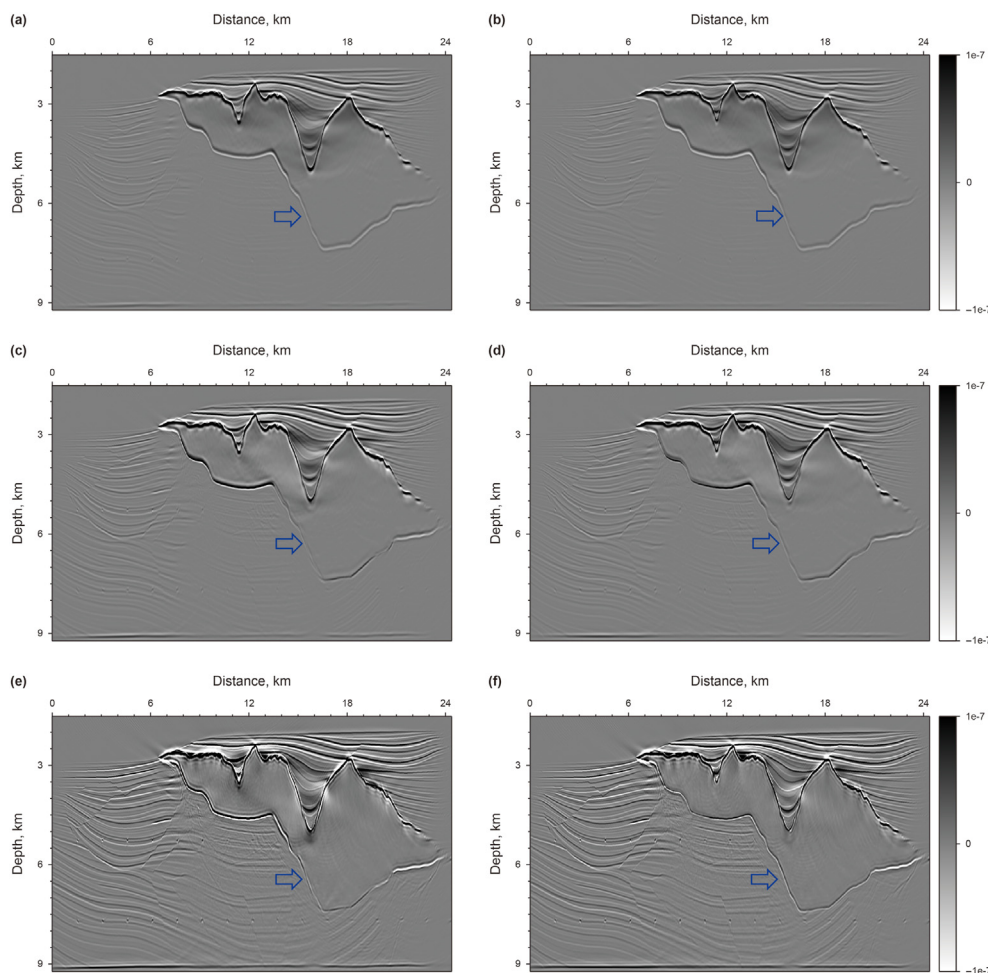


Fig. 17. (a) RTM with the de-primary imaging condition and the migration velocity model shown in Fig. 16. (b) The same as (a) but with the migration velocity model in Fig. 7b. (c) LSRTM with the de-primary imaging condition and the migration velocity model shown in Fig. 16 at the 1st iteration. (d) The same as (c) but with the migration velocity model in Fig. 7b. (e) and (f) are the counterpart of (c) and (d) at the 10th iteration, respectively.

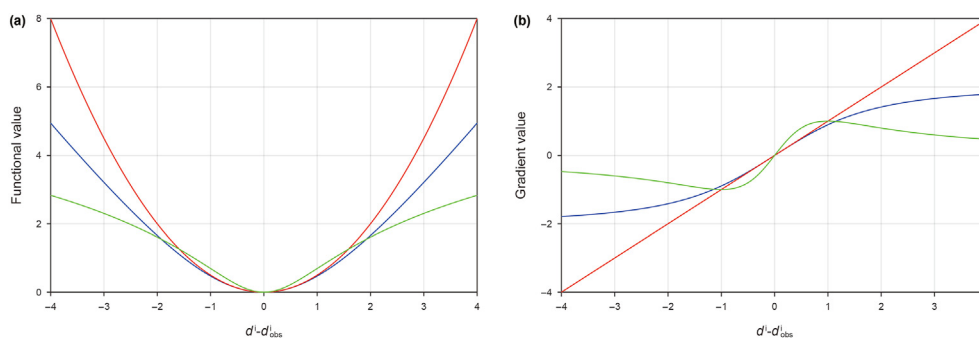


Fig. 18. A sketch of the L_2 -norm loss function (red), the Student's T-distribution loss function (green, $s = 1$, $l = 1$) and the Pseudo-Huber loss function (blue, $\delta = 1$). (a) The functional value and (b) their gradients.

the gradient affect convergence speed and imaging results; (3) high-wavenumber noise is also enhanced as iteration increases. To solve these problems, we formed the inversion-based RTM by combining the Huber-norm of data misfit and the L_1 -norm of reflectivity in the curvelet-domain as the objective function. The mixed norm functional is solved by using an improved IST method, which is updated along the Polak-Ribière conjugate-gradient direction by using a preconditioned non-linear conjugate-gradient method. The de-primary imaging condition is used to remove the low-wavenumber artifacts as well as high-wavenumber false reflectors in the gradient. The numerical examples, especially the Sigsbee2A model, demonstrate that the proposed method has greatly improved the migration image, especially the sub-salt image, compared with traditional RTM and LSRTM. In theory, all the techniques presented in the paper can be extended into 3D.

Acknowledgements

This work was supported by National Key R&D Program of China (No. 2018YFA0702502), NSFC (Grant No. 41974142, 42074129, and 41674114), Science Foundation of China University of Petroleum (Beijing) (Grant No. 2462020YXZZ005), and State Key Laboratory of Petroleum Resources and Prospecting (Grant No. PRP/indep-4-2012). The authors would like to thank reviewers for their comments and suggestions, which helped to improve and clarify the manuscript significantly.

References

Aravkin, A., van Leeuwen, T., Herrmann, F., 2011. Robust full-waveform inversion using the Student's t-distribution. In: 81st annual international meeting. SEG Technical Program Expanded Abstracts, 2669–2673. <https://doi.org/10.1190/1.3627747>.

Bai, M., Wu, J., Zhang, H., et al., 2019. Gaussian beam reconstruction of seismic data. *Geophysics* 84 (5), S373–S387. <https://doi.org/10.1190/geo2018-0664.1>.

Baysal, E., Kosloff, D.D., Sherwood, J.W.C., 1983. Reverse time migration. *Geophysics* 48 (11), 1514–1524. <https://doi.org/10.1190/1.1441434>.

Bleistein, N., Zhang, Y., Xu, S., et al., 2005. Migration/inversion: think image point coordinates, process in acquisition surface coordinates. *Inverse Probl.* 21 (5), 1715–1744. <https://doi.org/10.1088/0266-5611/21/5/013>.

Candès, E.J., Donoho, D.L., 2004. New tight frames of curvelets and optimal representations of objects with piecewise C^2 singularities: curvelets and Edges. *Commun. Pure Appl. Math.* 57 (2), 219–266. <https://doi.org/10.1002/cpa.10116>.

Candès, E.J., Guo, F., 2002. New multiscale transforms, minimum total variation synthesis: applications to edge-preserving image reconstruction. *Signal Process.* 82 (11), 1519–1543. [https://doi.org/10.1016/S0165-1684\(02\)00300-6](https://doi.org/10.1016/S0165-1684(02)00300-6).

Charbonnier, P., Blanc-Feraud, L., Aubert, G., et al., 1997. Deterministic edge-preserving regularization in computed imaging. *IEEE Trans. Image Process.* 6 (2), 298–311. <https://doi.org/10.1109/83.551699>.

Claerbout, J.F., 1971. Toward a unified theory of reflector mapping. *Geophysics* 36 (3), 467–481. <https://doi.org/10.1190/1.1440185>.

Claerbout, J.F., Doherty, S.M., 1972. Downward continuation of moveout-corrected seismograms. *Geophysics* 37 (5), 741–768. <https://doi.org/10.1190/1.1440298>.

Claerbout, J.F., Karrenbach, M., 1992. Electronic documents give reproducible research a new meaning. In: 62nd Annual International Meeting. SEG Technical Program Expanded Abstracts, pp. 601–604. <https://doi.org/10.1190/1.1822162>.

Dai, W., Schuster, G.T., 2013. Plane-wave least-squares reverse-time migration. *Geophysics* 78 (4), S165–S177. <https://doi.org/10.1190/geo2012-0377.1>.

Dutta, G., 2017. Sparse least-squares reverse time migration using seislets. *J. Appl. Geophys.* 136, 142–155. <https://doi.org/10.1016/j.jappgeo.2016.10.027>.

Dutta, G., Huang, H., Kanakamedala, K., et al., 2019. Practical strategies for interbed multiple attenuation. In: 89th Annual International Meeting. SEG Technical Program Expanded Abstracts, pp. 4550–4554. <https://doi.org/10.1190/segam2019-3216600.1>.

Etgen, J., Gray, S.H., Zhang, Y., 2009. An overview of depth imaging in exploration geophysics. *Geophysics* 74 (6), WCA5–WCA17. <https://doi.org/10.1190/1.3223188>.

Fang, J.W., Zhou, H.M., Chen, H., et al., 2019. Source-independent elastic least-squares reverse time migration. *Geophysics* 84 (1), S1–S16. <https://doi.org/10.1190/geo2017-0847.1>.

Fei, T.W., Luo, Y., Aramco, S., et al., 2010. De-blending reverse-time migration. In: 80th Annual International Meeting. SEG Technical Program Expanded Abstracts, pp. 3130–3134. <https://doi.org/10.1190/1.3513496>.

Fei, T.W., Luo, Y., Yang, J., et al., 2015. Removing false images in reverse time migration: the concept of de-primary. *Geophysics* 80 (6), S237–S244. <https://doi.org/10.1190/geo2015-0289.1>.

Gazdag, J., 1978. Wave equation migration with the phase-shift method. *Geophysics* 43 (7), 1342–1351. <https://doi.org/10.1190/1.1440899>.

Griffiths, M., Hembd, J., Prigent, H., 2011. Applications of interbed multiple attenuation. *Lead. Edge* 30 (8), 906–912. <https://doi.org/10.1190/1.3626498>.

Gu, B., Han, J., Ren, Z., et al., 2019. 2D least-squares elastic reverse time migration of multicomponent seismic data. *Geophysics* 84 (6), S523–S538. <https://doi.org/10.1190/geo2018-0720.1>.

Guo, S., Wang, H., 2020. Image domain least-squares migration with a Hessian matrix estimated by non-stationary matching filters. *J. Geophys. Eng.* 17 (1), 148–159. <https://doi.org/10.1093/jge/gxz098>.

Hill, N.R., 1990. Gaussian beam migration. *Geophysics* 55 (11), 1416–1428. <https://doi.org/10.1190/1.1442788>.

Hill, N.R., 2001. Prestack Gaussian-beam depth migration. *Geophysics* 66 (4), 1240–1250. <https://doi.org/10.1190/1.1487071>.

Huber, P.J., 1964. Robust estimation of a location parameter. *Ann. Math. Stat.* 35 (1), 73–101. <https://doi.org/10.1214/aoms/1177703732>.

Hu, Y., Pan, D.M., Wu, R.S., et al., 2021. Joint least square reverse time migration of phase and amplitude in the time-frequency domain. *Chin. J. Geophys.* 64 (1), 209–223. <https://doi.org/10.6038/cjg202100203>.

Jaramillo, H.H., Bleistein, N., 1999. The link of Kirchhoff migration and demigration to Kirchhoff and Born modeling. *Geophysics* 64 (6), 1793–1805. <https://doi.org/10.1190/1.1444685>.

Keho, T.H., Beydoun, W.B., 1988. Paraxial ray Kirchhoff migration. *Geophysics* 53 (12), 1540–1546. <https://doi.org/10.1190/1.1442435>.

Kim, Y.S., Almomin, A., Jeong, W., et al., 2019. Least-squares reverse time migration using analytic-signal-based wavefield decomposition. *Geophysics* 84 (3), S149–S157. <https://doi.org/10.1190/geo2018-0118.1>.

Li, C., Huang, J.P., Li, Z.C., et al., 2017. Regularized least-squares migration of simultaneous-source seismic data with adaptive singular spectrum analysis. *Petrol. Sci.* 14 (1), 61–74. <https://doi.org/10.1007/s12182-016-0134-1>.

Li, F., Gao, J., Gao, Z., et al., 2020. Least-squares reverse time migration with sparse regularization in the 2D wavelet domain. *Geophysics* 85 (6), S313–S325. <https://doi.org/10.1190/geo2018-0763.1>.

Liu, Y., Symes, W.W., Li, Z., 2013. Multisource least-squares extended reverse time migration with preconditioning guided gradient method. In: 83rd Annual International Meeting. SEG Technical Program Expanded Abstracts, pp. 3709–3715. <https://doi.org/10.1190/segam2013-1251.1>.

Liu, F., Zhang, G., Morton, S.A., et al., 2011. An effective imaging condition for reverse-time migration using wavefield decomposition. *Geophysics* 76 (1), S29–S39. <https://doi.org/10.1190/1.3533914>.

Lin, Y.Z., Lian, J., 2015. Least-squares reverse-time migration with modified total-variation regularization. In: 85th Annual International Meeting. SEG Technical Program Expanded Abstracts, pp. 4264–4269. <https://doi.org/10.1190/segam2015-5914034.1>.

- Liu, Q., Lu, Y., Zhang, H., 2020. Fast single-step least-squares reverse-time imaging via adaptive matching filters in beams. *IEEE Trans. Geosci. Rem. Sens.* 58 (3), 1913–1919. <https://doi.org/10.1109/TGRS.2019.2950645>.
- McMechan, G.A., 1983. Migration by exploration of time-dependent boundary values. *Geophys. Prospect.* 31 (3), 413–420. <https://doi.org/10.1111/j.1365-2478.1983.tb01060.x>.
- Murphy, G.E., Gray, S.H., 1999. Manual seismic reflection tomography. *Geophysics* 64 (5), 1546–1552. <https://doi.org/10.1190/1.1444658>.
- Neal, P., Stephen, B., 2014. Proximal algorithms. *Foundations & Trends in Optimization* 1 (3), 127–239. <https://doi.org/10.1561/2400000003>.
- Qu, Y.M., Li, J., 2019. Q-compensated reverse time migration in viscoacoustic media including surface topography. *Geophysics* 84, S201–S217. <https://doi.org/10.1190/geo2018-0313.1>.
- Qu, Y.M., Huang, C.P., Liu, C., et al., 2021. Full-path compensated least-squares reverse time migration of joint primaries and different-order multiples for deep-marine environment. *IEEE Trans. Geosci. Rem. Sens.* 59, 7109–7121. <https://doi.org/10.1109/TGRS.2020.3024189>.
- Paffenholz, J., McLain, B., Zaske, J., et al., 2002. Subsalt multiple attenuation and imaging: observations from the Sigsbee2B synthetic dataset. In: 72nd Annual International Meeting. SEG Technical Program Expanded Abstracts, pp. 2122–2125. <https://doi.org/10.1190/1.1817123>.
- Qu, Y.M., Huang, J.P., Li, Z.C., et al., 2017. Attenuation compensation in anisotropic least-squares reverse time migration. *Geophysics* 82, S411–S423. <https://doi.org/10.1190/geo2016-0677.1>.
- Schneider, W.A., 1978. Integral formulation for migration in two and three dimensions. *Geophysics* 43 (1), 49–76. <https://doi.org/10.1190/1.1440828>.
- Shi, Y., Zhang, W., Wang, Y., 2019. Seismic elastic RTM with vector-wavefield decomposition. *J. Geophys. Eng.* 16 (3), 509–524. <https://doi.org/10.1093/jge/gxz023>.
- Trad, D., 2020. Assumptions and goals for least squares migration. *Geophys. Prospect.* 68 (3), 802–814. <https://doi.org/10.1111/1365-2478.12879>.
- Wang, R., Wang, Y., 2017. Multichannel algorithms for seismic reflectivity inversion. *J. Geophys. Eng.* 14 (1), 41–50. <https://doi.org/10.1088/1742-2132/14/1/41>.
- Wang, X.Y., Zhang, J.J., Xu, H.Q., et al., 2021. Least-squares reverse time migration with wavefield decomposition based on the Poynting vector. *Chin. J. Geophys.* 64 (2), 645–655. <https://doi.org/10.6038/cjg202100120>.
- Whitmore, N.D., 1983. Iterative depth migration by backward time propagation. In: 53rd Annual International Meeting. SEG Technical Program Expanded Abstracts, pp. 382–385. <https://doi.org/10.1190/1.1893867>.
- Wu, D., Yao, G., Cao, J., et al., 2016. Least-squares RTM with L1 norm regularisation. *J. Geophys. Eng.* 13 (5), 666–673. <https://doi.org/10.1088/1742-2132/13/5/666>.
- Xu, S., Zhang, Y., Tang, B., 2011. 3D angle gathers from reverse time migration. *Geophysics* 76 (2), S77–S92. <https://doi.org/10.1190/1.3536527>.
- Yang, J., Elita Li, Y., Cheng, A., et al., 2019a. Least-squares reverse time migration in the presence of velocity errors. *Geophysics* 84 (6), S567–S580. <https://doi.org/10.1190/geo2018-0887.1>.
- Yang, J., Zhu, H., McMechan, G., et al., 2019b. Elastic least-squares reverse time migration in vertical transverse isotropic media. *Geophysics* 84 (6), S539–S553. <https://doi.org/10.1190/geo2019-0005.1>.
- Yao, G., Jakubowicz, H., 2016. Least squares reverse-time migration in a matrix-based formulation: LSRTM in a matrix-based formulation. *Geophys. Prospect.* 64 (3), 611–621. <https://doi.org/10.1111/1365-2478.12216>.
- Yao, G., Wu, D., 2015. Least-squares reverse-time migration for reflectivity imaging. *Sci. China Earth Sci.* 58 (11), 1982–1992. <https://doi.org/10.1007/s11430-015-5143-1>.
- Yao, G., da Silva, N.V., Wu, D., 2018. Forward modelling formulas for least-squares reverse-time migration. *Explor. Geophys.* 49 (4), 506–518. <https://doi.org/10.1017/EG16157>.
- Yao, G., Wu, D., Wang, S.X., 2020. A review on reflection-waveform inversion. *Petrol. Sci.* 17 (2), 334–351. <https://doi.org/10.1007/s12182-020-00431-3>.
- Zhang, R., Huang, J.P., Zhuang, S.B., et al., 2019. Target-oriented Gaussian beam migration using a modified ray tracing scheme. *Petrol. Sci.* 16 (6), 1301–1319. <https://doi.org/10.1007/s12182-019-00388-y>.
- Zhang, Y., Duan, L., Xie, Y., 2015. A stable and practical implementation of least-squares reverse time migration. *Geophysics* 80 (1), V23–V31. <https://doi.org/10.1190/geo2013-0461.1>.
- Zhang, Y.B., Liu, Y.K., Yi, J., et al., 2021. First-order multiples imaging aided by water bottom. *Petrol. Sci.* 18 (6), 1650–1661. <https://doi.org/10.1016/j.petsci.2021.09.036>.
- Zhang, Z., Tian, L., Zhou, D., et al., 2011. Application of Gaussian beam migration in imaging of the steep complex structure in Tan-Lu strike-slip fault zone of Bohai Sea in China. In: 81st Annual International Meeting. SEG Technical Program Expanded Abstracts, pp. 3435–3439. <https://doi.org/10.1190/1.3627911>.
- Zhao, Y., Niu, F.L., Fu, L., et al., 2021. Local events-based fast RTM surface-offset gathers via dip-guided interpolation. *Petrol. Sci.* 18 (3), 773–782. <https://doi.org/10.1007/s12182-021-00557-y>.