One-step data-domain least-squares reverse-time migration

Qiancheng Liu and Daniel Peter**

ABSTRACT

Least-squares reverse-time migration (LSRTM) is an iterative inversion algorithm for estimating the broadband-wavenumber reflectivity model. Although it produces superior results compared to conventional reverse-time migration (RTM), LSRTM is computationally expensive. Here, we introduce a one-step LSRTM method by considering the demigrated and observed data to design a deblurring preconditioner in the data domain using the Wiener filter. For the Wiener filtering, we further employ a stabilized division algorithm via Taylor expansion. The preconditioned observed data is then remigrated to obtain a deblurred image. The total cost of this method is about two RTMs. Through synthetic and real data experiments, we see that the one-step LSRTM is able to enhance image resolution and balance source illumination at low computational costs.

INTRODUCTION

Reverse-time migration (RTM) is a powerful tool in imaging complicated structures without dip limitation (Baysal et al., 1983; Whitmore, 1983; Zhang and Sun, 2009). However, RTM is defined as the adjoint operator of full wave-equation modeling and as a consequence the imaging quality suffers. The blurring effect is captured by the Hessian, also known as resolution function or point-spread function. Least-squares migration (LSM) approximates the pseudo-inverse operator in order to deblur the adjoint imaging step (Claerbout, 1992), where the LSM algorithm dates back to Lailly (1984). LSM was first applied to Kirchhoff migration (Schuster, 1993; Nemeth et al., 1999), then adapted to phase-shift migration (Kuehl and Sacchi, 2002; Clapp et al., 2005) and the one-way wave equation (Wang et al., 2005), and is implemented on top of RTM. Here, we focus on least-squares reverse-time migration (LSRTM).

Based on the two-way wave equation, LSRTM seeks for the broadband-wavenumber reflectivity model (Zeng et al., 2014) by minimizing the residual between predicted data (from Born modeling) and observed data via an iterative linearized inversion. Wong et al. (2011) use LSRTM to image ocean-bottom data. Dai et al. (2012) and Dai and Schuster (2013) improve the computational efficiency of LSRTM by source-encoding. Tan and Huang (2014) numerically illustrate that it is crucial to update source wavefields to image steep dips. Dutta and Schuster (2014) take Q-attention into consideration when performing LSRTM. Wong et al. (2015) jointly image reflection and free-surface-multiple data to increase the subsurface illumination. Zhang et al. (2015) introduce a practical LSRTM solution based on a cross-correlation objective function. Yao and Jakubowicz (2016) formulate LSRTM in a generalized matrix form. LSRTM indeed produces better images, however, even with additional techniques to improve the efficiency of LSRTM (Dai et al., 2012; Dai and Schuster, 2013), its computational burden remains high.

Preconditioning helps to make the LSM computationally cheaper and is usually applied either in the image or the data domain. In the image domain, several approximations to the Hessian (or inverse Hessian) are considered (Plessix and Mulder, 2004). Some researchers use a horizontally invariant approximation (Hu et al., 2001) or a diagonal approximation (Rickett, 2003). Guitton (2004) approximates the inverse Hessian with non-stationary matching filters. Lecomte (2008) approximates the Hessian using point-spread functions (PSFs). Tang (2009) explicitly computes the target-oriented PSFs using a phase-encoding technique to invert for the reflectivity. Aoki and Schuster (2009) use PSFs as well to design a deblurring filter that can be utilized in a regularization scheme to obtain acceptable results. Compared with the image-domain preconditioning, the data-domain approaches are more likely to be stationary and less costly. In the data domain, several methods like deconvolution (Robinson and Treitel, 1967; Peacock and Treitel, 1969) and automatic gain control (AGC) (Hunter et al., 1984; Rajagopalan and Milligan, 1994) are carried out. However, these methods are derived from a signal processing point of view.

In this paper, we propose a one-step data-domain LSRTM method using a Wiener filter. Unlike conventional data-domain methods, our preconditioner is derived from the wave-equation migration-demigration operator such that it becomes more suitable for LSRTM. The filtering is easy to carry out and its computational cost is negligible compared to that of the seismic modeling. Finally, the processed data is remigrated back to the image domain, yielding an imaging profile with minimized blurring effects. The total cost of our method is at two RTMs.

This paper first briefly reviews the method of conventional LSRTM. Then, we derive our one-step inversion theory in which the deblurring happens in the data domain using a Wiener filter. We design an appropriate division operation which enhances the filtering algorithm to be efficient and robust. Finally, we test our method using synthetic datasets and 2D marine data. Although our method is used in pre-stack LSRTM, it can potentially also be applied to general LSM and improve post-stack imaging profiles.

METHODS

Review of LSRTM

In this paper, we focus on acoustic wave-equation modeling. LSRTM mainly consists of three circular sections: (i) demigration (Born modeling), (ii) migration (RTM) and (iii) linearized optimization. The demigration and migration operators here are both based on the linearized two-way wave equation. Therefore, there is no dip limitation throughout this approach. All operators below are represented in matrix-vector notation.

Born modeling, a first-order approximation to the full wave equation, can be written as

$$\mathbf{L} \mathbf{m} \approx \mathbf{d}, \quad (1)$$

where $\mathbf{L}$ is the forward operator associated with a specific configuration of the survey geometry, the migration velocity and source wavelet; $\mathbf{d}$ is the observed data. In this paper we employ the acoustic wave equation with constant density and parameterize the logarithmic reflectivity as $\mathbf{m} = \delta v/v$ with δv being the velocity perturbation and v being the velocity model (Thierry et al., 1999; Operto et al., 2000; Tromp et al., 2005). In seismic migration,

only with data $\mathbf{d}$ and a migration velocity model at hand, one attempts to image subsurface structures. Due to an unbalanced illumination and a limited acquisition aperture, the resulting imaging profile is inferior compared to the true reflectivity model.

To overcome the imperfection of migration imaging and to gain an inverted reflectivity model of better quality, we adopt a least-squares inversion using the following misfit function in L2-norm:

$$E(\mathbf{m}) = \frac{1}{2} \|\mathbf{Lm} - \mathbf{d}\|_2^2, \quad (2)$$

the normal equation of which reads

$$\mathbf{L}^T \mathbf{Lm} = \mathbf{L}^T \mathbf{d}, \quad (3)$$

where $\mathbf{L}^T$ is the adjoint operator of the forward operator $\mathbf{L}$ (Plessix and Mulder, 2004; Virieux and Operto, 2009; Tromp et al., 2005). Conventionally, we have the RTM image

$$\mathbf{m}_{mig} = \mathbf{L}^T \mathbf{d} = \mathbf{L}^T \mathbf{Lm}, \quad (4)$$

from which we find that $\mathbf{L}^T \mathbf{L}$, the Hessian, acts as the blurring operator on $\mathbf{m}$ (as shown in Appendix A), resulting in a blurred image $\mathbf{m}_{mig}$. Unfortunately, the Hessian matrix for practical problem is prohibitive to compute or store. As a solution, conventional LSRTM usually chooses iterative inversions using a gradient optimization method, such as a steepest descent or a conjugate gradient method (Hestenes and Stiefel, 1952; Chavent, 1999; Nemeth et al., 1999), to deblur the migration file $\mathbf{m}_{mig}$ and to approach the reflectivity model. The computational cost of such a method depends on the number of iterations. Although Dai et al. (2012), Dai and Schuster (2013) improve the efficiency using a source-encoding technique, the cost of LSRTM remains high. For example, as suggested by Table 2 in Dai and Schuster (2013), the cost of plane-wave LSRTM is still at 12 RTMs.

Deblurring in the data domain

Equation 4 suggests that the RTM image can be directly deblurred if we find a good operator $(\mathbf{L}^T \mathbf{L})$ and get its inverse. Some studies propose approximations to $(\mathbf{L}^T \mathbf{L})^{-1}$, such as the local layered media assumption (Hu et al., 2001), diagonal approximation (Rickett, 2003), non-stationary matching filter (Guitton, 2004) and PSFs (Lecomte, 2008; Aoki and Schuster, 2009). All these methods are implemented in the image domain. Here on the contrary, we propose a method which can be carried out in the data domain.

First, we obtain the raw RTM image $\mathbf{m}_0$ with observed data $\mathbf{d}$

$$\mathbf{m}_0 = \mathbf{L}^T \mathbf{d}, \quad (5)$$

Second, we obtain the observed data $\mathbf{d}_1$ from the raw RTM image $\mathbf{m}_0$

$$\mathbf{d}_1 = \mathbf{Lm}_0. \quad (6)$$

The relationship between equations 5 and 6 is

$$\mathbf{d}_1 = \mathbf{L} \mathbf{L}^T \mathbf{d}. \quad (7)$$

We can now find a filter $\mathbf{F}$ which minimizes the difference between $\mathbf{d}_1$ and $\mathbf{d}$ in a least-squares sense

$$e(\mathbf{F}) = \frac{1}{2} \|\mathbf{F}\mathbf{d}_1 - \mathbf{d}\|_2^2. \quad (8)$$

The solution of $e(\mathbf{F})$ can be expressed approximately as

$$\mathbf{F} = (\mathbf{L}\mathbf{L}^T)^{-1}, \quad (9)$$

where $\mathbf{F}$ is a filter in the data domain. Note that $(\mathbf{L}\mathbf{L}^T)$ is the data resolution matrix (Tarantola, 1987). Here we simplify the inverse operation in equation 9 by using a diagonal approximation. This way, the method becomes applicable for practical problems at large scales. Finally, we precondition the observed data with the filter $\mathbf{F}$. We do not store the filter $\mathbf{F}$ itself but directly filter the observed data to get a new dataset $\mathbf{F}\mathbf{d}$. The remigrated RTM image $\mathbf{m}'$ by $\mathbf{F}\mathbf{d}$ reads

$$\mathbf{m}' = \mathbf{L}^T(\mathbf{F}\mathbf{d}) \approx \mathbf{L}^T(\mathbf{L}\mathbf{L}^T)^{-1}\mathbf{d} = \mathbf{L}^T(\mathbf{L}^T)^{-1}\mathbf{L}^{-1}\mathbf{L}\mathbf{m} = \mathbf{m}, \quad (10)$$

from which we have $\mathbf{m}' \approx \mathbf{m}$. That is, the remigrated RTM image $\mathbf{m}'$ approximates the true reflectivity model $\mathbf{m}$.

The total computational cost of this method is basically at 7 modeling runs. Three of them exist in the RTM by equation 5 if we exploit the boundary source-wavefield-reconstruction strategy. Two of them are required by the Born modeling of equation 6. The remaining two are present in the RTM by equation 10 because the source wavefield has been saved on the boundary, ready for the source wavefield reconstruction during the cross-correlation imaging operation.

Wiener filtering using a stabilized division via Taylor expansion

We implement the Wiener filtering in the frequency domain. The solution of equation 8 can be written as

$$\mathbf{F}(\omega) = \frac{\mathbf{d}(\omega) \mathbf{d}_1^*(\omega)}{\mathbf{d}_1(\omega) \mathbf{d}_1^*(\omega) + \epsilon} = \frac{\mathbf{d}(\omega) \mathbf{d}_1^*(\omega)}{|\mathbf{d}_1(\omega)|^2 + \epsilon}, \quad (11)$$

where $\mathbf{F}$ is the Wiener solution; ω denotes the angular frequency, $\mathbf{d}$ and $\mathbf{d}_1$ denote the observed and predicted data, respectively; $*$ represents the conjugate transpose, ϵ is a stabilizer to avoid division by zero. In the following, we will drop the explicit frequency dependency on ω of the Wiener filter in our notation. The division is carried out trace-by-trace because of the diagonal approximation in equation 9. The data preconditioned by the Wiener filter can be transformed back to the time domain. Although we use only a diagonal approximation, equation 11 has the ability to flatten the spectrum of the data as $\mathbf{d}_1$ has been cross-correlated by two more band-limited source wavelets within $\mathbf{L}$ and $\mathbf{L}^T$. We remove one cross-correlated source wavelet from $\mathbf{d}_1$ using the method in (Liu et al., 2016) such that only one more cross-correlated source wavelet remains. It follows that equation 11 coincides with the concept of spectrum flattening. We bandpass equation 11 with Hann windows to avoid boosting the high and low ends of the data spectrum.

Equation 11 is sensitive to the value of ϵ as it replaces small values on the denominator. To avoid the introduction of errors by ϵ , we could optionally consider a low-cut division, in

which the stabilization is achieved by zeroing $\mathbf{F}$ in case the denominator is smaller than ϵ . However, the low-cut division approach risks losing valuable information.

To alleviate the zero-division risk, we use a Taylor expansion to turn the division problem into a geometric series (Liu et al., 2016)(as proved in Appendix B):

$$\mathbf{F} = \begin{cases} \frac{\mathbf{d}\mathbf{d}_1^*}{|\mathbf{d}_1|^2 + \epsilon} & \text{if } |\mathbf{d}_1|^2 > \epsilon \\ \frac{\mathbf{d}\mathbf{d}_1^*}{\epsilon} \sum_{i=0}^n \left(1 - \frac{|\mathbf{d}_1|^2}{\epsilon}\right)^i & \text{if } |\mathbf{d}_1|^2 \leq \epsilon \end{cases} \quad (12)$$

where $\epsilon = \frac{\alpha}{n\omega}(\sum_{n\omega} |\mathbf{d}_1(\omega)|^2)$ with $(0 < \alpha < 1)$ is an adaptive stabilizer related to the average of the squared amplitude spectrum of $\mathbf{d}_1$ over all effective frequency points per trace, and n denotes the order of the Taylor expansion. To demonstrate the advantage of the stabilized division in equation 12 over equation 11 and a low-cut division, we look at a toy problem of signal deconvolution and reconstruction.

Consider the simple computation of $\tilde{\mathbf{w}} = (\mathbf{w}/|\mathbf{w}|) \cdot |\mathbf{w}|$ where $\mathbf{w}$ denotes the spectrum of a toy seismic trace indicated by the dashed line in Figure 1 for reference. $|\mathbf{w}|$ is its amplitude spectrum. We first divide $\mathbf{w}$ by $|\mathbf{w}|$, and then multiply $|\mathbf{w}|$ back. Obviously, a good numerical division algorithm should ensure that $\tilde{\mathbf{w}}$ and $\mathbf{w}$ are identical. For a clear demonstration, we set up a relatively high stabilization factor: $\epsilon = \frac{0.5}{n\omega}(\sum_{n\omega} |\mathbf{w}(\omega)|)$, where $n\omega$ denotes the number of frequencies. Figure 1 shows four results of $\tilde{\mathbf{w}}$ (solid lines) in the time domain for comparison, with the true $\mathbf{w}$ (dashed lines) for reference. Figures 1a and 1b are obtained by equation 7 and the low-cut division algorithm, respectively. Figures 1c and 1d are obtained by our stabilized division algorithm in equation 10, using $n = 2$ and $n = 6$, respectively. The comparison shows that the proposed division algorithm here provides a better, more stable result.

NUMERICAL EXAMPLES

In this section, we will test our method on the Marmousi model and a 2D marine data.

Marmousi benchmark

In the following, we test our approach on the 2D acoustic Marmousi model (Versteeg and Grau, 1990). The wave equation is computed by a staggered-grid, finite-difference stencil at $O(dt^2, dx^4)$ (Virieux, 1986) surrounded by convolutional perfectly matched layers (CPML) (Komatitsch and Martin, 2007). The observed data is generated from the true velocity model in Figure 2a. For this benchmark, 40 shots are excited evenly at an interval of 200 m. The source function is a 20-Hz Ricker wavelet. The observed data is acquired with a fixed-spreading acquisition geometry. We have 501 receivers at a 10 m interval for each shot gather. The recording duration is 4 s. Direct waves are muted. The migration velocity in Figure 2b is produced by smoothing the true velocity with a Gaussian smoother at a 400-m window length. High-wavenumber features are blurred. The reflectivity model in Figure 2c is taken for reference.

The observed data for the shot point at 4000 m is shown in Figure 3a. We first obtain a pre-stack RTM image as shown in Figure 4a. Then, the demigrated data, shown in Figure 3b, is generated by Born modeling. A Wiener filter is estimated in the frequency domain

Figure 1: Comparisons of $\tilde{\mathbf{w}} = (\mathbf{w}/|\mathbf{w}|) \cdot |\mathbf{w}|$ using three different division algorithms. Here $\mathbf{w}$ is the spectrum of the Ricker wavelet, $|\mathbf{w}|$ is its amplitude spectrum. In the time domain, the true and reconstructed wavelets are indicated by the solid and dashed lines, respectively. (a) and (b) correspond to the division algorithms by equation 7 and the low-cut form, respectively. Both (c) and (d) are the resulting trace by our stabilized division algorithm, with $n = 2$ and $n = 6$, respectively.

Figure 2: Marmousi benchmark: (a) true velocity used to generate the observed data, (b) migration velocity, (c) reference reflectivity model. We consider the logarithmic reflectivity $\delta v/v$ with δv being the velocity perturbation between (a) and (b), and v being the migration velocity in (b).

by comparing the observed and demigrated data trace-by-trace, in which our stabilized division is deployed. The stabilizer ϵ in equation 12 for this dataset is chosen with $\alpha = 0.01$ and $n = 6$. Note that what we get here is an approximation to $(\mathbf{L}\mathbf{L}^T)^{-1}$ rather than $(\mathbf{L}^T\mathbf{L})^{-1}$. We use the estimated band-limited filter $(\mathbf{L}\mathbf{L}^T)^{-1}$ to precondition the observed data. The result is shown in Figure 3c.

Remigrating the preconditioned data back to the image domain, a deblurred RTM image is obtained. Comparing Figure 4a with 4b, we observe that the image in Figure 4b obtained by the one-step LSRTM has a higher resolution and reveals more balanced amplitudes over the conventional RTM. Still, our method like others continues to suffer from sparse migration footprints which could be due to the off-diagonal components of the data resolution matrix. The Wiener filtering is implemented trace-by-trace, so its memory burden is only determined by the discrete sampling length of one trace. Additionally, the filtering algorithm can easily be run in parallel. The demigrated data from Figure 4b is shown in Figure 3d, which looks closer to Figure 3a than Figure 3b.

The conventional stacked RTM image (after source illumination) and the stacked deblurred RTM image are shown in Figures 5a and 5b, respectively. As illustrated in Figure 5b, the result by the one-step LSRTM is closer to the true reflectivity model in Figure 2c. Since the Wiener-filter estimation is much cheaper than one seismic modeling operation, the total cost of our method is about two RTMs.

To investigate in more detail the quantitative nature of our inverted result, we extract three well logs across the reservoir at $x = 5.7$ km in Figures 2c, 5a and 5b. We compare these vertical profiles in Figure 6a. The true reflectivity (plotted in black color) is taken as reference. After calibrating the RTM and one-step LSRTM curves, we observe that the one-step LSRTM curve (in red color) fits better the true reflectivity curve than the conventional RTM one (in blue color). We also compare the three well logs in the wavenumber domain in Figure 6b. Again, the one-step LSRTM result gets closer to the true model profile and has a broader spectrum than the conventional RTM profile.

2D Marine data

To further investigate the proposed one-step LSRTM approach, we test it on a real 2D marine dataset. The data is acquired by a single cable with 460 receiver groups at an interval of 12.5 m. The offset ranges from 100 m to 5750 m. There are 240 shot gathers at an interval of 37.5 m. The record length is 8 s with a sampling rate of 2 ms. The dominant frequency of this dataset is around 10 Hz. We also properly estimate the source wavelet and migration velocity. Figure 7 shows the migration velocity model used for this study.

In our method the wavelet deconvolution or amplitude compensation is unnecessary even for real data, because the deblurring filter will account for it. Thus, the workflow is similar to the one in the Marmousi benchmark above. In this example, we still correct for the half-integral filter caused by the 2D Green's function as a convention. Unlike iterative LSRTM, the convergence of the one-step LSRTM is insensitive to the cycle-skipping caused by slight phase-rotation mismatches in the source wavelet. The reason is that in our approach, traces from observed and demigrated data are matched as a whole. This however cannot gradually push out some noises, such as for example the residual multiples in the observed data, when compared to conventional iterative LSRTM.

Figure 3: Common-shot gathers in one-step LSRTM: (a) observed data, (b) demigrated data from the RTM image. Preconditioned data is shown in (c), using the Wiener filter from comparing (a) and (b). The preconditioned data in (c) is used for remigration, and (d) shows the demigrated data from the one-step LSRTM image. We find that (d) is closer to (a) than (b).

Figure 4: Single-shot pre-stack Marmousi RTM images: (a) raw RTM image, (b) deblurred RTM image after remigration.

Figure 5: Stacked Marmousi images by (a) conventional RTM and (b) one-step LSRTM.

Figure 6: Reflectivity comparison of well logs extracted at $x = 5.7$ km: (a) profile from true reflectivity model in Figure 2c (black), conventional RTM image in Figure 5a (blue) and one-step LSRTM image in Figure 5b (red); (b) the corresponding wavenumber spectra. The well log of one-step LSRTM matches the true one better than conventional RTM. We directly take the true reflectivity for reference, although strictly speaking, the reflectivity model should be converted to time-domain, band-pass filtered, and then converted back to depth-domain.

Figure 7: Migration velocity for 2D marine data

Figures 8a, 8b and 8c show the observed, demigrated and filtered data of the 100th common-shot-gather, respectively. The stabilizer ϵ in equation 12 for this dataset is chosen with $\alpha = 0.05$ and $n = 6$. Figures 9a and 9b show the original stacked RTM image (after source illumination) and our deblurred RTM image, respectively. We observe that Figure 9b shows a higher imaging resolution, but also more migration artifacts at deeper parts. These artifacts arise from the boosting of noises at the bottom of Figure 8c. We may further employ a regularization term to dip-filter these noises. In this example, we suppress the unwanted noise as following: (i) multiply each single-shot inverted image with its corresponding source illumination, (ii) stack them together, and (iii) normalize the stacked image with the stacked source illumination (Guitton et al., 2007). The remedy result is shown in Figure 9c. For further demonstration, we pick out three zoom-in sections between CDP number [100,800] and depth [1.25 km, 3.125 km] to show in Figure 10. We see that both Figures 10b and 10c have sharper and more balanced reflectors. The reflectors in Figure 10c using the remedied noise suppression workflow, become clearer than those in 10b.

Figure 8: The 2D marine datasets: (a) observed data, (b) demigrated data and (c) pre-conditioned data with Wiener filtering for remigration.

CONCLUSIONS

LSRTM is a powerful technique for improving the imaging quality of RTM using an iterative linearized inversion. However, it remains computationally costly. In this study, we presented a one-step data-domain LSRTM method by comparing the observed and demigrated data using a Wiener filter. The filtering was solved in the frequency domain using a stabilized division algorithm. The additional computational cost and memory burden of the data-domain filtering are much lower than those of the seismic modeling. The total computational cost of the approach presented here is found to be at two RTMs. Numerical results show that our method provides promising results with inverted images providing higher resolution and more balanced amplitudes. For the numerical benchmark however we used noise-free data. Some remedies have been implemented for a real marine dataset, still future work will involve the development of an improved cross-correlation objective function to handle noisy

Figure 9: Stack 2D Marine images by (a) conventional RTM, (b) one-step LSRTM and (c) the remedied one-step LSRTM from (b). We suppress the additional migration artifacts in (b) by multiplying each single-shot one-step LSRTM image with its corresponding source illumination, then stacking the processed images, and finally normalizing the stacked image with the stacked source illumination.

Figure 10: Zoomed-in views between CDP number [100,800] and depth [1.25 km, 3.125 km] from Figure 9. Compared to (a) conventional RTM, both (b) one-step LSRTM and (c) remedied one-step LSRTM have sharper and more balanced reflectors, with those in (c) being less noisy than those in (b).

data. Furthermore, we plan to incorporate source-encoding techniques into our method to further enhance it for very large datasets. Nevertheless, the one-step LSRTM approach presented here works in the data domain and can readily be applied to large-scale problems.

ACKNOWLEDGEMENTS

We are grateful to editors Stephane Operto, Arthur Chen, and reviewers Mandy Wong, Linan Xu and an anonymous reviewer for improving the initial manuscript. The research reported in this publication was supported by funding from King Abdullah University of Science and Technology (KAUST). For computer time, this research used the resources of the Information Technology Division and Extreme Computing Research Center (ECRC) at KAUST.

REFERENCES

- Aoki, N., and G. Schuster, 2009, Fast least-squares migration with a deblurring filter: *Geophysics*, **74**, WCA83–WCA93.
- Baysal, E., D. D. Kosloff, and J. W. Sherwood, 1983, Reverse time migration: *Geophysics*, **48**, 1514–1524.
- Chavent, G. and Plessix, R. E., 1999, An optimal true-amplitude least-squares prestack depth-migration operator: *Geophysics*, **64**, 508–515.
- Claerbout, J. F., 1992, *Earth soundings analysis: Processing versus inversion*: Blackwell Science.
- Clapp, M. L., R. G. Clapp, and B. L. Biondi, 2005, Regularized least-squares inversion for 3-d subsalt imaging: 75th Annual International Meeting, SEG, Expanded Abstracts, 1814–1817.
- Dai, W., P. Fowler, and G. Schuster, 2012, Multi-source least-squares reverse time migration: *Geophys. Prospect.*, **60**, 681–695.
- Dai, W., and G. T. Schuster, 2013, Plane-wave least-squares reverse-time migration: *Geophysics*, **78**, S165–S177.
- Dutta, G., and G. T. Schuster, 2014, Attenuation compensation for least-squares reverse time migration using the viscoacoustic-wave equation: *Geophysics*, **79**, S251–S262.
- Guitton, A., 2004, Amplitude and kinematic corrections of migrated images for nonunitary imaging operators: *Geophysics*, **69**, 1017–1024.
- Guitton, A., A. Valenciano, D. Bevc, and J. Claerbout, 2007, Smoothing imaging condition for shot-profile migration: *Geophysics*, **72**, S149–S154.
- Hestenes, M. R., and E. Stiefel, 1952, Methods of conjugate gradients for solving linear systems: **49**.
- Hu, J., G. Schuster, and P. A. Valasek, 2001, Poststack migration deconvolution: *Geophysics*, **66**, 939 – 952.
- Hunter, J., S. Pullan, R. Burns, R. Gagne, and R. Good, 1984, Shallow seismic reflection mapping of the overburden-bedrock interface with the engineering seismographs some simple techniques: *Geophysics*, **49**, 1381–1385.
- Komatitsch, D., and R. Martin, 2007, An unsplit convolutional perfectly matched layer improved at grazing incidence for the seismic wave equation: *Geophysics*, **72**, SM155–SM167.

- Kuehl, H., and M. Sacchi, 2002, Robust avp estimation using least-squares wave-equation migration: 72nd Annual International Meeting, SEG, Expanded Abstracts, 281 – 284.
- Lailly, P., 1984, The seismic inverse problem as a sequence of before stack migrations: Conference on Inverse Scattering.
- Lecomte, I., 2008, Resolution and illumination analyses in psdm: A ray-based approach: The Leading Edge, **27**, 650–663.
- Liu, Q., J. Zhang, and H. Zhang, 2016, Eliminating the redundant source effects from the cross-correlation reverse-time migration using a modified stablized division: Computers Geosciences, **92**, 49–57.
- Nemeth, T., C. Wu, and G. T. Schuster, 1999, Least-squares migration of incomplete reflection data: Geophysics, **61**, 208 – 221.
- Operto, S., S. Xu, and G. Lambar'e, 2000, Can we image quantitatively complex models with rays?: Geophysics, **65**, 1223–1238.
- Peacock, K., and S. Treitel, 1969, Predictive deconvolution: Theory and practice: Geophysics, **34**, 155–169.
- Plessix, R., 2006, A review of the adjoint-state method for computing the gradient of a functional with geophysical applications: Geophys. J. Int., **167**, 495–503.
- Plessix, R. E., and W. A. Mulder, 2004, Frequency-domain finite-difference amplitude-preserving migration: Geophys. J. Int., **157**, 975–987.
- Rajagopalan, S., and P. Milligan, 1994, Image enhancement of aeromagnetic data using automatic gain control: Exploration Geophysics, **25**, 173–178.
- Rickett, J. E., 2003, Illumination-based normalization for wave-equation depth migration: Geophysics, **68**, 1371–1379.
- Robinson, E. A., and S. Treitel, 1967, Principles of digital wiener filtering: Geophysical Prospecting, **15**, 311–332.
- Schuster, G. T., 1993, Least-squares cross-well migration: 63rd Annual International Meeting, SEG, Expanded Abstracts, 110–113.
- Schuster, G. T., and J. Hu, 2000, Greens function for migration: continuous recording geometry: Geophysics, **65**, 167–175.
- Tan, S., and L. Huang, 2014, Least-squares reverse-time migration with a wavefield-separation imaging condition and updated source wavefields: Geophysics, **79**, S195–S205.
- Tang, Y., 2009, Target-oriented wave-equation least-squares migration/inversion with phase-encoded hessian: Geophysics, **74**, WCA95–WCA107.
- Tarantola, A., 1987, verse problem theory: Methods for data fitting and model parameter estimation: Elsevier Science Publ. Co., Inc.
- Thierry, P., S. Operto, and G. Lambar'e, 1999, Fast 2D ray-Born inversion/migration in complex media: Geophysics, **64**, 162–181.
- Tromp, J., C. Tape, and Q. Liu, 2005, Seismic tomography, adjoint methods, time reversal, and banana-donut kernels: Geophys. J. Int., **160**, 195–216.
- Versteeg, R., and G. Grau, 1990, Practical aspects of seismic data inversion, the marmousi experience: Eur. Assoc. Expl. Geophys., Proceedings of 1990 EAEG Workshop, 52nd EAEG Meeting, 1–194.
- Virieux, A., and S. Operto, 2009, An overview of full-waveform inversion in exploration geophysics: Geophysics, **74**, WCC1–WCC26.
- Virieux, J., 1986, P-SV wave propagation in heterogeneous media: velocity-stress finite-difference method: Geophysics, **51**, 889–901.
- Wang, J., H. Kuehl, and M. D. Sacchi, 2005, High-resolution wave-equation ava imaging: Algorithm and tests with a data set from the western canadian sedimentary basin: Geo-

- physics, **70**, S91–S99.
- Whitmore, N., 1983, Iterative depth migration by backward time propagation: 382–385.
- Wong, M., B. Biondi, and S. Ronen, 2015, Imaging with primaries and free-surface multiples by joint least-squares reverse time migration: *Geophysics*, **80**, S223–S235.
- Wong, M., S. Ronen, and B. Biondi, 2011, Least-squares reverse time migration/inversion for ocean bottom data: a case study: SEG Technical Program Expanded Abstracts, 2369–2373.
- Yao, G., and H. Jakubowicz, 2016, Least-squares reverse-time migration in a matrix-based formulation: *Geophys. Prospect.*, **64**, 611–621.
- Zeng, C., S. Dong, J. Mao, and B. Wang, 2014, Broadband least-squares reverse time migration for complex structure imaging: 3715–3719.
- Zhang, Y., L. Duan, and Y. Xie, 2015, A stable and practical implementation of least-squares reverse time migration: *Geophysics*, **80**, V23–V31.
- Zhang, Y., and J. Sun, 2009, Practical issues of reverse time migration: True amplitude gathers, noise removal and harmonic-source encoding: 204–204.

APPENDIX

Appendix A: Adjoint formulation for migrated image

According to Schuster and Hu (2000), the solution of the acoustic wave equation can be represented with Green’s function in a summation form. Using the concept of exploding reflectors (Claerbout, 1992) in the frequency domain, observed data $d(\mathbf{x}_r, \omega; \mathbf{x}_s)$ from the reflectivity distribution model $m(\mathbf{x})$ in Born modeling can be written as

$$d(\mathbf{x}_r, \omega; \mathbf{x}_s) = \int \omega^2 G_0(\mathbf{x}, \omega; \mathbf{x}_s) G_1(\mathbf{x}_r, \omega; \mathbf{x}) m(\mathbf{x}) S(\omega) d\mathbf{x}, \quad (\text{A-1})$$

where $\mathbf{x}_r$ and $\mathbf{x}_s$ denotes the receiver and source locations, respectively; ω denotes the angular frequency, and $G_0(\mathbf{x}, \omega; \mathbf{x}_s)$ and $G_1(\mathbf{x}_r, \omega; \mathbf{x})$ describe propagations of the zero-order source wavefield (background source wavefield) from $\mathbf{x}_s$ to $\mathbf{x}$ and the first-order source wavefield (perturbed wavefield) from $\mathbf{x}$ to $\mathbf{x}_r$, respectively; $S(\omega)$ denotes the frequency spectrum of the source wavelet. The Born modeling in equation A-1 is simplified in matrix form as, namely, $\mathbf{Lm} = \mathbf{d}$.

Using the adjoint-state method (Claerbout, 1992; Plessix, 2006), the RTM image can be obtained by

$$m_{mig}(\mathbf{x}) = \int \omega^2 G_s(\mathbf{x}, \omega; \mathbf{x}_s) S(\omega) [G_r(\mathbf{x}, \omega; \mathbf{x}_r) d(\mathbf{x}_r, \omega; \mathbf{x}_s)]^* d\omega, \quad (\text{A-2})$$

where $G_s(\mathbf{x}, \omega; \mathbf{x}_s)$ and $G_r(\mathbf{x}, \omega; \mathbf{x}_r)^*$ describe the forward propagated source wavefield from $\mathbf{x}_s$ to $\mathbf{x}$ and the backward propagated receiver wavefield from $\mathbf{x}_r$ to $\mathbf{x}$, respectively.

Substituting equation A-1 into equation A-2 yields

$$m_{mig}(\mathbf{x}) = \int \omega^2 m(\mathbf{x}') G_s(\mathbf{x}, \omega; \mathbf{x}_s) S(\omega) [G_r(\mathbf{x}, \omega; \mathbf{x}_r) G_0(\mathbf{x}', \omega; \mathbf{x}_s) G_1(\mathbf{x}_r, \omega; \mathbf{x}') S(\omega)]^* d\omega, \quad (\text{A-3})$$

in which the kernel

$$H(\mathbf{x}; \mathbf{x}') = \int \omega^2 G_s(\mathbf{x}, \omega; \mathbf{x}_s) |S(\omega)|^2 [G_r(\mathbf{x}, \omega; \mathbf{x}_r) G_0(\mathbf{x}', \omega; \mathbf{x}_s) G_1(\mathbf{x}_r, \omega; \mathbf{x}')]^* d\omega, \quad (\text{A-4})$$

denotes the Hessian, or point-spread function (PSF), from $\mathbf{x}'$ to $\mathbf{x}$. In matrix form, the adjoint process in equation A-3 can be expressed as $\mathbf{m}_{mig} = \mathbf{L}^T \mathbf{d} = \mathbf{L}^T \mathbf{L} \mathbf{m}$, with the Hessian being $\mathbf{H} = \mathbf{L}^T \mathbf{L}$.

Appendix B: Stabilized division algorithm based on a Taylor expansion

Let us begin with a very simple example,

$$f(x) = \frac{1}{x}, \quad (0 < x < \epsilon) \quad (\text{A-5})$$

where ϵ is a small, positive number. The Taylor expansion of equation B.1 at $x = \epsilon$ reads

$$f(x) = \frac{1}{\epsilon} \frac{1}{x/\epsilon} = \frac{1}{\epsilon} \frac{1}{1 - (1 - x/\epsilon)} = \frac{1}{\epsilon} \sum_{i=0}^n (1 - \frac{x}{\epsilon})^i, \quad (0 < x < \epsilon) \quad (\text{A-6})$$

where n denotes the expansion order we employ. This approximation converges to the correct value in $0 < x < \epsilon$, where Taylor expansions with higher degrees have better approximations.

Similarly, for equation 10 within $0 < |\mathbf{d}_1|^2 < \epsilon$, a stabilized division can be written as

$$\mathbf{F} = \frac{\mathbf{d} \mathbf{d}_1^*}{\epsilon \frac{|\mathbf{d}_1|^2}{\epsilon}} \approx \frac{\mathbf{d} \mathbf{d}_1^*}{\epsilon} \sum_{i=0}^n (1 - \frac{|\mathbf{d}_1|^2}{\epsilon})^i. \quad (\text{A-7})$$

As a result, the division algorithm can be implemented in a practical and robust way.