GEOPHYSICS, VOL. 80, NO. 1 (JANUARY-FEBRUARY 2015); P. V23–V31, 7 FIGS.
10.1190/GEO2013-0461.1

A stable and practical implementation of least-squares reverse time migration

Yu Zhang[1], Lian Duan[2], and Yi Xie[3]

ABSTRACT

By adapting reverse time migration (RTM) and demigration as the migration and modeling operators to maximize the crosscorrelation between the simulated and the acquired seismic data, we introduced a new practical least-squares RTM (LSRTM) scheme and derived a steepest descent method in seeking the optimal image. Through synthetic and real data experiments, we determined that the proposed LSRTM provided high-quality images with balanced amplitudes, improved focusing, and enhanced resolution. The method was also capable of removing free surface ghosts caused by towed streamer acquisition, filling the structures and reducing crosstalk noise associated with simultaneous shooting.

INTRODUCTION

Migration is recognized as the most important process in seeking the subsurface structures and reflectivity. In the past, the development of prestack depth imaging has focused on improving its propagation operators to tackle the challenges of imaging in increasingly complicated subsurface structure, from ray-tracing-based Kirchhoff migration (Schneider, 1978) and Beam migration (Hill, 1990, 2001) to the wave-propagation-based one-way wave equation migration (Claerbout, 1971; Claerbout and Doherty, 1972; Gazdag, 1978) and reverse time migration (RTM) (Hemon, 1978; Baysal et al., 1983; McMechan, 1983; Whitmore, 1983). For complicated structures, RTM is the state-of-the-art imaging technology. Its true amplitude migration theory, which aims to automatically compensate for the geometric spreading occurring during the migration process and obtain the angle-dependent reflectivity, has been developed by Zhang and Sun (2009) and Xu et al. (2011). However, such a theory assumes perfect acquisition with regular surface sampling, infinite recording aperture, and unaliased seismic data, which cannot be achieved in practice. As a result and depending on the severity, these assumptions are not satisfied; the imaging quality of RTM could suffer even with a good velocity model.

To remove the acquisition footprint and to improve the quality of seismic imaging, least-squares migration (LSM) has been proposed to seek an inverted image, which generates the simulated data best matching the amplitude of the seismic data. The idea of LSM was first applied to Kirchhoff migration (Schuster, 1993; Nemeth et al., 1999), then generalized to one-way wave equation migration (Wang et al., 2005; Wang and Sacchi, 2007; Tang, 2008), and now applied to RTM (Dai et al., 2011). The method can be implemented in either the image domain (Tang, 2008; Aoki and Schuster, 2009; Dai et al., 2011) or the time domain (Tang and Biondi, 2009; Dai et al., 2010, 2011; Zhan and Schuster, 2010).

In practice, it is not an easy task to directly apply the (conventional) amplitude-matching-based least-squares RTM (LSRTM) (Dong et al., 2012; Yao and Jakubowicz, 2012). The earth is at least a viscoelastic medium with density variations; hence, it is much more complicated than the models we use to propagate acoustic wavefields in seismic imaging. As a result, the amplitude matching is never perfect. Also, it is difficult to define a good source signature in the modeling. The challenge of determining source strength, which can vary from source to source, is even greater. All these practical issues require a considerable effort in preprocessing the observed and simulated data to correctly use the conventional LSRTM formulation.

In this paper, we propose a new general framework of LSRTM, based on maximizing the crosscorrelation of the simulated and observed data at zero lag. The new method relaxes the amplitude constraints and thus can be applied to real data with a stable performance. We call our new method *correlative LSRTM* (CLSRTM). We derive the time-domain steepest descent direction and use it in an image-domain conjugate gradient formulation. Synthetic and real numerical results show that CLSRTM leads to improved image

Manuscript received by the Editor 11 December 2013; revised manuscript received 13 August 2014; published online 22 December 2014.
[1]Formerly CGG, Crompton Way, Crawley, West Sussex, UK; presently ConocoPhillips, Houston, Texas, USA. E-mail: yu.zhang@conocophillips.com.
[2]CGG, Compton Way, Crawley, West Sussex, UK. E-mail: lian.duan@cgg.com.
[3]CGG, Serangoon, Singapore. E-mail: yi.xie@cgg.com.

quality, revealing more structural details in the areas of poor illumination and imaging.

We begin with the introduction of the migration and modeling operators in the RTM setting before highlighting the practical issues and proposing the crosscorrelation objective function. As the core of this paper, the CLSRTM is then derived and summarized. We present synthetic and real examples to demonstrate the limitations and advantages of the proposed CLSRTM.

REVERSE TIME MIGRATION FORMULATIONS

A true amplitude migration intends to find the approximate angle-dependent reflectivity and is derived by seeking the inverse of its forward modeling problem (Bleistein et al., 2001). For a zero-phased and designatured shot record $D(x_r, y_r; x_s, y_s; t)$, with the shot at $\mathbf{x}_s = (x_s, y_s, z_s = 0)$ and receivers at $\mathbf{x}_r = (x_r, y_r, z_r = 0)$, the true amplitude RTM algorithm can be summarized as forward propagation of the source wavefield p_S (Zhang et al., 2007):

$$\begin{cases} \left(\frac{1}{v^2(\mathbf{x})}\frac{\partial^2}{\partial t^2} - \nabla^2\right) p_S(\mathbf{x}; t; \mathbf{x}_s) = 0, \\ p_S(x, y, z = 0; t; \mathbf{x}_s) = \delta(\mathbf{x} - \mathbf{x}_s) \int_0^t f(t')\mathrm{d}t', \end{cases} \tag{1}$$

and backward propagation of the receiver wavefield p_R by reducing time,

$$\begin{cases} \left(\frac{1}{v^2(\mathbf{x})}\frac{\partial^2}{\partial t^2} - \nabla^2\right) p_R(\mathbf{x}; t; \mathbf{x}_s) = 0, \\ p_R(x, y, z = 0; t; \mathbf{x}_s) = D(x, y; t; \mathbf{x}_s), \end{cases} \tag{2}$$

where $\mathbf{x} = (x, y, z)$ is the subsurface imaging location, v denotes velocity, $f(t)$ is the source wavelet with a flat spectrum, and ∇^2 denote the Laplacian operator. To obtain the true amplitude subsurface angle domain image, we apply the crosscorrelation imaging condition:

$$R(\mathbf{x}; \theta) = \iiint p_S(\mathbf{x}; t; \mathbf{x}_s) p_R(\mathbf{x}; t; \mathbf{x}_s) \delta(\theta - \theta') \mathrm{d}t \mathrm{d}\mathbf{x}_s \mathrm{d}\theta', \tag{3}$$

in 2D (Zhang and Sun, 2009) and

$$R(\mathbf{x}; \theta; \psi) = \iiint \frac{v(\mathbf{x})}{\sin\theta'} p_S(\mathbf{x}; t; \mathbf{x}_s) p_R(\mathbf{x}; t; \mathbf{x}_s) \delta(\theta - \theta') \delta(\psi - \psi') \mathrm{d}t \mathrm{d}\mathbf{x}_s \mathrm{d}\theta', \tag{4}$$

in 3D (Xu et al., 2011), where θ is reflection angle, and ψ is azimuth angle.

The RTM formulation in equations 1–4 provides true amplitude angle-dependent reflectivity if the acquisition is perfect, and the seismic wave propagation can be governed by the acoustic wave equation without density variation. However, both assumptions are not practical in the real world. On the other hand, as generating angle domain common image gathers is still computationally intensive, most of the time geophysicists use the full stacked image (Claerbout, 1971)

$$r(\mathbf{x}) = \iint p_S(\mathbf{x}; t; \mathbf{x}_s) p_R(\mathbf{x}; t; \mathbf{x}_s) \mathrm{d}t \mathrm{d}\mathbf{x}_s \tag{5}$$

to interpret subsurface structures. Hence, LSM is proposed here to improve the RTM imaging quality of the stacked image $r(\mathbf{x})$ rather than angle domain common image gathers.

CONVENTIONAL LEAST-SQUARES MIGRATION

To demonstrate the need for LSM, we generated the synthetic seismic record in Figure 1c using the Sigsbee2a velocity model (Paffenholz et al., 2002) in Figure 1a with conventional streamer acquisition of constant source and receiver depth and perform the described RTM with the smooth migration velocity in Figure 1b. The output stacked image in Figure 1e is suboptimal compared with the true reflectivity in Figure 1d. Limited recording aperture, unbalanced illumination under the overburden, and spectrum distortion due to the source and receiver ghosts all contribute to the poor imaging quality.

The existing LSM aims to overcome the migration image imperfectness caused by the acquisition footprint by minimizing the amplitude difference between the seismic data $D(\mathbf{x}_r; t; \mathbf{x}_s)$ and the simulated data $d(\mathbf{x}_r; t; \mathbf{x}_s)$ using the following objective function:

$$E(r) = \iiint (D(\mathbf{x}_r; t; \mathbf{x}_s) - d(\mathbf{x}_r; t; \mathbf{x}_s))^2 d\mathbf{x}_r \mathrm{d}t \mathrm{d}\mathbf{x}_s, \tag{6}$$

where d is generated by a demigration process; i.e.,

$$d(\mathbf{x}_r; t; \mathbf{x}_s) = M(r(\mathbf{x})). \tag{7}$$

The described LSM has been formulated and implemented based on the Kirchhoff migration (Schuster, 1993; Nemeth et al., 1999), the one-way wave equation migration (Tang, 2008), and the RTM (Dai et al., 2011). However, it is not an easy task to apply the conventional LSM to a real data project because (1) in most situations, a scaling difference between the recorded and simulated data, which might be shot, receiver, and time dependent, must be accounted for during the data matching and (2) data preprocessing, including proper scaling, source designature, denoising, and filtering, is compulsory and must be applied to recorded and simulated data. These preprocessing steps can be complicated and often vary greatly from one data set to another.

REVERSE TIME DEMIGRATION

RTM is the most advanced migration tool and is desirable to use as a migration engine in the LSM. Its demigration counterpart (M) must therefore defined accordingly in the RTM setting with the stacked image r as the input. Zhang and Duan (2012) use the following reverse time demigration (RTDM) to predict the seismic data d:

$$\begin{cases} \left(\frac{1}{v^2(\mathbf{x})}\frac{\partial^2}{\partial t^2} - \nabla^2\right) p_S(\mathbf{x}; t; \mathbf{x}_s) = \delta(\mathbf{x} - \mathbf{x}_s) f(t), \\ \left(\frac{1}{v^2(\mathbf{x})}\frac{\partial^2}{\partial t^2} - \nabla^2\right) p_R(\mathbf{x}; t; \mathbf{x}_s) = r(\mathbf{x}) \frac{\partial}{\partial t} p_S(\mathbf{x}; t; \mathbf{x}_s), \\ d(\mathbf{x}_r; t; \mathbf{x}_s) = p_R(\mathbf{x}_r; t; \mathbf{x}_s). \end{cases} \tag{8}$$

Note that the receiver wavefield p_R only propagates waves that are generated by the fictitious sources generated by the multiplication

of the time derivative of the source wavefield p_S and the stacked image r on the right side. The simulated data d are thus generated using the stacked image rather than the physical angle-dependent reflectivity. There are some immediate practical and compatibility issues when adapting the described RTDM in the conventional amplitude matching LSRTM.

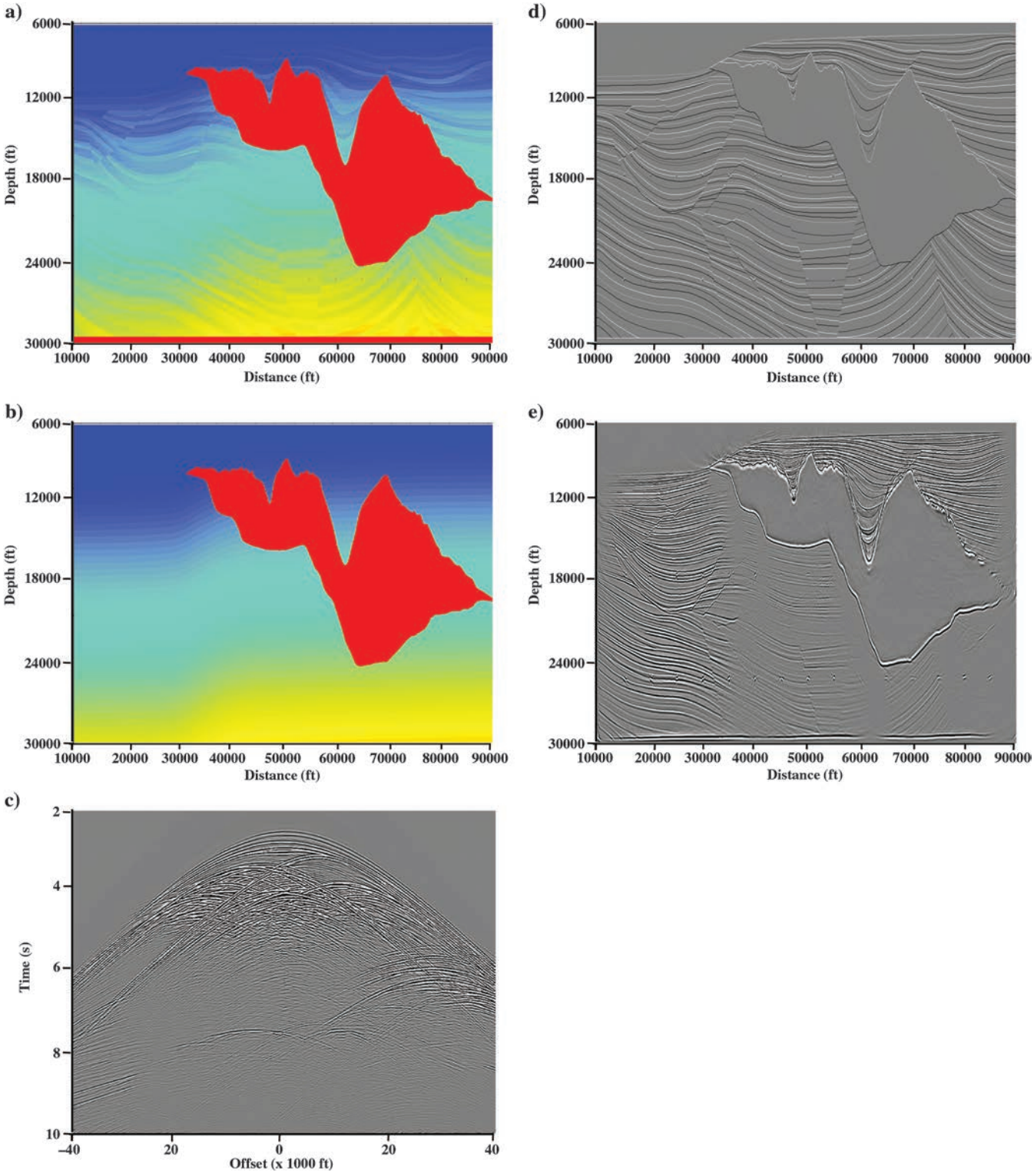


Figure 1. (a) The Sigsbee2a stratigraphic velocity model. (b) The migration velocity model. (c) A synthetic seismic record generated using the stratigraphic velocity model in panel (a). (d) The true reflectivity generated using the stratigraphic velocity model in panel (a). (e) The migration output from conventional RTM.

To simplify our discussion, we assume the recorded seismic data D can be modeled using the acoustic wave propagation of the wavefield $p(\mathbf{x};t;\mathbf{x}_s)$ through an isotropic medium with velocity $v(\mathbf{x})$ and density $\rho(\mathbf{x})$ as follows:

$$\begin{cases} \left(\frac{1}{v^2(\mathbf{x})}\frac{\partial^2}{\partial t^2} - \rho(\mathbf{x})\nabla\cdot\left(\frac{1}{\rho(\mathbf{x})}\nabla\right)\right)p(\mathbf{x};t;\mathbf{x}_s) = \delta(\mathbf{x}-\mathbf{x}_s)f(t), \\ D(\mathbf{x}_r;t;\mathbf{x}_s) = p(\mathbf{x}_r;t;\mathbf{x}_s). \end{cases} \tag{9}$$

Using the RTDM process in system 8, the conventional LSM objective function 6 is difficult in practice because (1) even in an ideal situation with known exact velocity and source information, the RTDM system 8 (relying purely on a stacked image without density variation) and the acoustic modeling system 9 (with density variation) produce different propagation amplitudes and (2) when the source information f is absent or the demigration velocity is not exact, the data matching step using amplitude information only is even more challenging. We will address these issues in the next section and propose our solutions.

In Appendix A, we prove that the operator M^T, which is the transpose of the demigration operator, is an imaging process defined by forward-propagating the source

$$\left(\frac{1}{v^2(\mathbf{x})}\frac{\partial^2}{\partial t^2} - \nabla^2\right)p_S(\mathbf{x};t;\mathbf{x}_s) = \delta(\mathbf{x}-\mathbf{x}_s)f(t), \tag{10}$$

backward-propagating the seismic data

$$\left(\frac{1}{v^2(\mathbf{x})}\frac{\partial^2}{\partial t^2} - \nabla^2\right)p_R(\mathbf{x};t;\mathbf{x}_s) = -\frac{\partial}{\partial t}d(x_r,y_r;t;\mathbf{x}_s)\delta(z), \tag{11}$$

and applying the crosscorrelation imaging condition 5 for all the shots to obtain the stacked image $r(\mathbf{x})$. Comparing the wave propagation formulation in equations 10 and 11 with the true amplitude RTM formulation in equations 1 and 2, the difference between the two wavefields is only in amplitude. Except for a global constant, the amplitude difference depends on surface velocity and propagation angles from the sources and the receivers. This gives rise to the classical statement that "migration is the transpose of a wave equation modeling."

CORRELATIVE LEAST-SQUARES REVERSE TIME MIGRATION

To overcome the practical issues using the amplitude-matching objective function 6 in conventional LSM, we propose the following crosscorrelation-based objective function in the time domain:

$$E(r) = -\int\int\int \frac{d(\mathbf{x}_r;t;\mathbf{x}_s)\cdot D(\mathbf{x}_r;t;\mathbf{x}_s)}{\sqrt{\int d^2(\mathbf{x}_r;t;\mathbf{x}_s)dt}\sqrt{\int D^2(\mathbf{x}_r;t;\mathbf{x}_s)\mathrm{d}t}}\mathrm{d}t\mathrm{d}\mathbf{x}_s\mathrm{d}\mathbf{x}_r. \tag{12}$$

The negative sign on the right side is introduced so that the optimal solution is achieved when the objective function E reaches its minimum. The crosscorrelation relaxes on the amplitude matching and uses phase information to measure the closeness between the simulated data and the observed seismic data. A similar idea has been introduced to full-waveform inversion to invert velocity (Van Leeuwen and Mulder, 2010). In CLSRTM, we assume that the velocity model, which is embedded in the modeling operator M, is correct or already optimized. Note that the value of E is unchanged by rescaling the modeled seismic data $M(r)$, where r is the full stacked image as defined in equation 5. Therefore, the exact overall scaling of the source strength can be ignored in CLSRTM.

Our goal is to find the optimal image r, which maximizes the crosscorrelation between the observed seismic and simulated data at zero lag, or equivalently, to minimize the objective function defined in equation 12. In the best scenario, in which the two data sets are identical or with a constant scaling difference, the objective function reaches its minimum -1. The numerical solution can be found using the steepest descent method. We start the derivation by seeking the gradient of the objective function 12 through a small perturbation of the reflectivity image δr as follows:

$$E(r+\delta r) - E(r) = -\frac{1}{\sqrt{\int D^2\mathrm{d}t}} \times \iint\left(\frac{\int M(r)\cdot D\mathrm{d}t}{\sqrt{\int (M(r))^2\mathrm{d}t}} - \frac{\int M(r+\delta r)\cdot D\mathrm{d}t}{\sqrt{\int (M(r)+M(\delta r))^2}}\right)\mathrm{d}\mathbf{x}_s\mathrm{d}\mathbf{x}_r. \tag{13}$$

Applying the Taylor expansion of δr to the first order, equation 13 can be approximated as

$$E(r+\delta r) - E(r) \approx \frac{1}{\sqrt{\int D^2\mathrm{d}t}\sqrt{\int (M(r))^2\mathrm{d}t}} \times \iint\left(\frac{\int M(r)\cdot D\mathrm{d}t\int M(r)\cdot M(\delta r)\mathrm{d}t}{\int (M(r))^2\mathrm{d}t} - \int M(\delta r)\cdot D\mathrm{d}t\right)\mathrm{d}\mathbf{x}_s\mathrm{d}\mathbf{x}_r. \tag{14}$$

We can thus define the gradient of the cost function as

$$\frac{dE}{dr} = M^T\left[\frac{1}{\sqrt{\int d^2\mathrm{d}t}\sqrt{\int D^2\mathrm{d}t}}\left(\frac{\int (d\cdot D)\mathrm{d}t}{\int d^2\mathrm{d}t}d - D\right)\right]. \tag{15}$$

The detailed steps in deriving equations 14 and 15 are enclosed in Appendix B.

Correlative least-squares reverse time migration scheme

We assume that the observed data D, the initial reflectivity image r_0, and the initial simulated data $d_0 = M(r_0)$ are provided. Setting $\delta s_0 = 0$, for iteration $i = 1,2,\ldots$, the conjugate gradient scheme of our CLSRTM can be summarized as follows:

1) Calculate the steepest descent direction first in the time domain and then convert it to the image domain using RTM as derived in equation 15: $\delta r_i = \frac{dE}{dr_i}$.
2) Compute the weighting coefficient: $\beta_i = \max\left(0, \frac{\int \delta r_i(\delta r_i - \delta r_{i-1})\mathrm{d}\mathbf{x}}{\int \delta r_i \delta r_i \mathrm{d}\mathbf{x}}\right)$.

3) Update the conjugate gradient direction: $\delta s_i = \delta r_i + \beta_i \delta s_{i-1}$.
4) Perform seismic modeling of the conjugate gradient direction: $\delta d_i = M(\delta s_i)$.
5) Perform an update on simulated seismic data using $d_i = d_i - \alpha \delta d_i$ and inverted image using $r_i = r_i - \alpha \delta s_i$, where the scalar α is found to minimize the objective function 12 using the linear searching method.
6) Check the objective function and stopping criterion. Either stop or go to the next iteration and repeat steps one through five.

NUMERICAL EXPERIMENTS

In the first example, the situation of simultaneous shooting (Beasley et al., 1998) is tested in which three identical sources are firing at the same time with a source depth of 12.5 m and source separation distances of 1 km. Streamer cables are towed in between these sources at the same depth, and shot and receiver ghosts are recorded (Figure 2b). Setting the background velocity to 2 km s^{-1} and the density model to contain a dipping staircaselike reflector in Figure 2a, we use an acoustic modeling code to generate the seismic data. The diffraction events from the sharp corners in the density model are evident in the data. In CLSRTM, we assume a prior knowledge of the source wavelet and use RTDM to simulate seismic data with ghosts. Figure 3b presents the inverted image at the end of the 20th iteration of CLSRTM. The migration artifacts and crosstalk noise associated with multiple sources are suppressed, and the wavelet at the dipping reflector is better focused and appears much sharper compared with the initial RTM image in Figure 3a. It appears that CLSRTM gradually removes ghost effects over the iterations and gives a wider frequency bandwidth and more balanced spectrum. In Figure 4, we verify our observations by plotting the frequency spectrum of the inverted amplitudes on the dipping reflector for every fifth iteration in Figure 4a and the numerical value of the objective function for every iteration in Figure 4b. Although the amplitude spectra indicates a gradual deghosting process over the iterations, converging to the correct solution with a flat spectrum from 0 to 32 Hz, the value of the objective function reduces rapidly for the first several iterations, and then steadily improves in the later iterations.

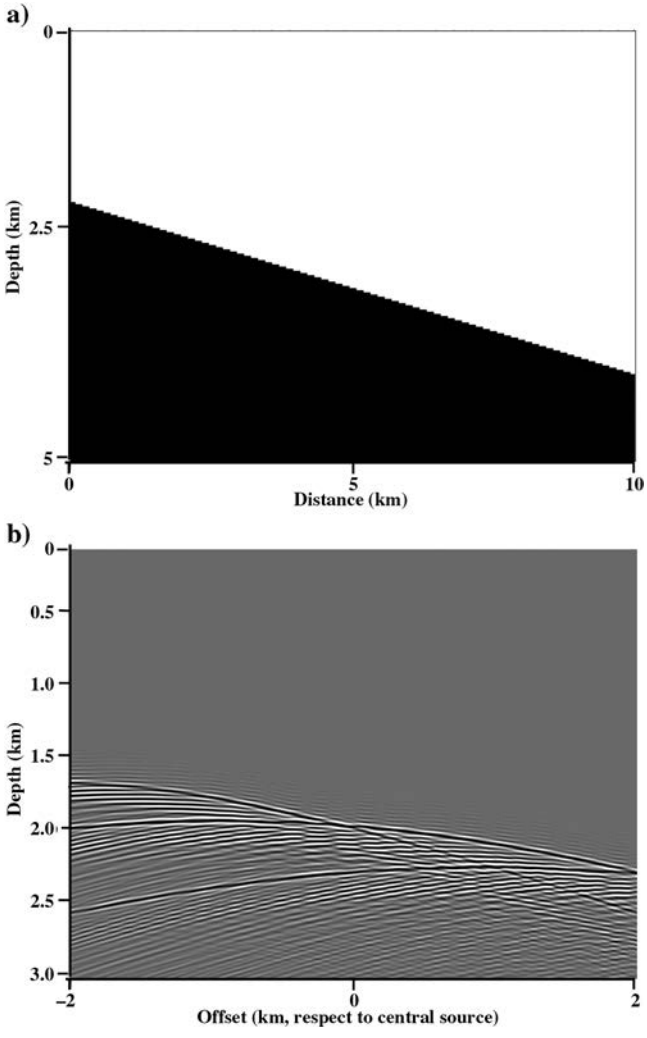


Figure 2. (a) The true density model, with sharp corners in the grid acting as diffractors. (b) A record containing energy scattered by the density contrasts, from three shot points fired simultaneously.

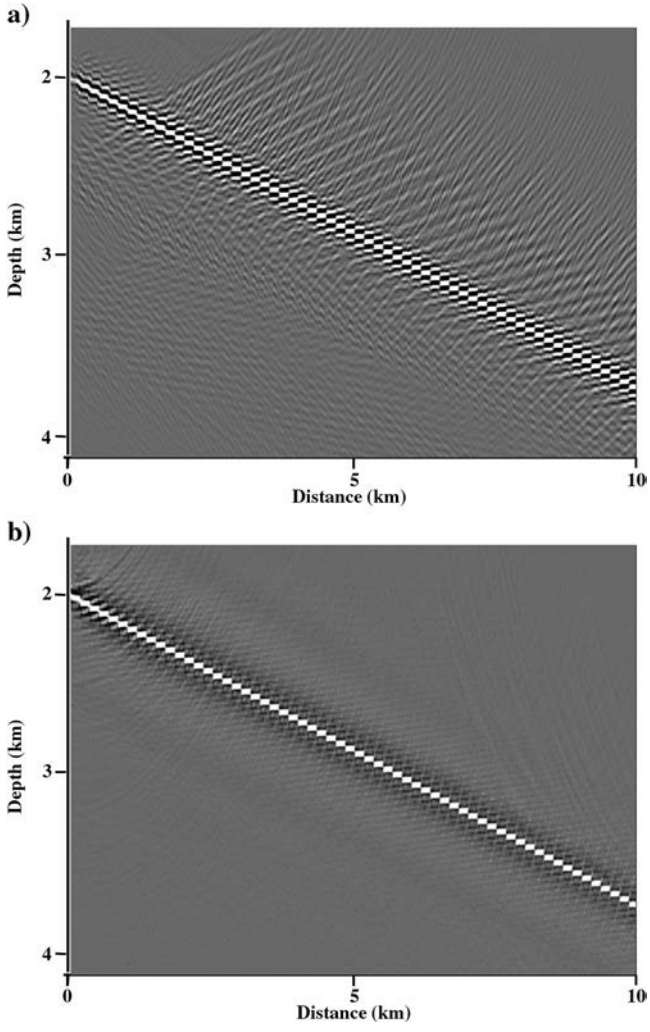


Figure 3. (a) The initial image using conventional RTM. (b) The inverted image after 20 iterations of CLSRTM.

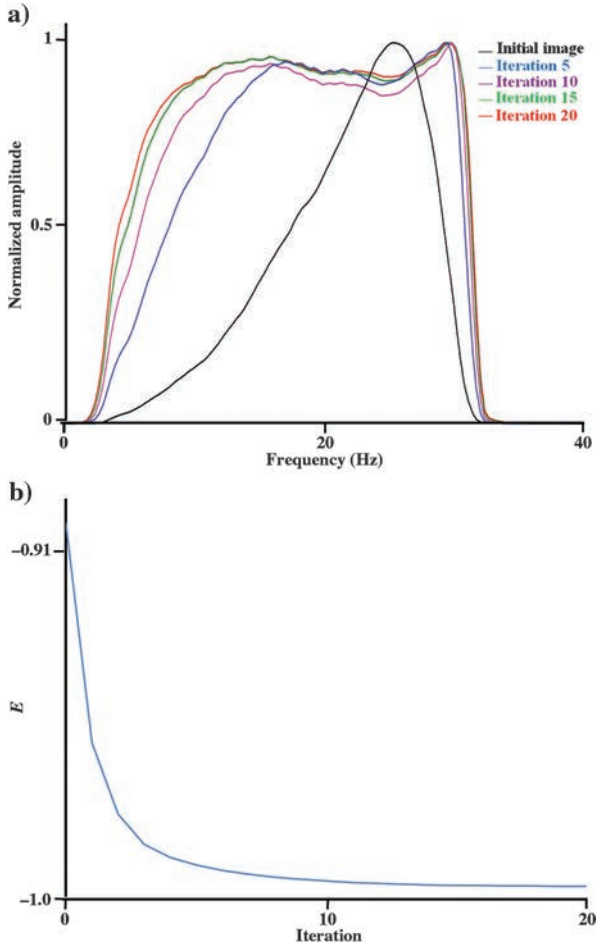


Figure 4. (a) The frequency spectrum variations of the dipping reflector in Figure 2. (b) The value of the objective function over the iterations of the CLSRTM.

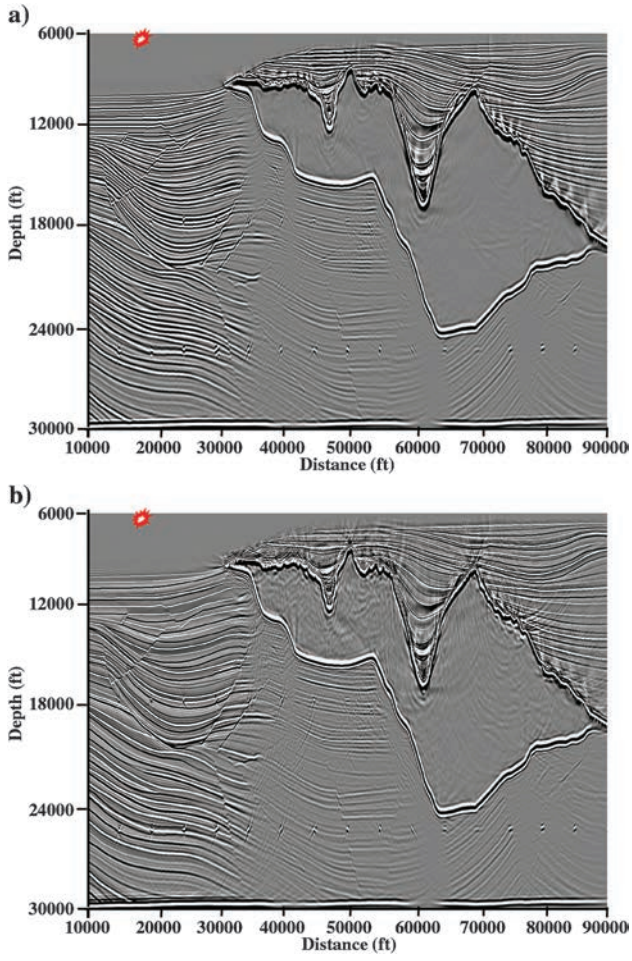
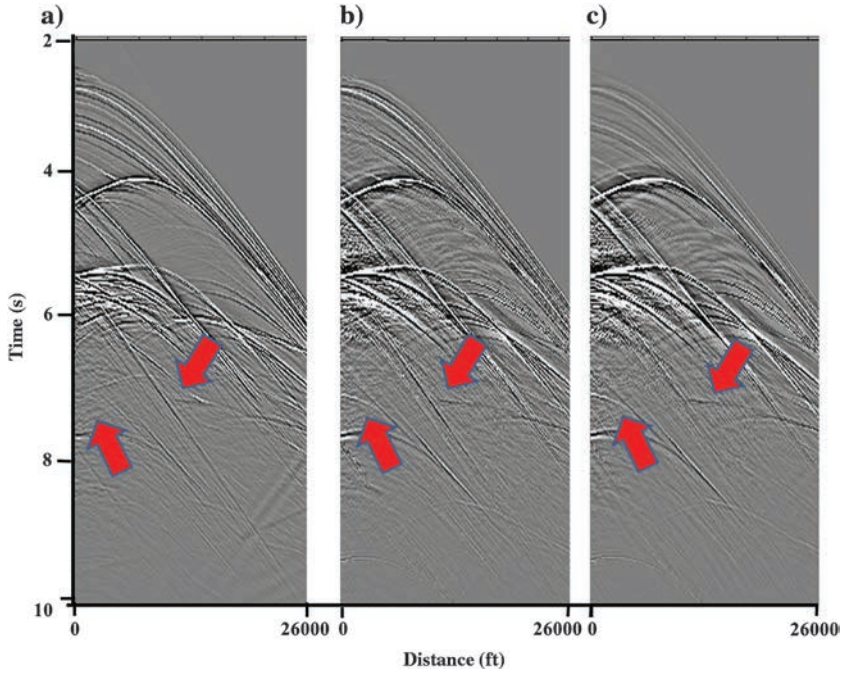


Figure 6. (a) The initial image using conventional RTM. (b) The inverted image after 10 iterations of CLSRTM using migration velocity.

Figure 5. At the indicated shot location in Figure 6, (a) the initial simulated shot record, (b) the simulated shot record after 10 CLSRTM iterations, and (c) the recorded seismic data.



The second example performs CLSRTM on the 2D synthetic data set Sigsbee2A (Paffenholz et al., 2002). The seismic data are generated using a fine stratigraphic velocity model as shown in Figure 1a, with 45.72-m (150-ft) shot spacing, 22.86-m (75-ft) receiver spacing, and 7924.8-m (26,000-ft) maximum offset. Source and receiver ghosts are recorded at a 7.62-m (25-ft) depth. We first migrate the seismic data using the smoothed migration velocity in Figure 1b. As illustrated in Figure 5, the initial simulated data in Figure 5a do not match the seismic data (Figure 5c) well, due to acquisition limitations and the imperfectness of the migration operator. However, over the iterations, CLSRTM gradually drives the simulated data to approach the seismic data. For example, as the arrows indicate in Figure 5, some reflections that are weak on the initial simulated data (Figure 5a) but strong on the seismic data (Figure 5c) are enhanced after 10 iterations of CLSRTM (Figure 5b). This demonstrates that our proposed method works well to match the seismic data by updating the reflectivity model when the migration velocity is essentially correct. In Figure 6, we compare the initial RTM image in Figure 6a with the inverted image after 10 iterations of CLSRTM in Figure 6b. The sedimentary structure on the left is much sharper because of the deghosting effect intrinsically built into CLSRTM, whereas the subsalt structures are enhanced to give a better overall amplitude balance. The equally spaced diffractors purposely embedded in the velocity model also provide references that the focusing in the inverted image is improved. However, we observe some artifacts in the sediments above the salt after CLSRTM due to the imperfect migration velocity used.

In the final example, we use real seismic data from the Central North Sea, which has been band-pass filtered from 5 to 45 Hz, and regularized to a 50×50 m receiver grid. Again, the source signature is unknown, and the migration/demigration model is tilted transversely isotropic (TTI). In the initial RTM image shown in Figure 7a, a salt dome is present in the center, but its steeply dipping flanks are weakly imaged. After nine iterations of CLSRTM, illumination of the image near the salt dome boundary is enhanced (Figure 7b), showing strong energy on the nearly

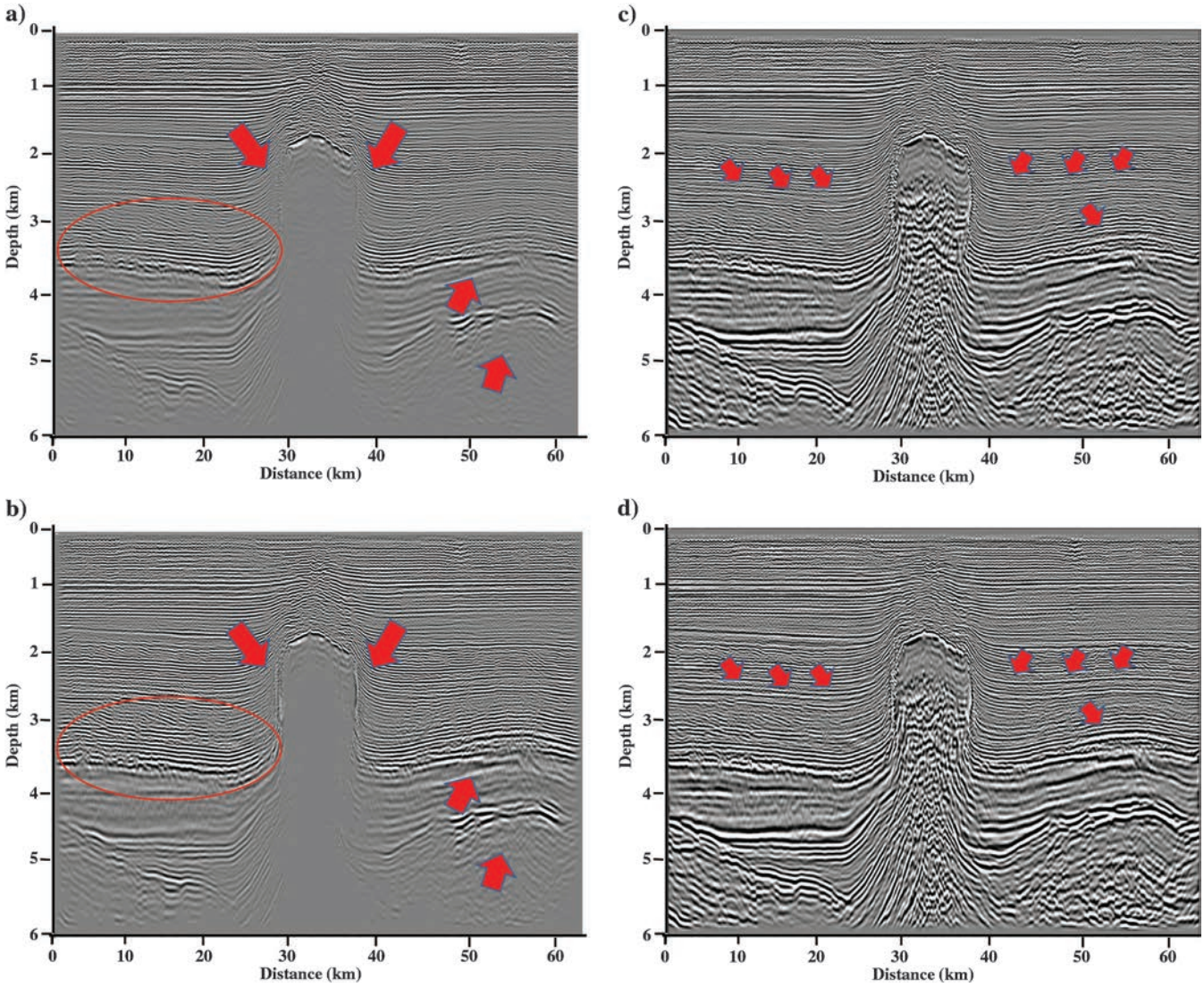


Figure 7. (a) The initial image using conventional RTM. (b) The inverted image after nine iterations of CLSRTM. (c) The initial image using conventional RTM with 500 ms AGC applied. (d) The inverted image after nine iterations of CLSRTM with 500 ms AGC applied.

vertical salt flanks and the termination structures, compared with the initial image. In addition, illumination of the chalk section on either side of the salt in the inverted image is improved. In many places, the structures become more continuous. With an application of automatic gain control (AGC) to the initial image in Figure 7a and the inverted image in Figure 7b, we further identify that the LSRTM (as shown in Figure 7d) is not only just balancing the overall amplitude of the initial image in Figure 7c, but it also reveals structural detail that is otherwise unable to be identified. Overall, the inverted image using CLSRTM is amplitude balanced with improved continuity and reveals more detailed structures.

CONCLUSIONS

LSM is an attractive technique for improving image resolution and illumination and suppressing migration artifacts. Yet, its application to daily production remains challenging. We have shown that using a crosscorrelation objective function, CLSRTM provides stable solutions even when the source signature is unknown. Because RTDM can simulate data with free surface boundary conditions for any acquisition, CLSRTM is capable of handling some of the difficult imaging issues caused by acquisition, such as free surface ghosts for towed streamers and crosstalk caused by simultaneous shooting. Our numerical experiments also indicate that LSRTM can be sensitive to velocity errors after a couple of iterations. A processing flow to combine model update and LSRTM is necessary to further improve the imaging quality.

ACKNOWLEDGMENTS

We thank CGG's Multi-Client and New Ventures Division for providing the North Sea data sets and our colleagues in CGG for their support, especially R. Wombell, A. Ratcliffe, G. Roberts, S. Gray, and J. Sun for their suggestions to improve the paper. Finally, we are grateful to the associate editors M. D. Sacchi and E. Slob and other anonymous reviewers for their help in improving the final quality of this paper.

APPENDIX A

THE TRANSPOSE OF DEMIGRATION OPERATOR *M*

The demigration process $d(\mathbf{x}_r;t;\mathbf{x}_s)=M(r(\mathbf{x}))$ is defined by the following formulations:

$$\begin{cases}\left(\frac{1}{v^2(\mathbf{x})}\frac{\partial^2}{\partial t^2}-\nabla^2\right)p_S(\mathbf{x};t;\mathbf{x}_s)=\delta(\mathbf{x}-\mathbf{x}_s)f(t),\\ \left(\frac{1}{v^2(\mathbf{x})}\frac{\partial^2}{\partial t^2}-\nabla^2\right)p_R(\mathbf{x};t;\mathbf{x}_s)=r(\mathbf{x})\frac{\partial}{\partial t}p_S(\mathbf{x};t;\mathbf{x}_s),\ \text{and}\times\\ d(\mathbf{x}_r;t;\mathbf{x}_s)=p_R(\mathbf{x}_r;t;\mathbf{x}_s).\end{cases} \tag{A-1}$$

Now, we try to find its transpose. For a function $\tilde{d}(\mathbf{x}_r;t;\mathbf{x}_s)$ in the data domain, the transpose $M^T(\tilde{d}(\mathbf{x}_r;t;\mathbf{x}_s))=\tilde{r}(\mathbf{x})$ is defined by

$$\int\tilde{r}(\mathbf{x})\cdot r(\mathbf{x})d\mathbf{x}=\iiint\tilde{d}(\mathbf{x}_r;t;\mathbf{x}_s)\cdot d(\mathbf{x}_r;t;\mathbf{x}_s)\mathrm{d}\mathbf{x}_r\mathrm{d}t\mathrm{d}\mathbf{x}_s. \tag{A-2}$$

If we define Green's function as

$$\left(\frac{1}{v^2(\mathbf{x})}\frac{\partial^2}{\partial t^2}-\nabla^2\right)G(\mathbf{x};t-t';\mathbf{x}')=\delta(\mathbf{x}-\mathbf{x}')\delta(t-t'), \tag{A-3}$$

then from the second equation in equation A-1 we have

$$p_R(\mathbf{x};t;\mathbf{x}_s)=\iint G(\mathbf{x};t-t';\mathbf{x}')r(\mathbf{x}')\frac{\partial}{\partial t}p_S(\mathbf{x}';t';\mathbf{x}_s)\mathrm{d}t'\mathrm{d}\mathbf{x}' \tag{A-4}$$

and

$$\begin{aligned}d(\mathbf{x}_r;t;\mathbf{x}_s)&=M(r(\mathbf{x}))\\&=\iint G(\mathbf{x}_r;t-t';\mathbf{x})r(\mathbf{x})\frac{\partial}{\partial t}p_S(\mathbf{x};t';\mathbf{x}_s)\mathrm{d}t'\mathrm{d}\mathbf{x}.\end{aligned} \tag{A-5}$$

Substituting equation A-5 into equation A-2, we have

$$\begin{aligned}&\int\tilde{r}(\mathbf{x})\cdot r(\mathbf{x})\mathrm{d}\mathbf{x}\\&=\iiint\tilde{d}(\mathbf{x}_r;t;\mathbf{x}_s)\left(\iint G(\mathbf{x}_r;t-t';\mathbf{x})r(\mathbf{x})\right.\\&\quad\left.\times\frac{\partial}{\partial t}p_S(\mathbf{x};t';\mathbf{x}_s)\mathrm{d}t'\mathrm{d}\mathbf{x}\right)\mathrm{d}\mathbf{x}_r\mathrm{d}t\mathrm{d}\mathbf{x}_s\\&=\int\left[\int\left(\iiint\tilde{d}(\mathbf{x}_r;t;\mathbf{x}_s)G(\mathbf{x}_r;t-t';\mathbf{x})d\mathbf{x}_r\mathrm{d}t\mathrm{d}\mathbf{x}_s\right)\right.\\&\quad\left.\times\frac{\partial}{\partial t}p_S(\mathbf{x};t';\mathbf{x}_s)\mathrm{d}t'\right]r(\mathbf{x})\mathrm{d}\mathbf{x}\\&=\int\left[\iint\left(\iint-\frac{\partial}{\partial t}\tilde{d}(\mathbf{x}_r;t';\mathbf{x}_s)G(\mathbf{x}_r;t'-t;\mathbf{x})\mathrm{d}\mathbf{x}_r\mathrm{d}t'\right)p_S\right.\\&\quad\left.\times(\mathbf{x};t;\mathbf{x}_s)\mathrm{d}t\mathrm{d}\mathbf{x}_s\right]r(\mathbf{x})\mathrm{d}\mathbf{x}.\end{aligned} \tag{A-6}$$

If we define

$$\tilde{p}_R(\mathbf{x};t;\mathbf{x}_s)=\iiint-\frac{\partial}{\partial t}\tilde{d}(\mathbf{x}_r;t';\mathbf{x}_s)G(\mathbf{x}_r;t'-t;\mathbf{x})\mathrm{d}\mathbf{x}_r\mathrm{d}t', \tag{A-7}$$

then $\tilde{r}(\mathbf{x})=M^T(\tilde{d}(\mathbf{x}_r;t;\mathbf{x}_s))$ can be generated by the following migration process:

1) Forward propagate $p_S(\mathbf{x};t;\mathbf{x}_s)$ by solving the wave equation

$$\left(\frac{1}{v^2(\mathbf{x})}\frac{\partial^2}{\partial t^2}-\nabla^2\right)p_S(\mathbf{x};t;\mathbf{x}_s)=\delta(\mathbf{x}-\mathbf{x}_s)f(t). \tag{A-8}$$

2) Backward propagate $p_R(\mathbf{x};t;\mathbf{x}_s)$ by solving the wave equation

$$\left(\frac{1}{v^2(\mathbf{x})}\frac{\partial^2}{\partial t^2}-\nabla^2\right)p_R(\mathbf{x};t;\mathbf{x}_s)=-\frac{\partial}{\partial t}\tilde{d}(x_r,y_r;t;\mathbf{x}_s)\delta(z). \tag{A-9}$$

3) Apply the crosscorrelation imaging condition for all the shots

$$r(\mathbf{x}) = \iint p_S(\mathbf{x};t;\mathbf{x}_s)p_R(\mathbf{x};t;\mathbf{x}_s)\mathrm{d}t\mathrm{d}\mathbf{x}_s. \qquad \text{(A-10)}$$

APPENDIX B

DERIVATION OF EQUATION 15

We start the derivation by seeking the gradient of the objective function 12 through a small perturbation of the reflectivity image δr as follows:

$$E(r+\delta r)-E(r) = \iint \frac{\int(M(r)\cdot D)\mathrm{d}t}{\sqrt{\int(M(r))^2dt}\sqrt{\int(D)^2\mathrm{d}t}} - \frac{\int(M(r+\delta r)\cdot D)\mathrm{d}t}{\sqrt{\int(M(r+\delta r))^2dt}\sqrt{\int(D)^2\mathrm{d}t}}\mathrm{d}\mathbf{x}_s\mathrm{d}\mathbf{x}_r. \qquad \text{(B-1)}$$

Applying Taylor expansion to equation B-1 and neglecting higher orders terms, we have

$$\begin{aligned}
&E(r+\delta r)-E(r)\\
&\approx \iint \frac{\int(M(r)\cdot D)\mathrm{d}t\sqrt{\int M(r)(M(r)+2M(\delta r))\mathrm{d}t}-\int(M(r+\delta r)\cdot D)\mathrm{d}t\sqrt{\int(M(r))^2\mathrm{d}t}}{\sqrt{\int(D)^2dt}\sqrt{\int(M(r))^2\mathrm{d}t}\sqrt{\int M(r)(M(r)+2M(\delta r))\mathrm{d}t}}\mathrm{d}\mathbf{x}_s\mathrm{d}\mathbf{x}_r\\
&= \iint \frac{\frac{2\int(M(r)\cdot D)dt\int(M(r)M(\delta r))\mathrm{d}t}{\sqrt{\int M(r)(M(r)+2M(\delta r))\mathrm{d}t}+\sqrt{\int(M(r))^2\mathrm{d}t}}-\int(M(\delta r)\cdot D)\mathrm{d}t\sqrt{\int(M(r))^2\mathrm{d}t}}{\sqrt{\int(D)^2\mathrm{d}t}\sqrt{\int(M(r))^2\mathrm{d}t}\sqrt{\int M(r)(M(r)+2M(\delta r))\mathrm{d}t}}\mathrm{d}\mathbf{x}_s\mathrm{d}\mathbf{x}_r\\
&\approx \iint \frac{\frac{\int(M(r)\cdot D)\mathrm{d}t\int(M(r)M(\delta r))\mathrm{d}t}{\sqrt{\int(M(r))^2\mathrm{d}t}}-\int(M(\delta r)\cdot D)\mathrm{d}t\sqrt{\int(M(r))^2\mathrm{d}t}}{\sqrt{\int(D)^2\mathrm{d}t}\int(M(r))^2\mathrm{d}t}\mathrm{d}\mathbf{x}_s\mathrm{d}\mathbf{x}_r\\
&= \iiint q(\mathbf{x}_r;t;\mathbf{x}_s)(M(\delta r))(\mathbf{x}_r;t;\mathbf{x}_s)\mathrm{d}t\mathrm{d}\mathbf{x}_r\mathrm{d}\mathbf{x}_s, \qquad \text{(B-2)}
\end{aligned}$$

where

$$\begin{aligned}
q(\mathbf{x}_r;t;\mathbf{x}_s) &= \frac{1}{\sqrt{\int(M(r))^2dt}\sqrt{\int D^2\mathrm{d}t}}\\
&\quad\times\left[\frac{\int(M(r)\cdot D)\mathrm{d}t}{\int(M(r))^2\mathrm{d}t}M(r)-D\right](\mathbf{x}_r;t;\mathbf{x}_s)\\
&= \frac{1}{\sqrt{\int d^2\mathrm{d}t}\sqrt{\int D^2\mathrm{d}t}}\left[\frac{\int(d\cdot D)\mathrm{d}t}{\int d^2\mathrm{d}t}d-D\right](\mathbf{x}_r;t;\mathbf{x}_s).
\end{aligned} \qquad \text{(B-3)}$$

Applying the relationship in equation A-2 to equation B-2, we have

$$E(r+\delta r)-E(r) \approx \int (M^T(q))(\mathbf{x})\cdot\delta r(\mathbf{x})\mathrm{d}\mathbf{x}. \qquad \text{(B-4)}$$

REFERENCES

Aoki, N., and G. Schuster, 2009, Fast least-squares migration with a deblurring filter: Geophysics, **74**, no. 6, WCA83–WCA93, doi: 10.1190/1.3155162.

Baysal, E., D. Kosloff, and J. Sherwood, 1983, Reverse time migration: Geophysics, **48**, 1514–1524, doi: 10.1190/1.1441434.

Beasley, C., R. Chambers, and Z. Jiang, 1998, A new look at simultaneous sources: 68th Annual International Meeting, SEG, Expanded Abstracts, 133–135.

Bleistein, N., J. Cohen, and J. Stockwell, 2001, Mathematics of multidimensional seismic inversion: Springer.

Claerbout, J., 1971, Toward a unified theory of reflector mapping: Geophysics, **36**, 467–481, doi: 10.1190/1.1440185.

Claerbout, J., and S. Doherty, 1972, Downward continuation of moveout-corrected seismograms: Geophysics, **37**, 741–768, doi: 10.1190/1.1440185.

Dai, W., C. Boonyasiriwat, and G. Schuster, 2010, 3D multi-source least-squares reverse time migration: 80th Annual International Meeting, SEG, Expanded Abstracts, 3120–3124.

Dai, W., X. Wang, and G. Schuster, 2011, Least-squares migration of multi-source data with a deblurring filter: Geophysics, **76**, no. 5, R135–R146, doi: 10.1190/geo2010-0159.1.

Dong, S., J. Cai, M. Guo, S. Suh, Z. Zhang, B. Wang, and Z. Li, 2012, Least-squares reverse time migration: Towards true amplitude imaging and improving the resolution: 83rd Annual International Meeting, SEG, Expanded Abstracts, 1425–1429.

Gazdag, J., 1978, Wave equation migration with the phase-shift method: Geophysics, **43**, 1342–1351, doi: 10.1190/1.1440899.

Hemon, C., 1978, Equations d'onde et modeles: Geophysical Prospecting, **26**, 790–821, doi: 10.1111/j.1365-2478.1978.tb01634.x.

Hill, R., 1990, Gaussian beam migration: Geophysics, **55**, 1416–1428, doi: 10.1190/1.1442788.

Hill, R., 2001, Prestack Gaussian-beam depth migration: Geophysics, **66**, 1240–1250, doi: 10.1190/1.1487071.

McMechan, G., 1983, Migration by extrapolation of time-dependent boundary values: Geophysical Prospecting, **31**, 413–420, doi: 10.1111/j.1365-2478.1983.tb01060.x.

Nemeth, T., C. Wu, and G. Schuster, 1999, Least-squares migration of incomplete reflection data: Geophysics, **64**, 208–221, doi: 10.1190/1.1444517.

Paffenholz, J., B. McLain, J. Zaske, and P. Keliher, 2002, Subsalt multiple attenuation and imaging: Observations from the Sigsbee2B synthetic dataset: 72nd Annual International Meeting, SEG, Expanded Abstracts, 2122–2125.

Schneider, W., 1978, Integral formulation for migration in two and three dimensions: Geophysics, **43**, 49–76, doi: 10.1190/1.1440828.

Schuster, G., 1993, Least-squares crosswell migration: 63rd Annual International Meeting, SEG, Expanded Abstracts, 110–113.

Tang, Y., 2008, Wave-equation Hessian by phase encoding: 78th Annual International Meeting, SEG, Expanded Abstracts, 2201–2205.

Tang, Y., and B. Biondi, 2009, Least-squares migration/inversion of blended data: 79th Annual International Meeting, SEG, Expanded Abstracts, 2859–2863.

Van Leeuwen, T., and W. A. Mulder, 2010, A correlation-based misfit criterion for wave-equation traveltime tomography: Geophysical Journal International, **182**, 1383–1394, doi: 10.1111/j.1365-246X.2010.04681.x.

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: Geophysics, **70**, no. 5, S91–S99, doi: 10.1190/1.2076748.

Wang, J., and M. D. Sacchi, 2007, High-resolution wave equation AVP imaging with sparseness constraints: Geophysics, **72**, no. 1, S11–S18, doi: 10.1190/1.2387139.

Whitmore, D., 1983, Iterative depth migration by backward time propagation: 53rd Annual International Meeting, SEG, Expanded Abstracts, 382–385.

Xu, S., Y. Zhang, and B. Tang, 2011, 3D angle gathers from reverse time migration: Geophysics, **76**, no. 2, S77–S92, doi: 10.1190/1.3536527.

Yao, G., and H. Jakubowicz, 2012, Least-squares reverse-time migration: 83rd Annual International Meeting, SEG, Expanded Abstracts, 3406–3410.

Zhan, G., and G. Schuster, 2010, Skeletonized least squares wave equation migration: 80th Annual International Meeting, SEG, Expanded Abstracts, 3380–3384.

Zhang, Y., and L. Duan, 2012, Predicting multiples using a reverse time demigration: 83rd Annual International Meeting, SEG, Expanded Abstracts, 520–524.

Zhang, Y., and J. Sun, 2009, Practical issues of reverse time migration: True-amplitude gathers, noise removal and harmonic-source encoding: First Break, **26**, 29–35.

Zhang, Y., S. Xu, N. Bleistein, and G. Zhang, 2007, True amplitude angle domain common image gathers from one-way wave equation migrations: Geophysics, **72**, no. 1, S49–S58, doi: 10.1190/1.2399371.