



Sparse Constrained Least-Squares Reverse Time Migration Based on Kirchhoff Approximation

Xu Hong-Qiao^{1,2,3}, Wang Xiao-Yi^{1,2,3}, Wang Chen-Yuan^{1,2,3} and Zhang Jiang-Jie^{1,2*}

¹Key Laboratory of Petroleum Resources Research, Institute of Geology and Geophysics, Chinese Academy of Sciences, Beijing, China, ²Innovation Academy for Earth Science, Chinese Academy of Sciences, Beijing, China, ³University of Chinese Academy of Sciences, Beijing, China

OPEN ACCESS

Edited by:

Hao Hu,
University of Houston, United States

Reviewed by:

Jincheng Xu,
Southern University of Science and
Technology, China
Chuang Li,
Xi'an Jiaotong University, China

*Correspondence:

Zhang Jiang-Jie
zhangjj@mail.iggcas.ac.cn

Specialty section:

This article was submitted to
Solid Earth Geophysics,
a section of the journal
Frontiers in Earth Science

Received: 28 June 2021

Accepted: 17 August 2021

Published: 03 September 2021

Citation:

Hong-Qiao X, Xiao-Yi W, Chen-Yuan W
and Jiang-Jie Z (2021) Sparse
Constrained Least-Squares Reverse
Time Migration Based on
Kirchhoff Approximation.
Front. Earth Sci. 9:731697.
doi: 10.3389/feart.2021.731697

Least-squares reverse time migration (LSRTM) is powerful for imaging complex geological structures. Most researches are based on Born modeling operator with the assumption of small perturbation. However, studies have shown that LSRTM based on Kirchhoff approximation performs better; in particular, it generates a more explicit reflected subsurface and fits large offset data well. Moreover, minimizing the difference between predicted and observed data in a least-squares sense leads to an average solution with relatively low quality. This study applies L1-norm regularization to LSRTM (L1-LSRTM) based on Kirchhoff approximation to compensate for the shortcomings of conventional LSRTM, which obtains a better reflectivity image and gets the residual and resolution in balance. Several numerical examples demonstrate that our method can effectively mitigate the deficiencies of conventional LSRTM and provide a higher resolution image profile.

Keywords: least-squares reverse time migration (LSRTM), kirchhoff approximation, L1-norm regularization, sparsity constraint, born approximation

INTRODUCTION

Seismic migration is an inverse procedure of forward modeling, which can restore the interior of the earth medium with record data. Specifically, migration attempts to eliminate the effects caused by the process of physical propagation and obtain an image that clearly depicts the structural information of interest. Reverse time migration (RTM), a state-of-the-art seismic imaging method (Baysal et al., 1983; McMechan, 1983), identifies the aforementioned acausal procedure appropriately. Based on two-way wave equation, RTM is powerful for handling complex geological settings and velocity with dramatic variation in the lateral direction. Therefore, it can deal with steep dips and salt dome better than conventional migration (Zhu and Lines, 1998; Yoon et al., 2003; Liu et al., 2010). However, most migration methods, including RTM, use the adjoint operator to compute the image instead of the inverse operator (Tarantola, 1984). Practical data suffers from many factors, such as irregular acquisition geometry and limited aperture of the acquisition system. These deficiencies generate artifacts and degrade the resolution. To overcome these limitations, least-squares migration (LSM) was proposed to combine with RTM (Liu et al., 2011; Dai et al., 2012). Therefore, seismic imaging can be regarded as a linearized inverse problem. With a proper initial velocity model, seismic records can be inverted to a more accurate profile. LSRTM iteratively reduces the residual between predicted data and observed data in a least-squares framework; therefore, the adjoint operator can keep approaching the inverse operator. Many results have indicated that LSRTM has a better performance than conventional RTM and migration (Zhang et al., 2015; Dutta and Schuster, 2014; Liu et al., 2016).

The precondition of seismic inversion is forward modeling, which maps the parameter model to seismic data. There are two main approaches to build linear approximation between physical model and wavefield (Yang and Zhang, 2019). One is the most commonly used Born approximation based on small perturbation (Beylkin, 1985; Bleistein, 1987). This requires that high-order scattered wavefields are much weaker than primary field. The Born operator describes a linear relationship between model perturbation and primary reflected wave. It divides the wavefield into two parts: background wavefield and perturbation wavefield. LSRTM based on Born approximation can achieve model perturbation with these two fields. In addition, an alternative scheme for modeling is Kirchhoff approximation (Bleistein, 1987). Compared with Born modeling, the Kirchhoff operator delineates the connection between primary reflected wave and reflectivity. Different operators lead to distinctive results under these two physical contexts. However, neither Born nor Kirchhoff approximation can avoid the impact on seismic image in a least-squares sense. Because minimizing the L2 norm only provides an average solution (Wang, 2016; Wu et al., 2016). It is essential to seek a balance between the residual and resolution. According to geological recognition, the earth medium usually presents a layered spatial distribution. The reflection coefficient that mirrors strata attributes should be sparsity, that is, the part of model that does not generate reflected wave ought to be zero. Therefore, the inverted model needs a sparse limitation.

This study implements a Kirchhoff modeling formula for LSRTM promoted by sparsity. The reflectivity model should be regularized with L1 norm while minimizing the residual of wavefield in the form of the L2 norm. Referring to 'least absolute shrinkage and selection operator' (Lasso) problem (Tibshirani, 1996), this reformed LSRTM can be solved by the algorithm of spectral projected gradient for L1 minimization (SPGL1), which is designed to solve sparse least squares (van den Berg and Friedlander, 2011). Examples show that our method can effectively overcome the problems mentioned above.

METHOD

RTM has great advantages in imaging steep structures such as salt dome. However, it suffers from low-frequency noise compared to conventional migration. Least-squares migration can get closer iteratively to the optimal solution and eventually obtain a relatively high signal-to-noise ratio, high resolution and amplitude equalized profile that eliminates the influence of the acquisition system. It contains three steps: constructing a linear modeling problem first, using the forward and backward propagation wavefields to image, and finally updating the physics model according to the residual.

Linear Modeling Operators

The linearization of nonlinear forward problem is essential to seismic inversion, making the physical progress more explicit; moreover, converting the medium parameter becomes easier. The

choice of a linear operator will lead to different physical significance and images. It is a common way to use Born approximation to realize linearization. The real velocity model is divided into two parts: background velocity v_0 and velocity perturbation δv . Given a perturbation δv , it generates a corresponding wavefield perturbation δu . The Born operator describes the relationship between reflected wave and model perturbation. Specifically, the incident wave interacting with model perturbation becomes a new source, namely the Huygens principle, and then the new source generates wavefield perturbations. This can be expressed as follows in time domain:

$$\left(\frac{1}{v_0(\mathbf{x})^2} \frac{\partial^2}{\partial t^2} - \nabla^2\right) u_0(\mathbf{x}, t; \mathbf{x}_s) = f(t; \mathbf{x}_s) \tag{1}$$

$$\left(\frac{1}{v_0(\mathbf{x})^2} \frac{\partial^2}{\partial t^2} - \nabla^2\right) \delta u_b(\mathbf{x}, t; \mathbf{x}_s) = m(\mathbf{x}) \frac{1}{v_0(\mathbf{x})^2} \frac{\partial^2 u_0(\mathbf{x}, t; \mathbf{x}_s)}{\partial t^2} \tag{2}$$

where u_0 represents the background field propagating in v_0 , $f(t; \mathbf{x}_s)$ is the source signature located at \mathbf{x}_s and excited at t , the model perturbation is denoted by $m(\mathbf{x}) = 2\delta v(\mathbf{x})/v_0(\mathbf{x})$, which describes velocity changes compared to background velocity. \mathbf{x} is a point in model. This study assumes that the density ρ is a constant (Eq. 1) and (Eq. 2) can be rewritten in form of an integral using Green's theorem:

$$u_0(\mathbf{x}, t; \mathbf{x}_s) = f(t; \mathbf{x}_s) G_0(\mathbf{x}, t; \mathbf{x}_s) \tag{3}$$

$$\delta u_b(\mathbf{x}_g, t; \mathbf{x}_s) = \int m(\mathbf{x}) \frac{1}{v_0^2(\mathbf{x})} \frac{\partial^2 u_0(\mathbf{x}, t; \mathbf{x}_s)}{\partial t^2} G_0(\mathbf{x}_g, t; \mathbf{x}) d\mathbf{x} \tag{4}$$

where $G_0(\mathbf{x}, t; \mathbf{x}_s)$ is the Green's function from \mathbf{x}_s to \mathbf{x} , $G_0(\mathbf{x}_g, t; \mathbf{x})$ propagates from \mathbf{x} to \mathbf{x}_g . Green's function is governed by:

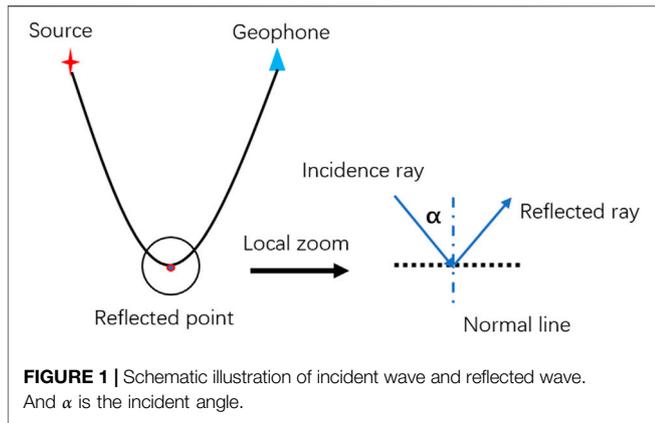
$$\left(\frac{1}{v_0(\mathbf{x})^2} \frac{\partial^2}{\partial t^2} - \nabla^2\right) G_0(\mathbf{x}, t; \mathbf{x}_s) = \delta(t; \mathbf{x}_s) \tag{5}$$

where the $\delta(t; \mathbf{x}_s)$ is Dirac function.

Born approximation represents scattered phenomenon caused by model perturbation, which could be a means of linearizing seismic inversion. However, this approximation is accurate when scattered field δu is much weaker than background field u_0 (Schuster, 2017), which is a disadvantage of Born approximation. It cannot describe kinematic and dynamic information of seismic waves well with strong reflector. And studies have shown that Born approximation has limited angle validity and it cannot appropriately predict the reflections generated with large incident angle (Yang and Zhang, 2019).

Compared to the Born operator, the Kirchhoff operator relates the reflectivity to wavefield perturbation. Therefore, it depicts the interaction between the incident field and reflectivity rather than velocity perturbation. There is a relationship between reflectivity and model perturbation when the perturbation and incident angle are small (Stolt and Weglein, 2012):

$$r(\mathbf{x}, \alpha) = \frac{i\omega}{2v_0(\mathbf{x}) \cos(\mathbf{x}, \alpha)} m(\mathbf{x}) \tag{6}$$



where the $r(\mathbf{x}, \alpha)$ is the reflection coefficient at point \mathbf{x} with incident angle α between the incidence and the normal line (Figure 1). This means that we can obtain the wavefield perturbation under the Kirchhoff approximation by substituting (Eq. 6) into (Eq. 4), and we have

$$\delta u_b(\mathbf{x}_g, t; \mathbf{x}_s) = \int \frac{2v_0(\mathbf{x})}{i\omega} r(\mathbf{x}, \alpha) \cos(\mathbf{x}, \alpha) \frac{1}{v_0^2(\mathbf{x})} \frac{\partial^2 u_0(\mathbf{x}, t; \mathbf{x}_s)}{\partial t^2} G_0(\mathbf{x}_g, t; \mathbf{x}) d\mathbf{x} \quad (7)$$

Here we turn Kirchhoff modeling equation into the same form as Born approximation. Then Eq. 7 can be rewritten as.

$$\left(\frac{1}{v_0(\mathbf{x})^2} \frac{\partial^2}{\partial t^2} - \nabla^2 \right) u_0(\mathbf{x}, t; \mathbf{x}_s) = - \int f(t; \mathbf{x}_s) dt \quad (8)$$

$$\left(\frac{1}{v_0(\mathbf{x})^2} \frac{\partial^2}{\partial t^2} - \nabla^2 \right) \delta u_k(\mathbf{x}, t; \mathbf{x}_s) = 2v_0(\mathbf{x}) r(\mathbf{x}, \alpha) \cos(\mathbf{x}, \alpha) \frac{1}{v_0(\mathbf{x})^2} \frac{\partial^2 u_0(\mathbf{x}, t; \mathbf{x}_s)}{\partial t^2} \quad (9)$$

It should be noted that the term $r(\mathbf{x}, \alpha)$ can be replaced by the generalized angle-dependent reflectivity model to get rid of the limitations of small perturbation and incident angle. Although there are some methods to solve the propagation direction of wave, such as Poynting vector and Plane Wave Decomposition (PWD), it is still tedious and time-consuming to obtain the angle term. Here we give an approximate scheme (Yang and Zhang, 2019).

Each shot can invert a reflectivity image, here we sum the images obtained by all shots. Then, we regard the summation as the final reflectivity model and use it to iterate. Approximately, we can get an averaged reflectivity model by multiple shots stacking. Therefore, we can get the predicted data by using this stacked reflectivity $R(\mathbf{x})$ rather than the angle-dependent term $r(\mathbf{x}, \alpha) \cos(\mathbf{x}, \alpha)$. Note that $R(\mathbf{x})$ is an averaged reflectivity over all illuminated angles.

$$R(\mathbf{x}) \approx \sum_{shot} r(\mathbf{x}, \alpha) \cos(\mathbf{x}, \alpha) \quad (10)$$

With this approximate reflectivity $R(\mathbf{x})$, we can express Eq. 9 as

$$\left(\frac{1}{v_0(\mathbf{x})^2} \frac{\partial^2}{\partial t^2} - \nabla^2 \right) \delta u_k(\mathbf{x}, t; \mathbf{x}_s) = 2v_0(\mathbf{x}) R(\mathbf{x}) \frac{1}{v_0(\mathbf{x})^2} \frac{\partial^2 u_0(\mathbf{x}, t; \mathbf{x}_s)}{\partial t^2} \quad (11)$$

In sum, with the relationship of reflectivity and model perturbation, two linear approximations have a similar form, which expresses their common ground. The difference between two approximations is also evident. From Eq. 6, $\cos \alpha$ approximately equals to one and can be ignored for a small incident angle. Therefore, reflectivity can be regarded as the spatial derivative of model perturbation. The inverted model after spatial derivation has a higher resolution, that is, the spectrum has been improved. More details are provided in the numerical tests.

Least-Squares With Sparse Optimization

In contrast to full waveform inversion (FWI) (Liu, et al., 2020), LSRTM first establishes a linear relationship between physical model and corresponding response (Tarantola, 1984), then it implements the inverse problem. The least-squares method (LSM) only requires the construction of a migration operator and inverse migration operator, which is conjugated to each other. It can reduce the residual between the observed and predicted data iteratively to approach the optimal solution of the inverse problem gradually. According to the linear approximation above, we can express them in the form of a matrix:

$$\mathbf{d} = \mathbf{Lm} \quad (12)$$

where the \mathbf{d} is predicted data, such as background or perturbation fields. \mathbf{L} represents modeling operator and \mathbf{m} is the physics model. Usually, it is assumed that the background velocity has been obtained in advance, and then the data can be predicted. Hence, the misfit function can be expressed as:

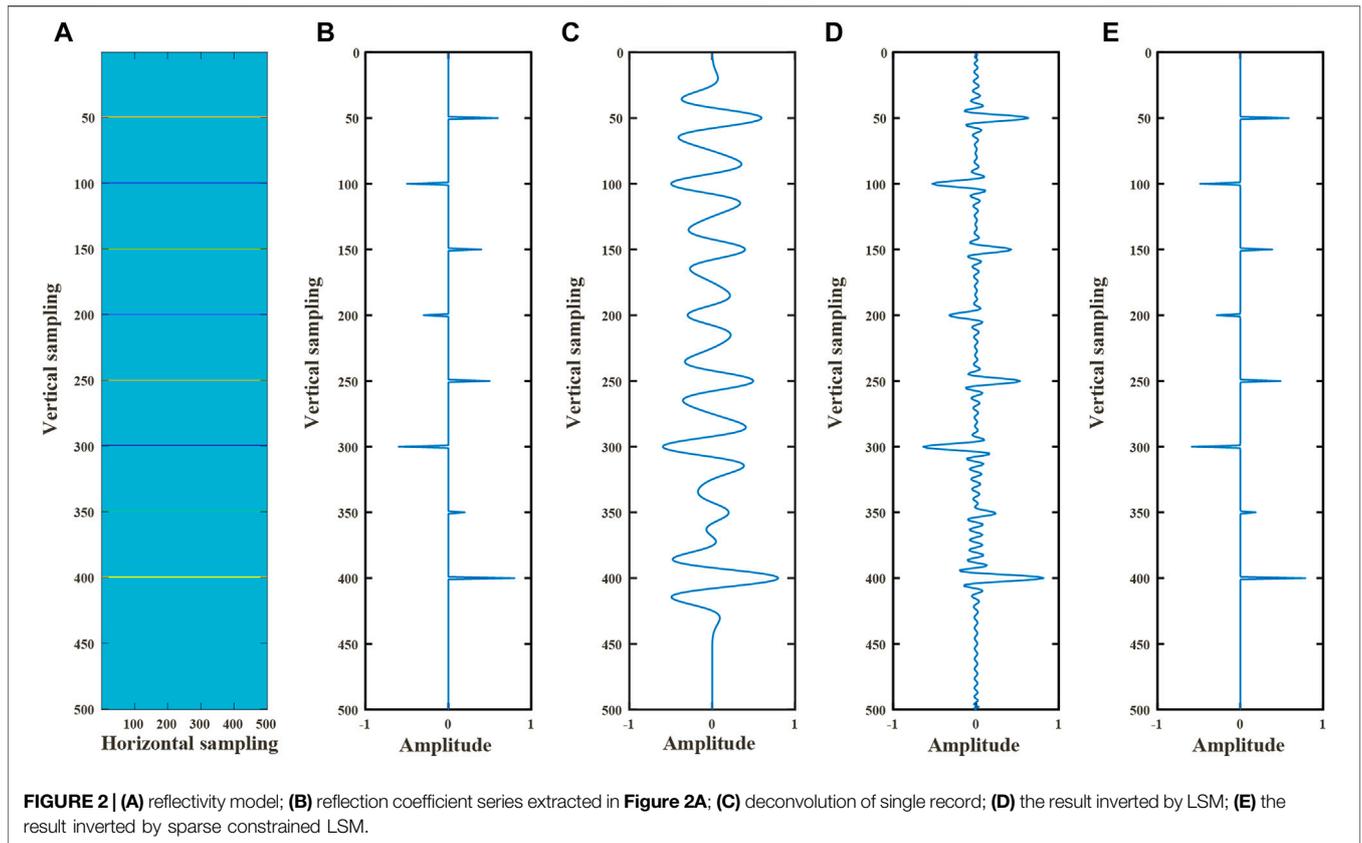
$$E(\mathbf{m}) = \|\mathbf{Lm} - \mathbf{d}_{obs}\|_2 \quad (13)$$

The model \mathbf{m} , which makes $\frac{\partial E(\mathbf{m})}{\partial \mathbf{m}}$ (the Jacobian matrix) equal to 0, is the optimal solution of Eq. 13. However, the computation of the Jacobian matrix is quite time consuming, particularly for seismic exploration. We adopt the adjoint-state method to calculate the adjoint operator \mathbf{L}^T of modeling operator \mathbf{L} . Specifically, the gradient of $E(\mathbf{m})$ can be obtained by back propagation of the wavefield residual and background field, here we give the gradient based on Kirchhoff approximation (Plessix, 2006; Wang et al., 2021):

$$\mathbf{m}_{mig} = \mathbf{L}^T \mathbf{d} = 2v_0(\mathbf{x}) \sum_{shot} \int \frac{1}{v_0(\mathbf{x})^2} \frac{\partial^2 u_0(\mathbf{x}, t; \mathbf{x}_s)}{\partial t^2} q(\mathbf{x}, t; \mathbf{x}_s) dt \quad (14)$$

Where $q(\mathbf{x}, t; \mathbf{x}_s)$ is the adjoint wavefield governed by:

$$\frac{1}{v_0(\mathbf{x})^2} \frac{\partial^2 q(\mathbf{x}, t; \mathbf{x}_s)}{\partial t^2} - \nabla^2 q(\mathbf{x}, t; \mathbf{x}_s) = \delta u(\mathbf{x}_r, t; \mathbf{x}_s) \quad (15)$$



According to **Eq. 13**, we can obtain a least-squares solution $\mathbf{m} = (\mathbf{L}^T \mathbf{L})^{-1} \mathbf{L}^T \mathbf{d}$. Note that LSM provides a smooth solution of the model, which is determined by the properties of the L2 norm. As a result, LSM has a limited ability to improve the quality of the image. Here we give a simple model to display the impact of LSM. In this example, we use the Ricker wavelet with a center frequency of 30 Hz and a time sampling interval of 1 ms. With convolution model theory, we can get seismic records via the convolution of Ricker wavelet and reflection coefficients, which can be obtained by $\mathbf{d} = \mathbf{Lm}$. Conversely, reflection coefficients can be obtained by the deconvolution of seismic records and wavelet, that is, $\mathbf{m}_{mig} = \mathbf{L}^T \mathbf{d}$. **Figure 2C** is the result of deconvolution, and it is hard to identify the reflectors. Compared to deconvolution, **Figure 2D** shows that LSM improves the resolution obviously. However, many oscillations caused by $(\mathbf{L}^T \mathbf{L})^{-1}$ near the real reflection coefficients should not exist. That's why we regard the least-squares solution as a smooth or average solution. The actual model indicates that the medium presents a layered spatial distribution, as shown in **Figure 2A** or **Figure 2B**, that is, the sub-surfaces are sparse. In **Figure 2E**, the inversion result with sparse constrained LSM performs quite well, and these oscillations generated by LSM are suppressed; thus, the resolution and sparsity of the reflection coefficient series are improved effectively.

Due to the feasibility and sparse property of L1-norm, we modify the objective function with L1 norm to realize sparse

reconstruction of the model in this study. Generally, **Eq. 13** can be reformed with two new problems.

1 Basis Pursuit (BP) problem

$$(BP) \quad \min \|\mathbf{m}\|_1, \quad \text{subject to } \mathbf{Lm} = \mathbf{d} \quad (16)$$

Eq. 16 depicts a BP problem that comes from compressed sense theory, and it aims to seek a sparse solution that satisfies $\mathbf{Lm} = \mathbf{d}$. However, practical seismic data inevitably contain noise, and **Eq. 16** can be modified as a basis pursuit denoising (BPDN) problem:

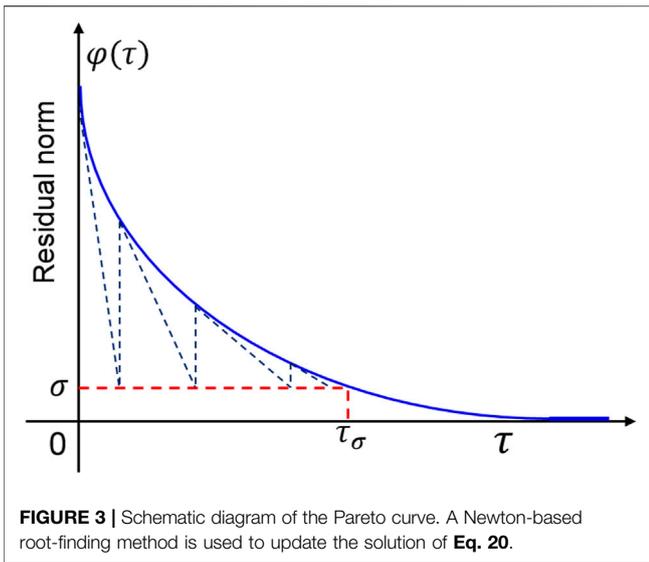
$$(BP_\sigma) \quad \min \|\mathbf{m}\|_1, \quad \text{subject to } \|\mathbf{Lm} - \mathbf{d}_{obs}\|_2 \leq \sigma \quad (17)$$

where the σ describes the noise level in the data, and **Eq. 16** and **Eq. 17** are equivalent to each other when $\sigma = 0$.

2 Least Absolute Shrinkage and Selection Operator (LASSO) problem

$$(LS_\tau) \quad \min \|\mathbf{Lm} - \mathbf{d}_{obs}\|_2, \quad \text{subject to } \|\mathbf{m}\|_1 \leq \tau \quad (18)$$

where the $\tau \geq 0$ is an explicit limitation of sparsity on \mathbf{m} . Problems (BP_σ) and (LS_τ) are different descriptions of the same question. They are equivalent in the sense that there exists a solution \mathbf{m}^* of (BP_σ) for a given σ , and there exists a corresponding τ that makes \mathbf{m}^* also be a solution of (LS_τ) .



Both problems mentioned above can be solved by the algorithm of spectral projected gradient for L1 minimization (SPGL1). Given a constraint τ , we can obtain the residual norm from Eq. 18:

$$\varphi(\tau) = \|\mathbf{r}_\tau\|_2 \text{ with } \mathbf{r}_\tau = \mathbf{L}\mathbf{m}_\tau - \mathbf{d}_{obs} \quad (19)$$

Let

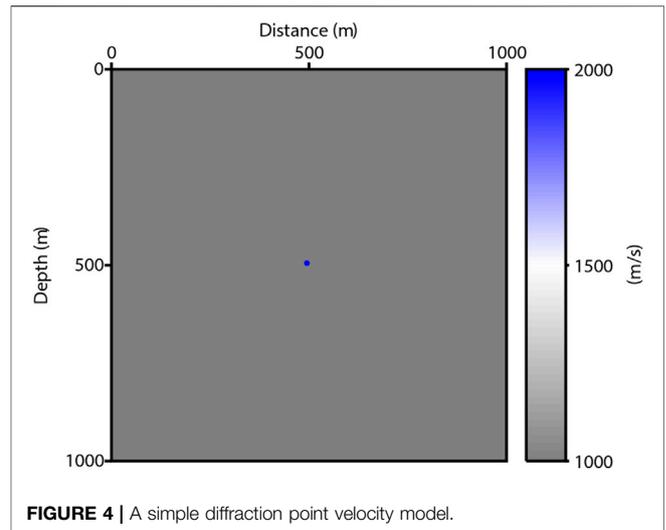
$$\varphi(\tau) = \sigma \quad (20)$$

Eq. 20 recasts (LS_τ) as a problem of finding the root of a nonlinear equation and defines a continuous curve, the Pareto curve (Figure 3).

For a given σ , SPGL1 uses the Newton method to approach the root, and as the τ updates iteratively, the optimal solution \mathbf{m}_{τ_σ} of problems (BP_σ) and (LS_τ) can be obtained. Therefore, we balance the 2-norm of the residual against the 1-norm of the solution eventually (van den Berg and Friedlander, 2009). From Figure 3, the question is degraded to a simple Lasso problem (LS_τ) when the noise level factor σ is equal to 0. In this study, we set $\sigma = 0$. Note that synthetic seismic records do not contain noise in general, so we set the noise level factor to be zero. Actually, the algorithm of SPGL1 can deal with noisy data, and we can add some random noise or set some traces to be zero in synthetic data. Besides, the determination of parameter τ is quite important. According to the theoretical model, we can calculate the perturbation model or reflectivity model and make a rough estimate of τ . In general, it is appropriate to set the value of tau to tens of times that of the calculated perturbation model or reflectivity model. Then, the parameter τ can be adjusted according to the inversion results.

Here we summarize the workflow of L1-regularized LSRTM as follows:

- 1) Obtain the predicted data \mathbf{d}_0 with migration velocity v_0 , and get the \mathbf{d}_{res} with $\mathbf{d}_{res} = \mathbf{d}_{obs} - \mathbf{d}_0$;
- 2) Set the initial model \mathbf{m}_0 and predict the data $\mathbf{L}\mathbf{m}_0$ based on Born or Kirchhoff approximation, therefore we can get the residual $\mathbf{r}_0 = \mathbf{L}\mathbf{m}_0 - \mathbf{d}_{res}$ and gradient operator $\mathbf{g}_0 = \mathbf{L}^T \mathbf{r}_0$;



- 3) Input the parameters of τ , σ and set $k = 0$;
- 4) Solve the Lasso problem (LS_τ) with the algorithm of SPGL1, and update the \mathbf{m}_k , \mathbf{r}_k and \mathbf{g}_k until $k = k_{max}$;
- 5) Output the result $\mathbf{m}_{k_{max}}$.

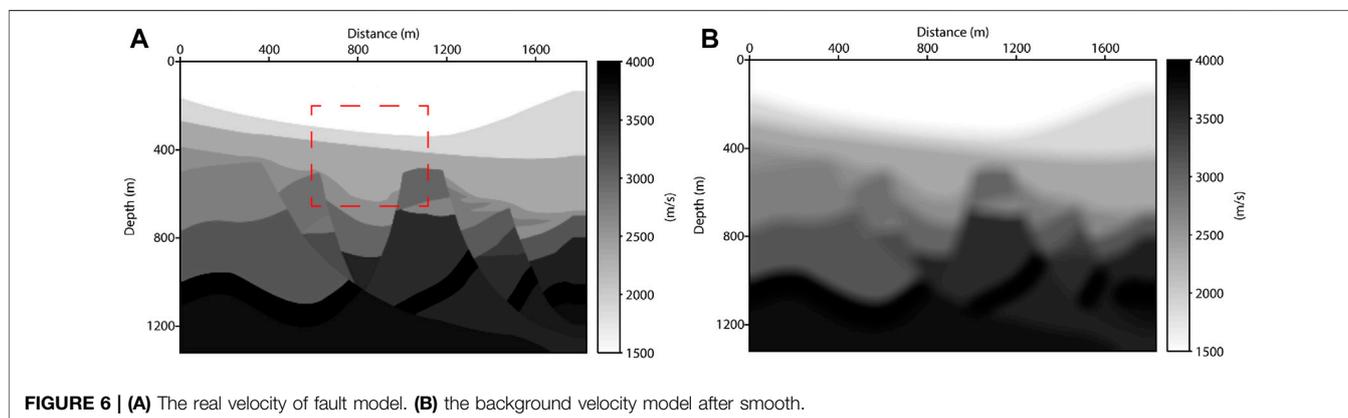
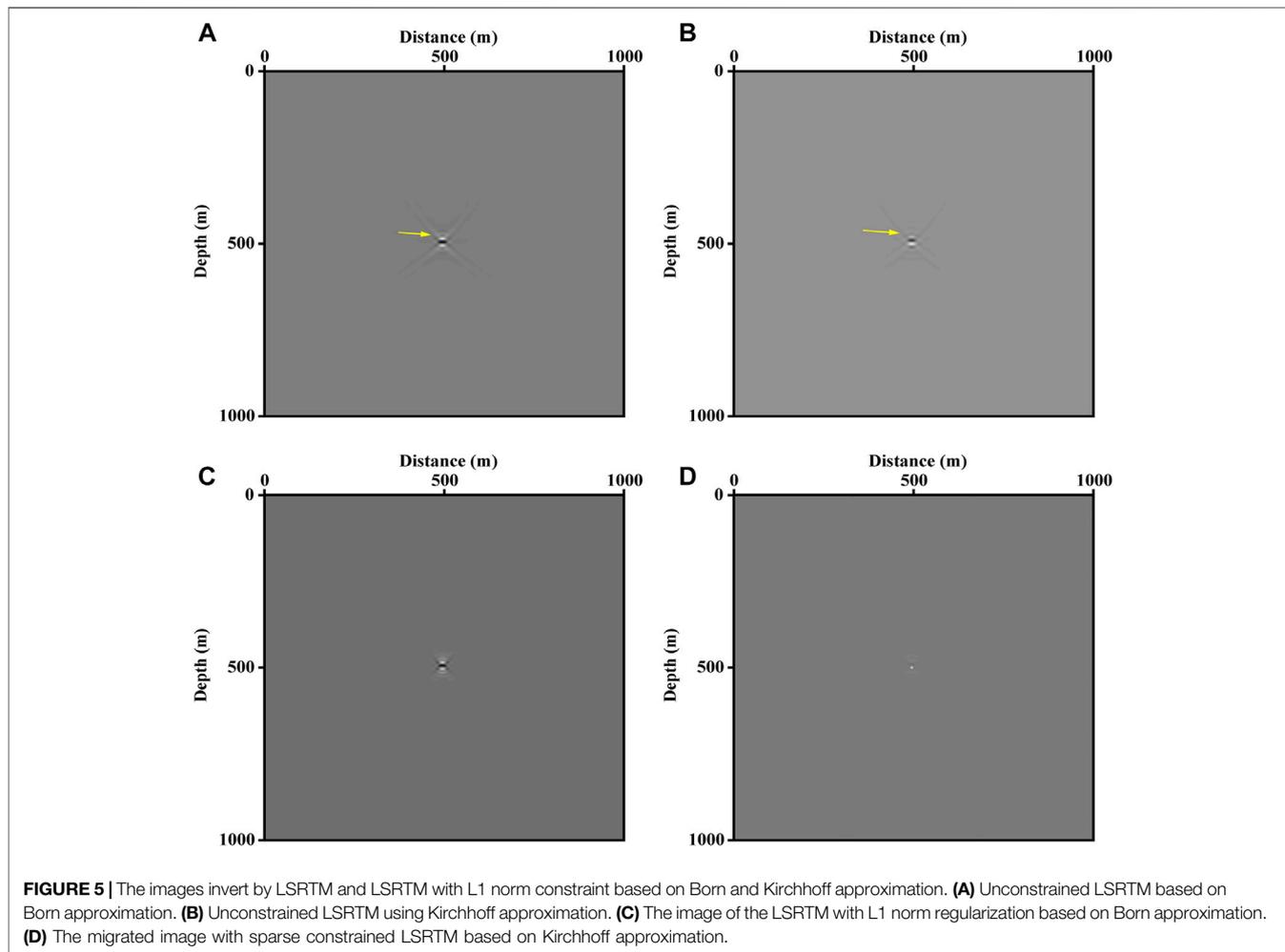
NUMERICAL EXAMPLE

In this study, two theoretical models are used to test the validity of the proposed method, including a single diffraction point and complex fault model. Both are based on the two-way acoustic wave equation. Here we use the finite difference method on regular grid.

Single-Diffraction Point

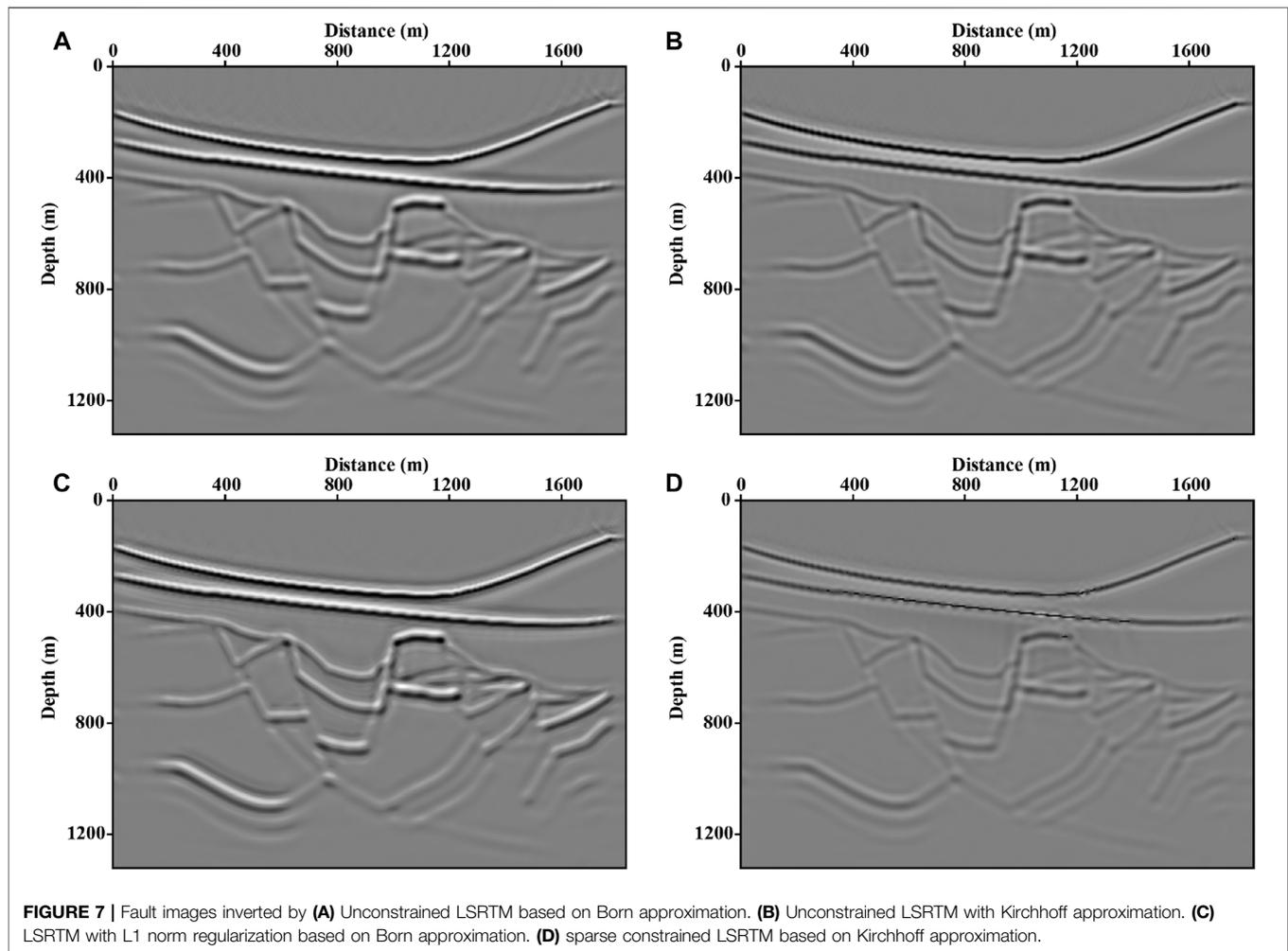
To verify the effectiveness of this method, we first set a simple model with a diffraction point of 2000 m/s embedded in the background velocity of 1,000 m/s (Figure 4), and the entire model has been discretized into 201×201 grids in the horizontal and vertical directions, respectively, with the same interval of 5m. The geometry system is arranged as follows: a total of 21 shots are uniformly distributed on the surface of this model with an interval of 50m. Geophones are also placed on each grid point on the surface. We use the Ricker wavelet with a center frequency of 25 Hz for modeling, and the sampling interval is 0.5 ms. In this example, we set 1,000 m/s as the migration velocity.

As shown in Figure 5A, the image is obtained by LSRTM based on Born approximation. This is consistent with the actual situation to a certain degree. The single scatter point is blurred with a disturbing cross pattern (marked by a yellow arrow). However, it should be a dot on the image (Lecomte and Kjeller, 2008). This is because we use the adjoint operator to migrate rather than the inverse operator in Eq. 14. Specifically, Eq. 13 defines a normal equation with $\mathbf{L}^T \mathbf{L}\mathbf{m} = \mathbf{L}^T \mathbf{d}$. The term $\mathbf{L}^T \mathbf{L}$, Hessian matrix, is equivalent to a blur operator acting on the true image \mathbf{m} . Furthermore, $\mathbf{L}^T \mathbf{L}$ includes the influence of irregular acquisition, limited acquisition aperture, band



limited source, etc., which generate artifacts and degrade the resolution of the image (Jiang and Zhang, 2019). Compared to LSRTM, the same method in **Figure 5C** with the L1 constraint performs better; it mitigates the distortion caused by the blur operator. Therefore, with the promotion of sparsity, the resolution in the least-squares method has been improved significantly, and the image looks more like a scatter point.

The LSRTM based on the Kirchhoff operator inverts the reflectivity directly from the seismic records. **Figure 5B** displays the image produced by Kirchhoff approximation. Compared to **Figure 5A**, least-squares RTM based on Kirchhoff operator suffers from the same problems. Similarly, we implement the L1 norm on LSRTM, which is shown in **Figure 5D**. The cross pattern is eliminated clearly, and we



obtain an explicit dot rather than a blurred spot. Therefore, for a simple model, the sparsity-promoting LSRTM based on Kirchhoff approximation can effectively improve the resolution of the image. The results calculated by LSRTM in **Figures 5A,B** iterate 5 times. **Figures 5C,D** use the SPGL1 algorithm for iterating 10 times with $\tau = 2$.

Fault Model

We also test the other relative complex model. In this fault model (**Figure 6A**), there are some classical geological structures, including folds, fault blocks, and depressions. Therefore, it appropriately shows the complex structure of near-surface media. The maximum and minimum velocities are 4,000 m/s and 1,500 m/s, respectively. Similarly, we discretize it into 265×367 grids with an interval of 5 m. Thus, a total of 25 shots are uniformly located at the surface of this model. The modeling seismic wavelet is same to last experiment.

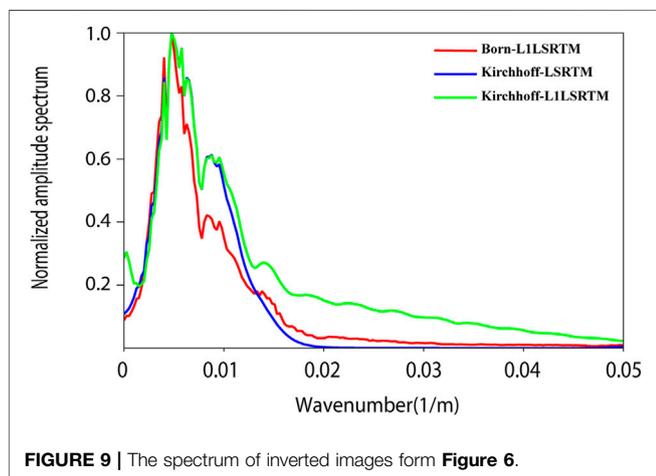
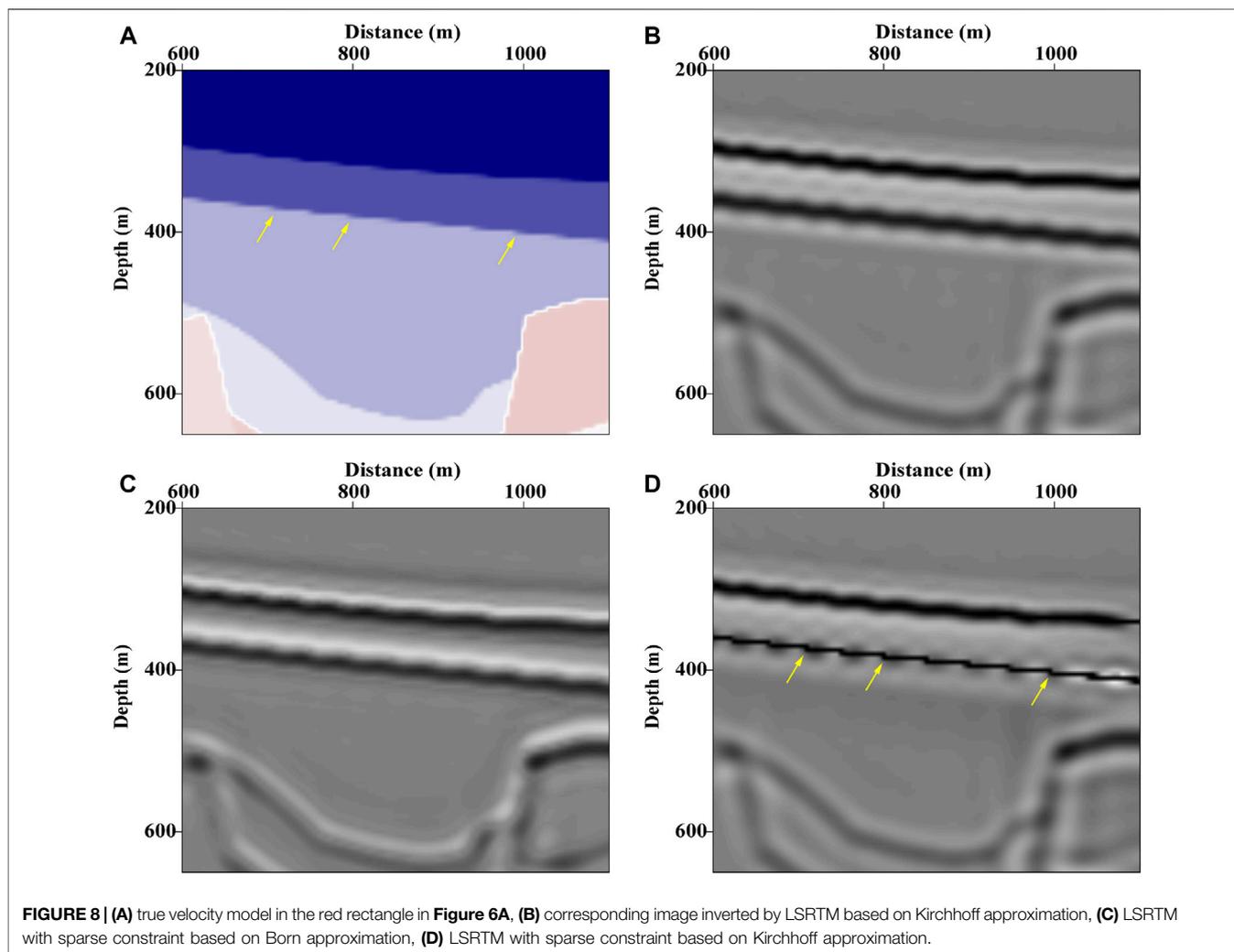
As shown in **Figures 7A,B**, images inverted by LSRTM based on two approximations fit the fault model well, and the contact relationship between structures can be clearly depicted. To further improve the resolution of these images, we combined L1 norm regularization with LSRTM to reconstruct the model. From **Figures 7C,D**, the method based on Kirchhoff approximation

recovers the stratum's sparsity more effectively. The results in **Figures 7A,B** are calculated by LSRTM with 10 iterations. **Figures 7C,D** use the SPGL1 algorithm for iterating 10 times with $\tau = 2000$ and $\tau = 500$, respectively. Note that the amplitude of inverted results is different because of the value of parameter τ .

Furthermore, we enlarge the model framed by red rectangle in **Figure 6A**, which has step-like strata (marked by yellow arrows in **Figure 8A**). After inverting, the reflectivity image in **Figure 8D** produced by constrained LSRTM based on Kirchhoff operator agrees with the actual situation.

Furthermore, images inverted by two different approximations have different phases. According to **Eq. 6**, reflectivity can be derived from model perturbation. With the assumption of a small incident angle, $\cos(\alpha, \alpha)$ is roughly equivalent to 1 and can be ignored. Then, **Eq. 6** can be rewritten as $r(\mathbf{x}) = ikm(\mathbf{x})/2$, where the wavenumber $k = \frac{\omega}{c}$. Therefore, reflectivity $r(\mathbf{x})$ is the spatial derivative of model perturbation $m(\mathbf{x})$. As a result, the image inverted by Kirchhoff performs sparser and sharper, and there is a phase shift of 90° between perturbation model and reflectivity model.

Figure 9 shows the amplitude spectra of the images in **Figures 7B–D**, respectively. The spectra curves are the sum of each trace by the spatial Fourier transform along the depth. The



red one is generated from the image inverted by L1-regularized LSRTM based on Born approximation. The blue and green spectrum curves are produced by unconstrained and constrained LSRTM with Kirchhoff approximation, respectively. Because of the spatial

derivative and sparse constraint, the spectrum of the image inverted by L1-LSRTM with Kirchhoff approximation has more high-wavenumber components than that of Born approximation, which explains that Kirchhoff approximation improves the resolution of the image.

CONCLUSION

The LSRTM recasts classical seismic inversion as a linear inverse problem. By means of linear approximation, physical model is related to the corresponding wavefield. Thereafter, we can reduce the residual between predicted and observed data iteratively to directly invert the interest parameters. This study introduces two linearization methods. Born approximation obtains the relationship between the model and physical response based on perturbation theory. With the help of the Born operator, we derive another type of linear method, namely the Kirchhoff operator, which relates the reflectivity to wavefield explicitly. Moreover, these two methods have a relationship of a spatial derivative, and there is a phase shift

between perturbation model and reflectivity model. Although two operators are different physical quantities, the resolution can be improved by Kirchhoff approximation.

LSRTM can mitigate the shortcomings of other migration methods, while the solution is smooth and deviates from the true model. Specifically, there are redundant oscillatory axes in the strata that should be sparsely distributed. Therefore, we reform the question as a sparsity-promoting LSRTM. The SPGL1 algorithm can effectively solve this problem and invert a sparse image that matches the model well. Examples prove the validity of our study.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

REFERENCES

- Baysal, E., Kosloff, D. D., and Sherwood, J. W. C. (1983). Reverse Time Migration. *Geophysics* 48, 1514–1524. doi:10.1190/1.1441434
- Beylkin, G. (1985). Imaging of Discontinuities in the Inverse Scattering Problem by Inversion of a Causal Generalized Radon Transform. *J. Math. Phys.* 26, 99–108. doi:10.1063/1.526755
- Bleistein, N., Cohen, J. K., and Hagin, F. G. (1987). Two and One-half Dimensional Born Inversion with an Arbitrary Reference. *Geophysics* 52, 26–36. doi:10.1190/1.1442238
- Bleistein, N. (1987). On the Imaging of Reflectors in the Earth. *Geophysics* 52, 931–942. doi:10.1190/1.1442363
- Dai, W., Fowler, P., and Schuster, G. T. (2012). Multi-source Least-Squares Reverse Time Migration. *Geophys. Prospecting* 60, 681–695. doi:10.1111/j.1365-2478.2012.01092.x
- Dutta, G., and Schuster, G. T. (2014). Attenuation Compensation for Least-Squares Reverse Time Migration Using the Viscoacoustic-Wave Equation. *Geophysics* 79, S251–S262. doi:10.1190/geo2013-0414.1
- Jiang, B., and Zhang, J. (2019). Least-squares Migration with a Blockwise Hessian Matrix: A Prestack Time-Migration Approach. *Geophysics* 84, R625–R640. doi:10.1190/geo2018-0533.1
- Lecomte, I. (2008). Resolution and Illumination Analyses in PSDM: A ray-based Approach. *The Leading Edge* 27, 650–663. doi:10.1190/1.2919584
- Liu, H. W., Li, B., Liu, H., Tong, X. -L., and Liu, Q. (2010). The Algorithm of High Order Finite Difference Pre-stack Reverse Time Migration and GPU Implementation. *Chin. J. Geophys.* 53, 1725–1733. doi:10.1002/cjg2.1530(in Chinese)
- Liu, Y., Chang, X., Jin, D., He, R., Sun, H., and Zheng, Y. (2011). Reverse Time Migration of Multiples for Subsalt Imaging. *Geophysics* 76, WB209–WB216. doi:10.1190/geo2010-0312.1
- Liu, Y., He, B., and Zheng, Y. (2020). Controlled-order Multiple Waveform Inversion. *Geophysics* 85, R243–R250. doi:10.1190/geo2019-0658.1
- Liu, Y., Liu, X., Osen, A., Shao, Y., Hu, H., and Zheng, Y. (2016). Least-squares Reverse Time Migration Using Controlled-Order Multiple Reflections. *Geophysics* 81, S347–S357. doi:10.1190/geo2015-0479.1
- McMechan, G. A. (1983). Migration by Extrapolation of Time-dependent Boundary Values*. *Geophys. Prospect* 31, 413–420. doi:10.1111/j.1365-2478.1983.tb01060.x
- Plessix, R.-E. (2006). A Review of the Adjoint-State Method for Computing the Gradient of a Functional with Geophysical Applications. *Geophys. J. Int.* 167, 495–503. doi:10.1111/j.1365-246X.2006.02978.x
- Schuster, G. T. (2017). *Seismic Inversion*. Tulsa: Society of Exploration Geophysicists. doi:10.1190/1.9781560803423
- Stolt, R. H., and Weglein, A. B. (2012). *Seismic Imaging and Inversion: Application of Linear Inverse Theory*. New York: Cambridge University Press, 1–404. doi:10.1017/CBO9781139056250
- Tarantola, A. (1984). Inversion of Seismic Reflection Data in the Acoustic Approximation. *Geophysics* 49, 1259–1266. doi:10.1190/1.1441754

AUTHOR CONTRIBUTIONS

XH-Q: Methodology, Investigation, Formal analysis, Visualization, Writing—Original Draft WX-Y: Conceptualization, Writing—Review and Editing, Project administration WC-Y: Validation, Data Curation, Writing—Review and Editing ZJ-J: Supervision, Writing—Review and Editing, Funding acquisition.

ACKNOWLEDGMENTS

We thank the National Natural Science Fund of China (under grant 42074158) for supporting this work. And we are grateful for the comments provided by reviewers, which improve the quality of this paper a lot.

- Tibshirani, R. (1996). Regression Shrinkage and Selection via the Lasso. *J. R. Stat. Soc. Ser. B (Methodological)* 58, 267–288. doi:10.1111/j.2517-6161.1996.tb02080.x
- van den Berg, E., and Friedlander, M. P. (2009). Probing the Pareto Frontier for Basis Pursuit Solutions. *SIAM J. Sci. Comput.* 31, 890–912. doi:10.1137/080714488
- van den Berg, E., and Friedlander, M. P. (2011). Sparse Optimization with Least-Squares Constraints. *SIAM J. Optim.* 21, 1201–1229. doi:10.1137/100785028
- Wang, X. Y., Zhang, J. J., Xu, H. Q., and Tian, B. Q. (2021). Least-squares Reverse Time Migration with Wavefield Decomposition Based on the Poynting Vector. *Chin. J. Geophys. (in Chinese)* 64, 645–655. doi:10.6038/cjg202100120
- Wang, Y. H. (2016). *Seismic Inversion: Theory and Applications*. Oxford: Wiley, 85–86. doi:10.1002/9781119258032.ch7
- Wu, D., Yao, G., Cao, J., and Wang, Y. (2016). Least-squares RTM with L1 Norm Regularisation. *J. Geophys. Eng.* 13, 666–673. doi:10.1088/1742-2132/13/5/666
- Yang, K., and Zhang, J. (2019). Comparison between Born and Kirchhoff Operators for Least-Squares Reverse Time Migration and the Constraint of the Propagation of the Background Wavefield. *Geophysics* 84, R725–R739. doi:10.1190/geo2018-0438.1
- Yoon, K., Shin, C., Suh, S., Lines, L. R., and Hong, S. (2003). 3D Reverse-Time Migration Using the Acoustic Wave Equation: An Experience with the SEG/EAGE Data Set. *The Leading Edge* 22, 38–41. doi:10.1190/1.1542754
- Zhang, Y., Duan, L., and Xie, Y. (2015). A Stable and Practical Implementation of Least-Squares Reverse Time Migration. *Geophysics* 80, V23–V31. doi:10.1190/geo2013-0461.1
- Zhu, J., and Lines, L. R. (1998). Comparison of Kirchhoff and Reverse-time Migration Methods with Applications to Prestack Depth Imaging of Complex Structures. *Geophysics* 63, 1166–1176. doi:10.1190/1.1444416

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Hong-Qiao, Xiao-Yi, Chen-Yuan and Jiang-Jie. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.