AN EFFICIENT WAVEFIELD SIMULATION AND RECONSTRUCTION METHOD FOR LEAST-SQUARES REVERSE TIME MIGRATION

XUAN KE, YING SHI and WEIHONG WANG

[1] *Earth Science College of Northeast Petroleum University, Daqing Heilongjiang 163318, P.R. China. sskgyth@sina.com*
[2] *Science and Technology Innovation Team on Fault Deformation, Sealing and Fluid Migration, Daqing Heilongjiang 163318, P.R. China.*

ABSTRACT

Ke, X., Shi, Y. and Wang, W.H., 2018. An efficient wavefield simulation and reconstruction method for least-squares reverse time migration. *Journal of Seismic Exploration*, 27: 183-200.

The least-squares reverse time migration (LSRTM) could be seen as an iterative processing of the reverse time migration (RTM). Therefore, a large storage requirement and accurate wavefield reconstruction scheme are also needed by LSRTM when the cross-correlated imaging condition is adopted. We propose a scheme that can reconstruct the wavefield backward accurately in the time domain. The variable order finite difference is introduced near the boundary region, which could reduce the storage requirement needed for the wavefield reconstruction. When the wavefield storage is greatly reduced, the proposed accurate wavefield reconstruction method and the precision of the reconstructed wavefield are feasible for LSRTM. We compare two wavefield reconstruction strategies based on the acoustic wave equation, then give a theoretical comparison of the data storage and computation consumption for the two strategies. At last, we perform a numerical experiment of LSRTM to prove the feasibility of our scheme.

KEY WORDS: least-squares reverse time migration, LSRTM, wavefield reconstruction, variable order finite difference, storage and computation consumption.

INTRODUCTION

Reverse time migration (RTM) has been a widely used seismic imaging method for complex structures. Its imaging condition always requires the cross-correlation calculation of the source and receiver wavefields (Claerbout, 1971). But the cross-correlation imaging condition are under high-frequency assumption, Weglein (2011a,b) proposed a predicted source and receiver experiment imaging condition, which could avoid the high-frequency assumption and be used in RTM (Liu, 2014; Weglein, 2016a,b). Zou (2017) provide a wedge resolution comparison between RTM (with Claerbout II imaging condition) and the first migration method (with Stolt extended Clarebout III imaging condition) with both conventional and broadband data. Zou's paper (Zou, 2017) demonstrates that the Stolt extended Claerbout III result has better resolution than the RTM result due to the smaller side lobes in the image from each reflector. The application of the imaging conditions above require that either the source or receiver wavefield must be stored or reconstructed. The simplest method is to store all the wavefield data, but the large data storage is almost rigorous in practice. Symes (2007) presented an optimal checkpointing scheme, which sets some checkpoints to achieve a small recomputation ratio, but this method needs to store many intermediate states of the wavefield. Clapp (2009) presented a random boundary condition, which allows the historical source wavefield to be obtained by time reversal computing, but this method also brings in some random noise. With the development of seismic data processing methods, RTM has been also used as a core operator in the iterative process, such as in least-squares reverse time migration (LSRTM) and full waveform inversion (FWI). This requires a wavefield reconstruction strategy with low storage consumption and high accuracy. Clapp (2008) presented a boundary saving method, which could reconstruct the wavefield accurately with a relatively small amount of storage, but the storage consumption depends on the number of time samples and finite difference orders. In practical usage, it is not feasible enough. Yang (2014) presented a graphic processing unit (GPU) implementation about Clapp's (2008) method and used it for RTM. Feng (2012) presented a single layer boundary saving strategy, reconstructing the wavefield by changing the finite difference order near the boundaries. This method effectively reduces the wavefield storage, but because the wavefield simulation operator is not matched completely with the reconstruction operator, it would also generate some noise. This makes the reconstructing precision fine for RTM, but not high enough for LSRTM and FWI.

In this paper, we first review LSRTM and then present a wavefield simulation and reconstruction strategy with variable order finite difference for LSRTM. The wavefield simulation and reconstruction operator could be matched completely, making the reconstructed wavefield accurate enough for some iterative processing based on RTM, such as LSRTM. Some noise

produced in the algorithm is analyzed in numerical examples. Two wavefield reconstruction strategies were tested with constant density acoustic wave equations. The theoretical data storage consumption and computation are analyzed in numerical examples. At last we achieve LSRTM with our wavefield reconstruction approach.

THEORY

Least-squares reverse time migration

In this section, we review LSRTM based on the constant density acoustic wave equation,

$$\frac{1}{v^2(\mathbf{x})}\frac{\partial^2 p(\mathbf{x},t;\mathbf{x}_s)}{\partial t^2} - \nabla^2 p(\mathbf{x},t;\mathbf{x}_s) = f(\mathbf{x},t;\mathbf{x}_s) \quad , \tag{1}$$

where $v(\mathbf{x})$ is the velocity model, $p(\mathbf{x}, t; \mathbf{x}_s)$ is the pressure field at position $\mathbf{x}$, and $f(\mathbf{x}, t; \mathbf{x}_s)$ is the source function.

Applying the Born approximation, we could derive the forward Born modeling operator,

$$\begin{cases} \dfrac{1}{v_b^2(\mathbf{x})}\dfrac{\partial^2 p_b(\mathbf{x},t;\mathbf{x}_s)}{\partial t^2} - \nabla^2 p_b(\mathbf{x},t;\mathbf{x}_s) = f(\mathbf{x},t;\mathbf{x}_s) \\[2mm] \dfrac{1}{v_b^2(\mathbf{x})}\dfrac{\partial^2 p_s(\mathbf{x},t;\mathbf{x}_s)}{\partial t^2} - \nabla^2 p_s(\mathbf{x},t;\mathbf{x}_s) = m(\mathbf{x})\dfrac{1}{v_b^2(\mathbf{x})}\dfrac{\partial^2 p_b(\mathbf{x},t;\mathbf{x}_s)}{\partial t^2} \end{cases} \quad , \tag{2}$$

where v_b is the background velocity, p_b is the background wavefield, ps denotes the perturbation wavefield, and $\mathbf{m}(\mathbf{x})$ is the model parameter of LSRTM. The detail of the derivation is presented in Appendix A. Generally, the forward Born modeling operator could be compactly expressed as a matrix-vector notation $dcal = \mathbf{L}\mathbf{m}$, where $dcal$ is the reflection data vector recorded from ps and L is the forward Born modeling operator that linearly relates $\mathbf{m}$ and $dcal$ (Dai, 2013).

Based on the adjoint state method (Plessix, 2006), the adjoint operator of the forward Born modeling operator could be expressed as follows,

$$\begin{cases} \dfrac{1}{v_b^2(\mathbf{x})}\dfrac{\partial^2 p_b(\mathbf{x},t;\mathbf{x}_s)}{\partial t^2} - \nabla^2 p_b(\mathbf{x},t;\mathbf{x}_s) = f(\mathbf{x},t;\mathbf{x}_s) \\[2mm] \dfrac{1}{v_b^2(\mathbf{x})}\dfrac{\partial^2 p_r(\mathbf{x},t;\mathbf{x}_s)}{\partial t^2} - \nabla^2 p_r(\mathbf{x},t;\mathbf{x}_s) = d_{cal}(\mathbf{x}_g,t;\mathbf{x}_s) \\[2mm] \mathbf{m}(\mathbf{x};\mathbf{x}_s) = \displaystyle\sum_t \dfrac{1}{v_b^2(\mathbf{x})}\dfrac{\partial^2 p_b(\mathbf{x},t;\mathbf{x}_s)}{\partial t^2}\, p_r(\mathbf{x},t;\mathbf{x}_s) \end{cases} \qquad (3)$$

Eq. (3) could also be expressed compactly as $\mathbf{m} = \boldsymbol{L}^T d_{cal}$, with the model parameter $\mathbf{m}$ indicating the migration image and $\boldsymbol{L}^T$ denoting the adjoint operator of the forward Born modeling operator.

For conventional LSRTM, the model parameter $\mathbf{m}$ is sought to minimize the misfit functional,

$$E(\mathbf{m}) = \frac{1}{2}\left\| \boldsymbol{L}\mathbf{m} - d_{obs} \right\|^2 . \qquad (4)$$

Eq. (4) could be solved by optimization algorithms. Clearbout (1992) proposed an iterative process of conjugate gradient method as follows

$$\begin{aligned} &\mathbf{r} \leftarrow \boldsymbol{L}\mathbf{m} - d_{obs} \\ &\text{iterate} \\ &\{ \\ &\quad \Delta\mathbf{m} \leftarrow \boldsymbol{L}^T\mathbf{r} \\ &\quad \Delta\mathbf{r} \leftarrow \boldsymbol{L}\Delta\mathbf{m} \\ &\quad \alpha = \frac{\left\langle\langle S_k \cdot S_k\rangle \cdot \langle \mathbf{r}\cdot\Delta\mathbf{r}\rangle\right\rangle - \left\langle\langle S_k \cdot \Delta\mathbf{r}\rangle \cdot \langle S_k \cdot \mathbf{r}\rangle\right\rangle}{\left\langle\langle \Delta\mathbf{r}\cdot\Delta\mathbf{r}\rangle \cdot \langle S_k \cdot S_k\rangle\right\rangle - \left\langle\langle S_k \cdot \Delta\mathbf{r}\rangle \cdot \langle S_k \cdot \Delta\mathbf{r}\rangle\right\rangle} \\ &\quad \beta = \frac{\left\langle\langle S_k \cdot \mathbf{r}\rangle \cdot \langle \Delta\mathbf{r}\cdot\Delta\mathbf{r}\rangle\right\rangle - \left\langle\langle \mathbf{r}\cdot\Delta\mathbf{r}\rangle \cdot \langle S_k \cdot \Delta\mathbf{r}\rangle\right\rangle}{\left\langle\langle \Delta\mathbf{r}\cdot\Delta\mathbf{r}\rangle \cdot \langle S_k \cdot S_k\rangle\right\rangle - \left\langle\langle S_k \cdot \Delta\mathbf{r}\rangle \cdot \langle S_k \cdot \Delta\mathbf{r}\rangle\right\rangle} \\ &\quad s_{k+1} = \alpha\Delta\mathbf{x} + \beta s_k \\ &\quad S_{k+1} = \alpha\Delta\mathbf{r} + \beta S_k \\ &\quad \mathbf{m} \leftarrow \Delta\mathbf{m} + s_{k+1} \\ &\quad \mathbf{r} \leftarrow \Delta\mathbf{r} - S_{k+1} \\ &\} \end{aligned}$$

where r is the data residual, $\Delta\mathbf{m}$ is the gradient, $\Delta\mathbf{r}$ is the conjugate gradient,

α and β are two scalars used to construct the update steps, and s_k and S_k are the update steps in the model and residual data domain for the k-th iteration.

When we perform operator L^T, it is similar to RTM, which needs large wavefield storage to guarantee the accuracy of the wavefield reconstruction. Additionally, because LSRTM can be regarded as a RTM iterative process, the data storage enhancement and the I/O (Input/Output) efficiency of the algorithm have more significant meaning.

Wavefield simulation and reconstruction based on variable order finite difference

If the cross-correlation imaging condition was employed in LSRTM or RTM, the source and receiver wavefields should be correlated at zero-offset in the time domain. However, because the two kinds of wavefields have opposite directions, we must reconstruct one of them in reverse. Feng's (2011) scheme could reconstruct the source wavefield with a relatively small amount of storage, but the reconstructed wavefields still have some residuals. The residual is mainly caused by two reasons. The first one is the change of finite difference orders would introduce a boundary effect. Because the boundary effect happens near the boundary location, its occurrence for the left, right, and bottom boundaries is relatively late. This could be solved by expanding the scope of the calculation appropriately, but the effect of the top boundary would introduce some residuals that are difficult to remove. The second reason is that the finite difference operators for the wavefield simulation and reconstruction are different. This means the wavefield cannot be reconstructed perfectly and the residual between simulation and reconstruction wavefield exist all over the wavefield.

In this paper, we propose that a variable orders finite difference operator, which is used for both wavefield simulation and reconstruction. This can ensure that no unwanted residual is introduced during the wavefield reconstruction. At the top boundary, we adapt the same strategy as Clapp (2008), which avoids the boundary effect due to the difference order variation at the top boundary. In fact, the wavefield simulation requires the addition of certain boundary conditions, so the wavefield region we have just discussed is the non-absorbing region within the absorbing region.

When using a variable order finite difference scheme, we obtain the discrete form of eq. (1) by ignoring the source item:

$$p_{m,n}^{k+1} = 2p_{m,n}^{k} - p_{m,n}^{k-1} + \frac{v^2 \Delta t^2}{\Delta x^2} \sum_{i=-M}^{M} C_i^M p_{m+i,n}^{k} + \frac{v^2 \Delta t^2}{\Delta z^2} \sum_{j=-N}^{N} C_j^N p_{m,n+j}^{k} \,, \quad (5)$$

where M and N are the half finite difference orders on the horizontal and vertical, respectively. C_i^M and C_j^N are the corresponding differential coefficients.

It is worth noting that M and N are not global constants. In most regions, we choose a high order finite difference scheme to simulate the wavefield, for example $M = N = 5$. As the simulating point approaches the boundaries, we change the finite difference scheme from high order to second order ($M = N = 1$) gradually. In addition, M and N are not necessarily equal, they just reduce gradually in their respective directions near the boundary. This ensures that when the difference order of one direction changes, the differential precision of the calculated point in the other direction is not affected. Fig. 1 shows the scheme of the variable order finite difference.
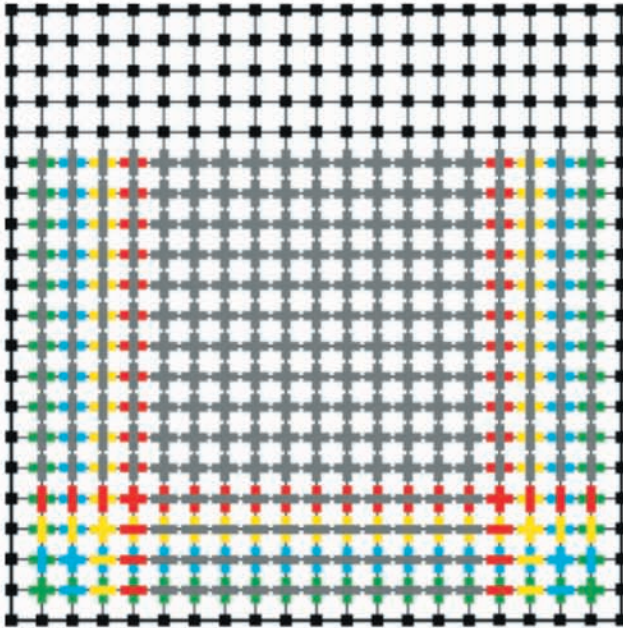
Fig. 1. A sketch of the finite difference orders adopted by the wavefield simulation and reconstruction. The black squares represent the stored layer of wavefields near the boundary. The thick solid colored lines represent the orders and direction of the finite difference. Two, 4, 6, 8 and 10 orders of finite difference are denoted by green, blue, yellow, red, and gray, respectively.
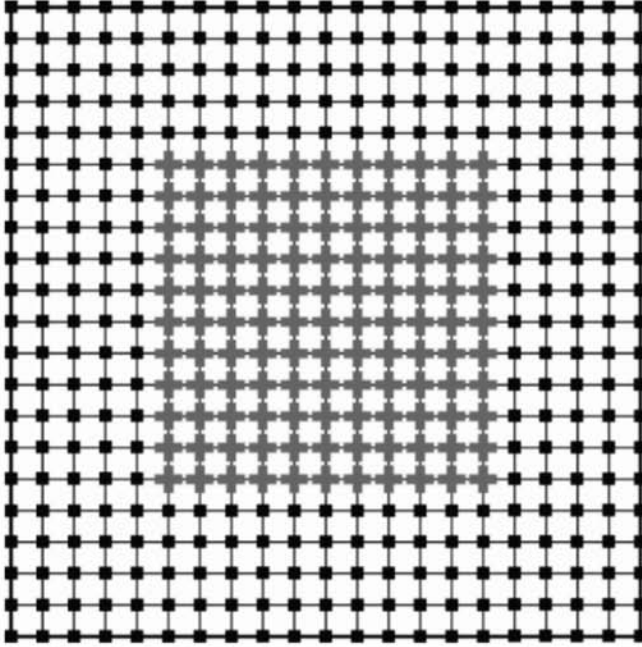
Fig. 2. A sketch of the simulation about the accurate source wavefield reconstruction scheme. The black squares represent the stored layer of wavefields near the boundary. The thick solid gray lines represent the grids, which need to be implemented by the high order finite difference.

Fig. 2 shows a sketch of the simulation of Clapp's (2008) method. Comparing Fig. 1 and Fig. 2, we see that the key ideas of our method are that we adopt the variable orders of finite difference to avoid the multi-layer storage of the wavefield at the left, right, and bottom boundaries; the multiple layers of the wavefield at the top boundary are saved to keep the accuracy of the wavefield from being effected by the changing finite difference order.

Accurate reconstruction of the wavefield using our approach can be achieved by storing the wavefield and the inverse operation of the variable order finite difference operator. This process ensures that the residual in the wavefield calculation is minimized and that no additional precision loss is introduced during the LSRTM iteration.

EXAMPLES

Wavefield simulation based on variable order finite difference

In this paper, we perform different orders of finite difference in both the horizontal and vertical directions where necessary, which are marked by different colors in Fig. 1 so that we can guarantee the wavefield simulation accuracy as much as possible.

Isotropic media of constant velocity are used for the numerical test. The velocity is 3500 m/s, the distance between grids is 15 m, and the time sampling interval is 1 ms. A Ricker wavelet with a 40 Hz dominant frequency is set as the source at the location (1 km,1 km). Fig. 3 shows that when the difference order is constant in the horizontal direction, the frequency dispersion of the wavefield is suppressed with the increase of the vertical difference order, and the overall accuracy is improved. This also explains why we use a higher order finite difference as much as possible in the different directions near the boundary where the algorithm allows.
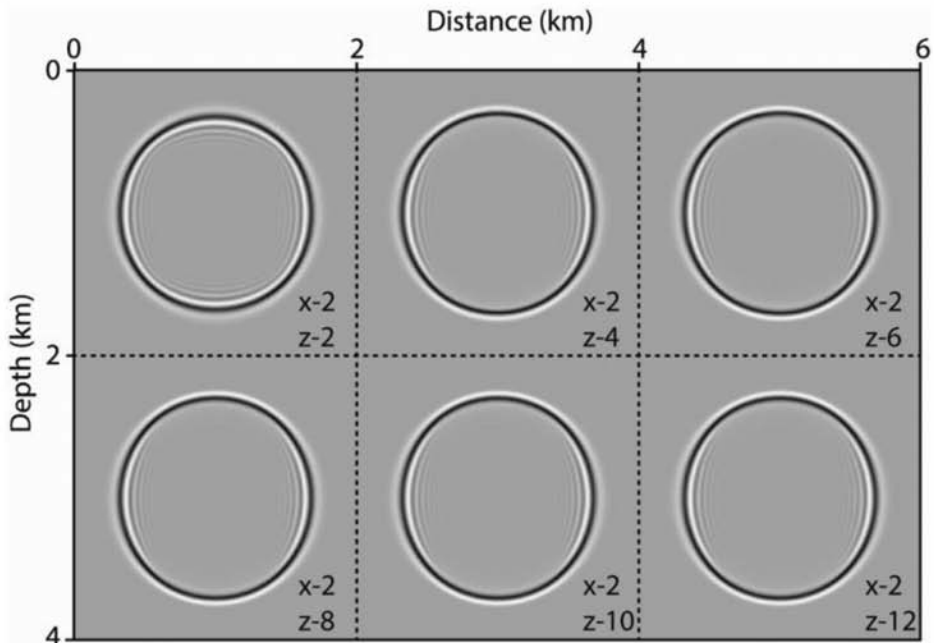
Fig. 3. Snapshots at 0.2 s, the order of the finite difference on horizontal direction is 2, the orders on vertical direction are 2, 4, 6, 8, 10, and 12 in turn.

In the previous section, we discussed that some reflections will be introduced when the difference order changes along a certain direction. Therefore, we have performed the following numerical experiments on this phenomenon. We adopt the same base model parameters as the test above and a 12-order finite difference is performed in most of the model, but the difference order is gradually reduced from 12 to 2 from the 1.5 km depth, which is marked the by red dotted line in Fig. 4a. Fig. 4b shows the amplitude extracted from the data marked by the white dotted line in Fig. 4a. From Fig. 4b, we see that, in the case where the difference coefficient changes, some energy reflections do occur, but the reflected energy is very small relative to the effective wavefield energy.
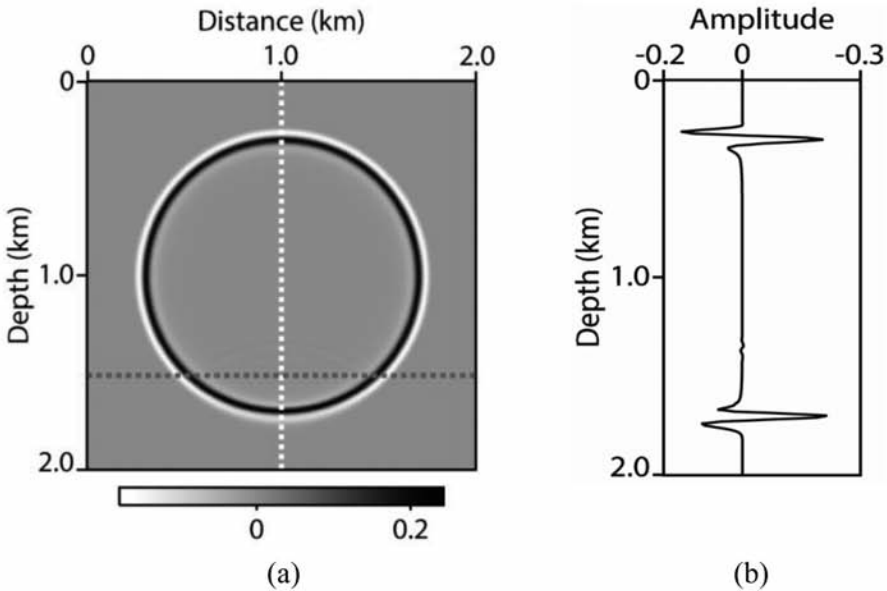
Fig. 4. A sketch of the reflection caused by the change of the difference order. (a) Snapshot with false reflections at 0.2 s, the change of difference order happens at 1.5 km depth, marked by the horizontal dotted line. (b) The amplitude of the data marked by the white dotted line in Fig. 4a.

Source wavefield reconstruction

The Partial Marmousi2 model showed in Fig. 5 is used for further numerical tests. We use a grid size of 401 in the horizontal direction and 287

in the vertical direction with a spacing of 15 m. We choose the Ricker wavelet as the source at 4.5 km on the ground. Fig. 6 shows the forward and reconstructed source wavefield at 1.8 s using our method. Fig. 7a shows the difference of the wavefield shown in Fig. 6. Fig. 7b shows the difference between the forward and reconstructed wavefield at 1.8 s, where the source wavefield is reconstructed by Clapp's (2008) method, which is known as the accurate source wavefield reconstruction scheme. We see that the magnitude of the residuals from our method is within the float number precision and the distribution is random, so the wavefield reconstruction method we proposed could be considered accurate.
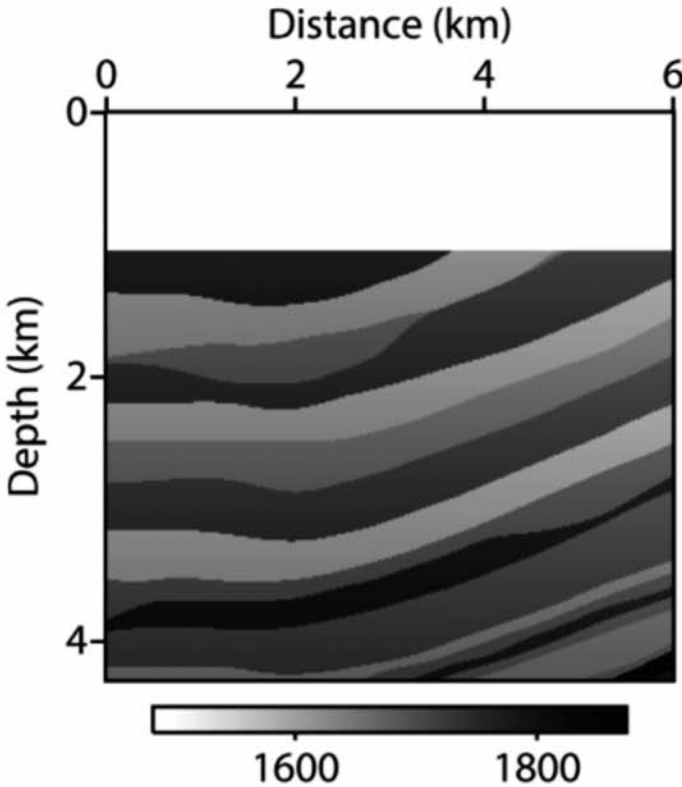
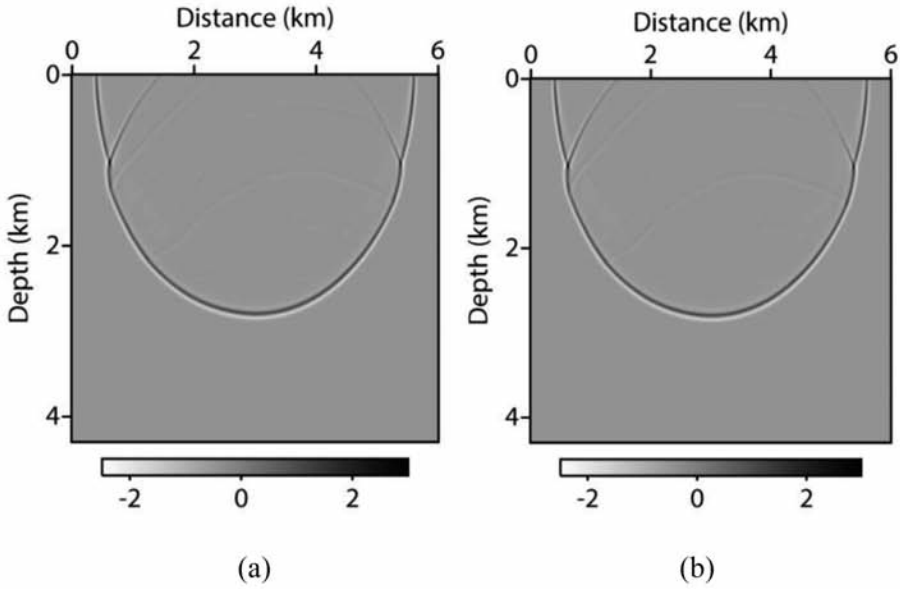Fig. 5. The velocity of the partial Marmousi2 model.

Fig. 6. Snapshot of the source wavefield at 1.8 s. (a) The forward wavefield. (b) The reconstructed wavefield.
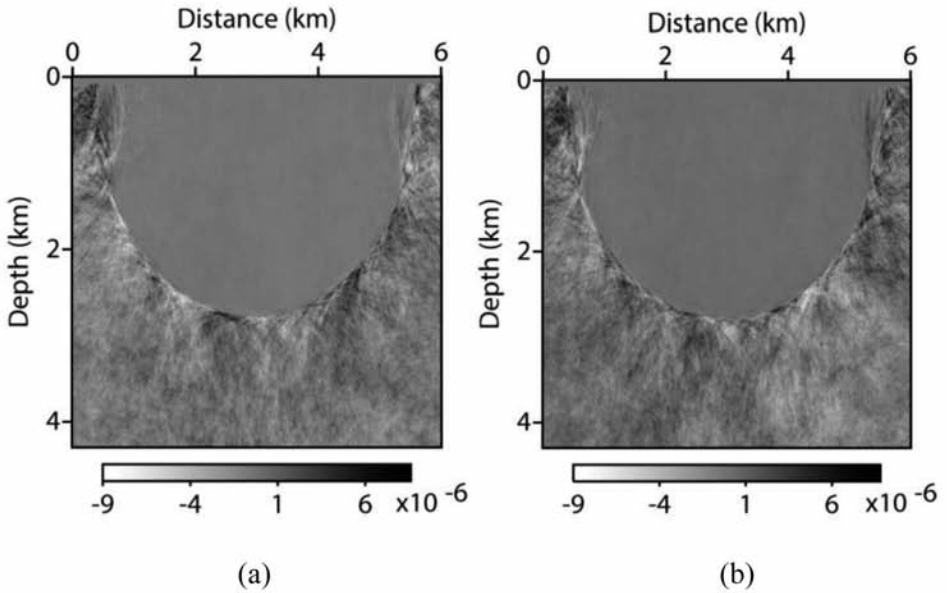

Fig. 7. The difference between the forward and reconstructed source wavefields at 1.8 s. (a) The scheme in this paper. (b) The accurate wavefield reconstruction scheme.

Comparison of storage and computation consumption

The exact storage and computation performance will vary with the size of the model, the scheme adopted, and so on. Therefore, we theoretically compare the storage and computation consumption of the two strategies as follows:

Strategy 1: The accurate wavefield reconstruction scheme (store the $d/2$ layers of the wavefield near the boundary), d is the finite difference order.

Strategy 2: Our scheme, the maximum finite difference order is d, the minimum order is 2.

For simplicity, we assume that the height and width of the simulation domain are both N. Considering that LSRTM is an iterative process and we need to store the wavefield data in the implementation of eq. (3), which we use as an example to compare the data storage and computation consumption. We define the computation consumption of the wavefield simulation from 0 to the maximum sampling time as the unit 1. The adoption of the absorbing boundary condition will expand the simulation region of the wavefield and the wavefield reconstruction's computation does not include the boundary zone, which is why the two strategies' computation consumptions are both less than 2. However, because at the left, right, and bottom boundaries, only a single layer of wavefield is saved in strategy 2, the wavefield grid, which needs to be reconstructed, is larger than in strategy 1 (there are more thick solid lines in Fig. 1 than in Fig. 2). This means that the computation consumption of strategy 2 is a little larger than strategy 1.

We summarize the computation comparison and data storage in Table 1, by using $N = 1000$, $d = 10$, and the time steps $N_S = 5000$.

Table 1. Comparison of the computation consumption and data storage for the source wavefield's backward propagation based on a two dimensional (2D) isotropic acoustic wave equation. We set the propagation of the wavefield in N_S time steps as a unit of computation consumption.

Strategy	Computation Consumption	Data storage
1	$<$ strategy 2 < 2	$\{[N^2 - (N - d)^2] \times N_s - 2(N - d)^2\} \times 4(\text{Byte}) \approx 387$ MB
2	< 2	$\{[N^2 - (N - d/2 - 1)(N - 2)] \times N_s - 2(N - d/2 - 1)(N - 2)\} \times 4(\text{Byte}) \approx 160$ MB

Least-squares reverse time migration

We use the same model defined in section 3.2 for the LSRTM test. Twenty shots are distributed from 0 to 6000 m on the ground, and the source interval is 300 m. One hundred ninety-nine receivers are evenly distributed beside each source, the maximum offset is 1485 m, and the trace interval is 15 m. We perform 5200 time steps for each shot, the time interval is 1 ms.

Fig. 8 shows the smoothed velocity model, which is involved in the calculation of LSRTM as background velocity. Fig. 9a denotes the model parameter **m**, and it could be used as the theoretical value to evaluate the effect of LSRTM. The result of LSRTM after 30 iterations is shown in Fig. 9b. The amplitude comparison between the theoretical and calculated value, shown in Fig. 10, could demonstrate that the LSRTM algorithm based on our approach works well. Fig. 11 reveals that the energy ratio of the data residual and original data decreases gradually as the number of iterations increases.
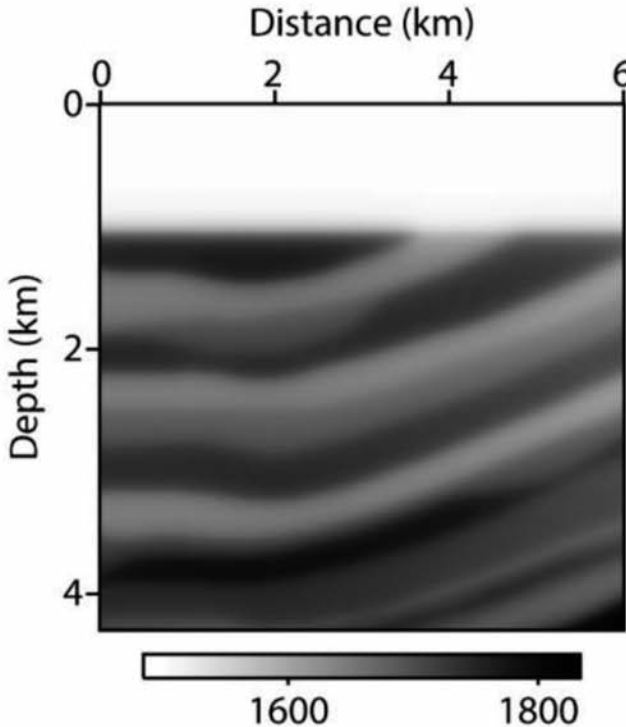

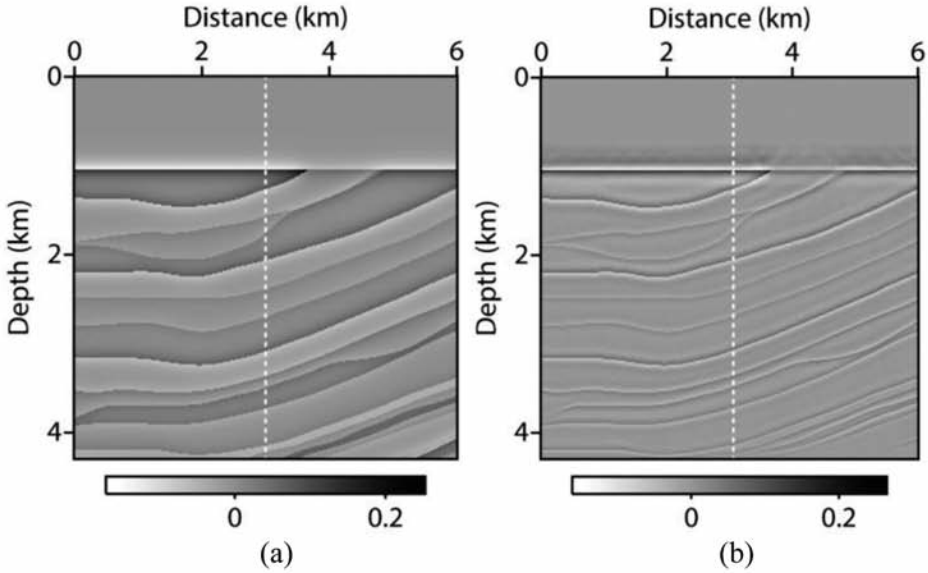Fig. 8. The smoothed velocity model.

Fig. 9. The result about LSRTM. (a) The model parameter **m** , which could be considered the theoretical value of LSRTM. (b) The calculated result of LSRTM.
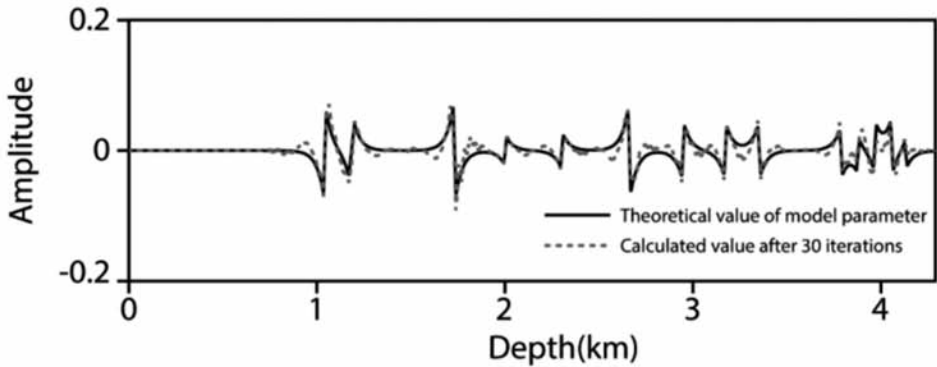

Fig. 10. The amplitude comparison of LSRTM, the black dashed line is the theoretical value marked by the white dotted line in Fig. 9a, and the dotted line is the calculated value marked by the white dotted line in Fig. 9b.
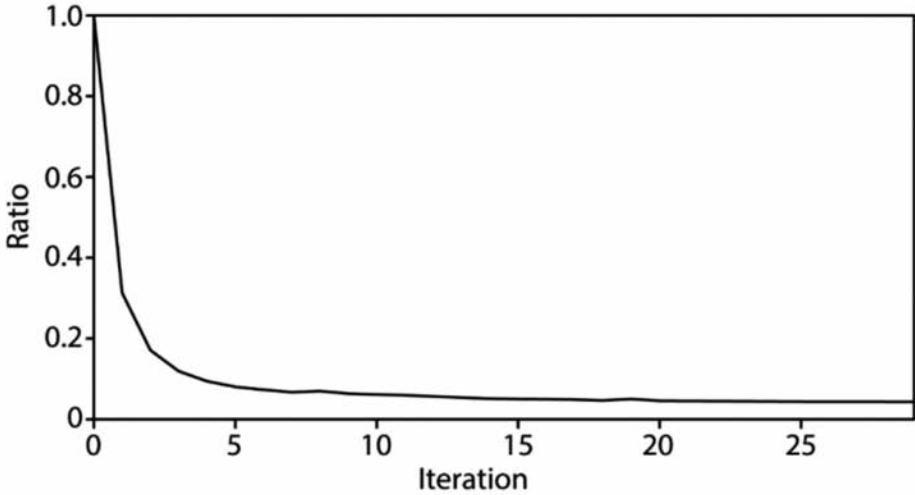
Fig. 11. The energy ratio of the data residual and original data with the iterations of LSRTM.

CONCLUSION AND DISCUSSION

We present an approach that accurately reconstructs the source wavefield backward in the time domain. Compared with the traditional reconstruction method, the proposed method greatly reduces the storage cost. Consequently, it is feasible for LSRTM, and, due to the reduced storage, our method is very promising for GPU acceleration and LSRTM for three dimensional data. However, the scheme we presented is only feasible for the wave equation in a non-dissipative heterogeneous medium, it does not work well for the visco-acoustic wave equation.

It is noteworthy that the currently proposed approach about LSRTM and FWI are both under high-frequency assumption. In the next researches, we hope to place our method within some new imaging principle based on the predicted source and receiver experiment imaging condition as mentioned in the 'Introduction' of this paper, which could avoid the high-frequency assumption and promote the theoretical rigor and practical application value of our research.

ACKNOWLEDGMENTS

We thank the editor and the anonymous reviewers for their valuable comments, which benefit us from a broader perspective on the limitations of

REFERENCES

Claerbout, J., 1971. Toward a unified theory of reflector mapping. Geophysics, 36(3): 467-481.

Claerbout, J., 2004. Earth Soundings Analysis: Processing Versus Inversion. Blackwell Scientific Publications, Boston: 131-153.

Clapp, R.G., 2008. Reverse time migration: Saving the boundaries. Stanford Exploration Project: 136-144.

Clapp, R.G., 2009. Reverse time migration with random boundaries. Expanded Abstr., 79th Ann. Internat. SEG Mtg., Houston: 2809–2813.

Dai, W. and Schuster, G.T., 2013. Plane-wave least-squares reverse-time migration. Geophysics, 78(4): S165-S177.

Feng, B. and Wang, Y.H., 2012. Reverse time migration with source wavefield reconstruction strategy. J. Geophys. Engineer., 9: 69-74.

Liu, F. and Weglein, A.B., 2014. The first wave equation migration RTM with data consisting of primaries and internal multiples: theory and 1D examples. J. Seismic Explor., 23: 357-366.

Plessix, R.E., 2006. A review of the adjoint-state method for computing the gradient of a functional with geophysical applications. Geophys. J. Internat., 167: 495-503.

Symes, W.W., 2007. Reverse time migration with optimal checkpointing. Geophysics, 72(5): 213- 221.

Weglein, A.B., Stolt, R.H. and Mayhan, J.D., 2011. Reverse-time migration and Green's theorem: Part I - The evolution of concepts, and setting the stage for the new RTM method. J. Seismic Explor., 20: 73-90.

Weglein, A.B., Stolt, R.H. and Mayhan, J.D., 2011. Reverse time migration and Green's theorem: Part II - A new and consistent theory that progresses and corrects current RTM concepts and methods. J. Seismic Explor., 20: 135-159.

Weglein, A.B., Zou, Y.L., Fu, Q., Liu, F., Wu, J., Ma, C., Stolt, R.H., Lin, X.L. and Mayhan, J.D., 2016. The first migration method that is equally effective for acquired frequencies for imaging and inverting at the target and reservoir. Expanded Abstr., 86th Ann. Internat. SEG Mtg., Dallas: 4266-4272.

Weglein, A.B., 2016. Multiples: Signal or noise? Geophysics., 81(4): V283-V302.

Yang, P.L., Gao, J.H. and Wang, B.L., 2014. RTM using effective boundary saving: A staggered grid GPU implementation. Comput. Geosci., 68: 64-72.

Zou, Y.L, Fu, Q. and Weglein, A.B., 2017. A wedge resolution comparison between RTM and the first migration method that is equally effective at all frequencies at the target: tests and analysis with both conventional and broadband data. Expanded Abstr., 87th Ann. Internat. SEG Mtg., Houston: 4468-4472.

APPENDIX

The constant density acoustic wave equation

$$\frac{1}{v^2(\mathbf{x})}\frac{\partial^2 p(\mathbf{x},t;\mathbf{x}_s)}{\partial t^2} - \nabla^2 p(\mathbf{x},t;\mathbf{x}_s) = f(\mathbf{x},t;\mathbf{x}_s) \qquad . \tag{A-1}$$

Assuming the velocity $v(\mathbf{x})$ is the stack of the background velocity perturbation velocity $v_S(\mathbf{x})$:

$$v(\mathbf{x}) = v_b(\mathbf{x}) + v_S(\mathbf{x}) \quad , \tag{A-2}$$

which will generate the wavefield:

$$p(\mathbf{x},t;\mathbf{x}_S) = p_b(\mathbf{x},t;\mathbf{x}_S) + p_S(\mathbf{x},t;\mathbf{x}_S) \quad . \tag{A-3}$$

$p_b(\mathbf{x},t;\mathbf{x}_S)$ denotes the background wavefield and $p_S(\mathbf{x},t;\mathbf{x}_S)$ is the perturbation wavefield. Then from eq. (A-1), we have

$$\frac{1}{[v_b(\mathbf{x}) + v_s(\mathbf{x})]^2}\frac{\partial^2 [p_b(\mathbf{x},t;\mathbf{x}_s) + p_s(\mathbf{x},t;\mathbf{x}_s)]}{\partial t^2} - \nabla^2 [p_b(\mathbf{x},t;\mathbf{x}_s) + p_s(\mathbf{x},t;\mathbf{x}_s)] = f(\mathbf{x},t;\mathbf{x}_s) \, .$$

$$\tag{A-4}$$

The second-order Taylor series expansion of the velocity term could be expressed as the following equation:

$$\frac{1}{[v_b(\mathbf{x}) + v_s(\mathbf{x})]^2} \approx \frac{1}{v_b^{\,2}(\mathbf{x})} - \frac{2v_s(\mathbf{x})}{v_b(\mathbf{x})} \quad .$$

Substituting the equation above into eq. (A-4), we have

$$\frac{1}{v_b^2(\mathbf{x})}\frac{\partial^2 p_b(\mathbf{x},t;\mathbf{x}_s)}{\partial t^2} + \frac{1}{v_b^2(\mathbf{x})}\frac{\partial^2 p_s(\mathbf{x},t;\mathbf{x}_s)}{\partial t^2} - \frac{2v_s(\mathbf{x})}{v_b^3(\mathbf{x})}\frac{\partial^2 p_b(\mathbf{x},t;\mathbf{x}_s)}{\partial t^2} - \frac{2v_s(\mathbf{x})}{v_b^3(\mathbf{x})}\frac{\partial^2 p_s(\mathbf{x},t;\mathbf{x}_s)}{\partial t^2}$$

$$-\nabla^2 p_b(\mathbf{x},t;\mathbf{x}_s) - \nabla^2 p_s(\mathbf{x},t;\mathbf{x}_s) = f(\mathbf{x},t;\mathbf{x}_s) \tag{A-5}$$

Ignoring the high order perturbation wavefield term $\dfrac{2v_s(\mathbf{x})}{v_b{}^3(\mathbf{x})}\dfrac{\partial^2 p_b(\mathbf{x},t;\mathbf{x}_s)}{\partial t^2}$

the eq. (A-5) could be expressed as follows:

$$\begin{cases} \dfrac{1}{v_b{}^2(\mathbf{x})}\dfrac{\partial^2 p_b(\mathbf{x},t;\mathbf{x}_s)}{\partial t^2}-\nabla^2 p_b(\mathbf{x},t;\mathbf{x}_s)=f(\mathbf{x},t;\mathbf{x}_s) \\[2ex] \dfrac{1}{v_b{}^2(\mathbf{x})}\dfrac{\partial^2 p_s(\mathbf{x},t;\mathbf{x}_s)}{\partial t^2}-\nabla^2 p_s(\mathbf{x},t;\mathbf{x}_s)=\dfrac{2v_s(\mathbf{x})}{v_b{}^3(\mathbf{x})}\dfrac{\partial^2 p_b(\mathbf{x},t;\mathbf{x}_s)}{\partial t^2} \end{cases} \qquad . \text{ (A-6)}$$

Defining $\mathbf{m}(\mathbf{x})=\dfrac{2v_s(\mathbf{x})}{v_b(\mathbf{x})}$, eq. (A-6) could be expressed as:

$$\begin{cases} \dfrac{1}{v_b^2(\mathbf{x})}\dfrac{\partial^2 p_b(\mathbf{x},t;\mathbf{x}_s)}{\partial t^2}-\nabla^2 p_b(\mathbf{x},t;\mathbf{x}_s)=f(\mathbf{x},t;\mathbf{x}_s) \\[2ex] \dfrac{1}{v_b^2(\mathbf{x})}\dfrac{\partial^2 p_s(\mathbf{x},t;\mathbf{x}_s)}{\partial t^2}-\nabla^2 p_s(\mathbf{x},t;\mathbf{x}_s)=\mathbf{m}(\mathbf{x})\dfrac{1}{v_b^2(\mathbf{x})}\dfrac{\partial^2 p_b(\mathbf{x},t;\mathbf{x}_s)}{\partial t^2} \end{cases} \qquad . \text{ (A-7)}$$