GEOPHYSICS, VOL. 83, NO. 3 (MAY-JUNE 2018); P. S261–S270, 7 FIGS.
10.1190/GEO2017-0466.1

Check for updates

Downloaded 03/21/18 to 138.67.129.87. Redistribution subject to SEG license or copyright; see Terms of Use at http://library.seg.org/

3D acoustic least-squares reverse time migration using the energy norm

Daniel Rocha[1], Paul Sava[1], and Antoine Guitton[2]

ABSTRACT

We have developed a least-squares reverse time migration (LSRTM) method that uses an energy-based imaging condition to obtain faster convergence rates when compared with similar methods based on conventional imaging conditions. To achieve our goal, we also define a linearized modeling operator that is the proper adjoint of the energy migration operator. Our modeling and migration operators use spatial and temporal derivatives that attenuate imaging artifacts and deliver a better representation of the reflectivity and scattered wavefields. We applied the method to two Gulf of Mexico field data sets: a 2D towed-streamer benchmark data set and a 3D ocean-bottom node data set. We found LSRTM resolution improvement relative to RTM images, as well as the superior convergence rate obtained by the linearized modeling and migration operators based on the energy norm, coupled with inversion preconditioning using image-domain nonstationary matching filters.

INTRODUCTION

Wavefield migration delivers an image of the subsurface structures using wavefield extrapolation methods (Sun et al., 2003; Biondi, 2012). For complex geologic settings, the two-way wave equation is generally used for extrapolation and the migration procedure is known as reverse time migration (RTM) (Baysal et al., 1983; Lailly, 1983; McMechan, 1983; Levin, 1984; Zhang and Sun, 2009). In practice, however, data recording is always incomplete and possibly aliased and irregular, causing wavefield migration to degrade image quality and resolution especially for greater depths (Zhang et al., 2015). In addition, the obtained image does not fully explain data at receiver locations because migration represents the adjoint operator of single-scattering modeling, and, therefore, it is not a good approximation of the inverse operator that correctly reverses propagation of seismic data (Claerbout, 1992).

Considering these imaging quality issues, least-squares migration (LSM) is proposed to deliver images with more accurate amplitudes, illumination compensation, and reduced footprint of the acquisition geometry (Chavent and Plessix, 1999; Nemeth et al., 1999; Kuehl and Sacchi, 2003; Aoki and Schuster, 2009). If the RTM engine is used for migration, the method is called least-squares RTM (LSRTM) (Dai et al., 2010; Dai and Schuster, 2012; Dong et al., 2012; Yao and Jakubowicz, 2012). LSM involves a forward operator (single-scattering modeling), an adjoint operator (migration), and an objective function, which work together to achieve an image that best explains in a least-squares sense the reflection data acquired at receivers. The objective function has the role of comparing modeled data with observed data using the image as reflectivity. To achieve the least-squares solution, we use efficient gradient methods that decrease the objective function iteratively (Hestenes and Stiefel, 1952; Scales, 1987).

Because of the high computational cost of LSRTM, which is at least an order of magnitude higher than RTM, techniques that expedite LSRTM convergence toward the true reflectivity model lead to reduction in computational cost while also achieving a satisfying solution in fewer iterations. For instance, a common procedure to obtain faster rates of convergence is to use an approximate of the Hessian operator (Aoki and Schuster, 2009; Tang, 2009; Dai et al., 2010; Kazemi and Sacchi, 2014, 2015; Hou and Symes, 2016; Huang et al., 2016). Here, we demonstrate that modeling and migration operators based on an imaging condition that delivers more accurate amplitudes and attenuates artifacts, such as the one derived from the energy norm (Douma et al., 2010; Whitmore and Crawley, 2012; Brandsberg-Dahl et al., 2013; Pestana et al., 2013; Rocha et al., 2016), also expedites convergence. This migration operator attenuates low-wavenumber artifacts, delivering a better representation of subsurface reflectivity. The corresponding modeling operator uses spatial and temporal derivatives based on the wave equation itself to generate scattered wavefields from the energy image. We also incorporate preconditioning

Manuscript received by the Editor 21 July 2017; revised manuscript received 2 December 2017; published ahead of production 14 January 2018; published online 16 March 2018.
[1]Colorado School of Mines, Center for Wave Phenomena, Golden, Colorado, USA. E-mail: drocha@mines.edu; psava@mines.edu.
[2]DownUnder GeoSolutions, West Perth, Australia and Colorado School of Mines, Center for Wave Phenomena, Golden, Colorado, USA. E-mail: aguitton@gmail.com.
© 2018 Society of Exploration Geophysicists. All rights reserved.

Downloaded 03/21/18 to 138.67.129.87. Redistribution subject to SEG license or copyright; see Terms of Use at http://library.seg.org/

with an approximation of the Hessian operator in our inversion, which expedites LSRTM convergence even more. This approximation of the Hessian is achieved by a nonstationary multidimensional filter that corrects the blurring effect in the image caused by the wavefield modeling and migration operators (Guitton, 2004; Aoki and Schuster, 2009; Fletcher et al., 2016).

THEORY

We can define acoustic wavefield migration as an operator such that

$$\mathbf{m}_0 = \mathbf{L}^{\mathrm{T}}\mathbf{d}_r^{\mathrm{obs}}, \tag{1}$$

where $\mathbf{L}^{\mathrm{T}}$ is the migration operator based on some imaging condition, $\mathbf{d}_r^{\mathrm{obs}}$ is the single-scattered data observed at receiver locations, and $\mathbf{m}_0$ is an image (an estimate of the earth reflectivity). The operator $\mathbf{L}^{\mathrm{T}}$ involves back-propagation of $\mathbf{d}_r^{\mathrm{obs}}$ through an earth model, thus generating a receiver wavefield $\mathbf{u}_r$, and the application of an imaging condition comparing $\mathbf{u}_r$ with the source wavefield $\mathbf{u}_s$. The migration operator is the adjoint operator of single-scattering modeling (or commonly known as linearized modeling). Therefore, $\mathbf{L}$ is a linear operator such that

$$\mathbf{d}_r = \mathbf{L}\mathbf{m}, \tag{2}$$

and simulates single-scattered data $\mathbf{d}_r$ at receiver locations using an image containing reflectors that act as sources under the action of the source wavefield $\mathbf{u}_s$.

Therefore, we define $\mathbf{m}$ as reflectivity that depends on a certain imaging condition and is not necessarily defined in terms of contrasts in the earth model. The same principle applies to the linearized modeling operator $\mathbf{L}$, which we define as an adjoint operator to a certain imaging condition, and this operator is not directly related to the physics of single scattering.

Conventional imaging condition and linearized modeling

The conventional imaging condition is defined as the zero-lag crosscorrelation between source and receiver wavefields $\mathbf{u}_s$ and $\mathbf{u}_r$:

$$\mathbf{m}_0 = \mathbf{u}_s^{\mathrm{T}}\mathbf{u}_r. \tag{3}$$

Because we know both wavefields are generated by extrapolation using an earth model and observed data at receivers, we can rewrite equation 3 as

$$\mathbf{m}_0 = (\mathbf{E}_+\mathbf{K}_s\mathbf{d}_s)^{\mathrm{T}}\mathbf{E}_-\mathbf{K}_r\mathbf{d}_r^{\mathrm{obs}}, \tag{4}$$

where $\mathbf{E}_+$ and $\mathbf{E}_-$ are the forward and backward extrapolator operators, $\mathbf{K_s}$ and $\mathbf{K_r}$ are the source and receiver injection operators (see Appendix A), respectively, and $\mathbf{d}_s$ is the source function. The subscripts in the extrapolator operators indicate extrapolation direction for either positive (+) or negative (−) times. Note the two important relations between the extrapolator operators: $\mathbf{E}_+ = \mathbf{E}_-^{\mathrm{T}}$ and $\mathbf{E}_- = \mathbf{E}_+^{\mathrm{T}}$. Then, one can write equation 4 similarly to equation 1:

$$\mathbf{m}_0 = \mathbf{L}^{\mathrm{T}}\mathbf{d}_r^{\mathrm{obs}}, \tag{5}$$

where

$$\mathbf{L}^{\mathrm{T}} = (\mathbf{E}_+\mathbf{K}_s\mathbf{d}_s)^{\mathrm{T}}\mathbf{E}_-\mathbf{K}_r. \tag{6}$$

We can obtain the operator $\mathbf{L}$ (adjoint of $\mathbf{L}^{\mathrm{T}}$) if we apply the adjoint for each individual operator and reverse the order of operators. Therefore, conventional linearized modeling is defined as

$$\mathbf{d}_r = \mathbf{L}\mathbf{m}, \tag{7}$$

where

$$\mathbf{L} = \mathbf{K}_r^{\mathrm{T}}\mathbf{E}_+(\mathbf{E}_+\mathbf{K}_s\mathbf{d}_s) = \mathbf{K}_r^{\mathrm{T}}\mathbf{E}_+\mathbf{u}_s. \tag{8}$$

In other words, single-scattered data $\mathbf{d}_r$ are obtained by extraction at the receiver locations ($\mathbf{K}_r^{\mathrm{T}}$), and wavefield extrapolation ($\mathbf{E}_+$) using $\mathbf{u}_s\mathbf{m}$ as the source term.

Energy imaging condition and linearized modeling

For an extrapolated wavefield that satisfies the acoustic-wave equation, we can compute its energy norm in discretized space and time as (Rocha et al., 2016)

$$\|\mathbf{u}\|_E^2 = \sum_{\mathbf{x},t}\left[\frac{\dot{\mathbf{u}}^2}{v^2(\mathbf{x})} + |\nabla\mathbf{u}|^2\right], \tag{9}$$

where $v(\mathbf{x})$ is the migration velocity, superscript dot indicates the time differentiation, and ∇ is the spatial gradient. On the right side of equation 9, the first term inside the brackets corresponds to the kinetic energy of the wavefield, and the second term corresponds to its potential energy. Based on this norm, we can define an imaging condition that ignores wave events between source and receiver wavefields that share the same propagation direction, thus suppressing low-wavenumber artifacts in RTM that do not characterize subsurface reflectivity:

$$\mathbf{m}_0(\mathbf{x}) = \sum_t\left[\frac{\dot{\mathbf{u}}_s\dot{\mathbf{u}}_r}{v^2} - \nabla\mathbf{u}_s\cdot\nabla\mathbf{u}_r\right]. \tag{10}$$

A more compact form of equation 10 uses the energy operator

$$\square = \left\{\nabla, \frac{1}{v(\mathbf{x})}\frac{\partial}{\partial t}\right\}, \tag{11}$$

which computes a 4D vector field from the scalar wavefield. A negative time derivative is implied for the back-propagated receiver wavefield $\mathbf{u}_r$. If we compute the dot product between such vector fields obtained from source and receiver wavefields at each spatial location (implying summation over time), we can rewrite equation 10 as

$$\mathbf{m}_0 = (\square\mathbf{u}_s)^{\mathrm{T}}\square\mathbf{u}_r. \tag{12}$$

Using the operators introduced in equations 4 and 12 becomes

$$\mathbf{m}_0 = (\square\mathbf{E}_+\mathbf{K}_s\mathbf{d}_s)^{\mathrm{T}}\square\mathbf{E}_-\mathbf{K}_r\mathbf{d}_r^{\mathrm{obs}}. \tag{13}$$

Because equation 13 is a function of $\mathbf{d}_r$, we can write the corresponding migration operation as

$$\mathbf{m}_0 = \mathbf{L}^{\mathrm{T}}\mathbf{d}_r^{\mathrm{obs}}, \tag{14}$$

Downloaded 03/21/18 to 138.67.129.87. Redistribution subject to SEG license or copyright; see Terms of Use at http://library.seg.org/

where

$$\mathbf{L}^{\mathrm{T}} = (\square\mathbf{E}_{+}\mathbf{K}_s\mathbf{d}_s)^{\mathrm{T}}\square\mathbf{E}_{-}\mathbf{K}_r. \tag{15}$$

Therefore, the linearized modeling operation based on the energy norm is defined as

$$\mathbf{d}_r = \mathbf{L}\mathbf{m}, \tag{16}$$

where

$$\mathbf{L} = \mathbf{K}_r^{\mathrm{T}}\mathbf{E}_{+}\square^{\mathrm{T}}\square(\mathbf{E}_{+}\mathbf{K}_s\mathbf{d}_s) = \mathbf{K}_r^{\mathrm{T}}\mathbf{E}_{+}\square^{\mathrm{T}}\square\mathbf{u}_s. \tag{17}$$

In other words, single-scattered data $\mathbf{d}_r$ are obtained by extraction at the receiver locations ($\mathbf{K}_r^{\mathrm{T}}$), and wavefield extrapolation ($\mathbf{E}_{+}$) with $\square^{\mathrm{T}}\square\mathbf{u}_s\mathbf{m}$ as the source term, which for a point in space and time can explicitly be written as

$$[\square^{\mathrm{T}}\square\mathbf{u}_s\mathbf{m}](\mathbf{x},t) = \frac{\mathbf{m}(\mathbf{x})}{v^2(\mathbf{x})}\ddot{\mathbf{u}}_s(\mathbf{x},t) - \nabla\cdot[\mathbf{m}(\mathbf{x})\nabla\mathbf{u}_s(\mathbf{x},t)]. \tag{18}$$

The same procedure to find a proper adjoint operator is applicable to other imaging conditions. For instance, applying a Laplacian operator on a conventional image is theoretically equivalent to the application of the energy imaging condition but in the far field (Douma et al., 2010). Knowing that the Laplacian operator is self-adjoint, the imaging condition with Laplacian filtering can be written as

$$\mathbf{m}_0 = \nabla^2\mathbf{u}_s^{\mathrm{T}}\mathbf{u}_r. \tag{19}$$

The corresponding linearized modeling is written as

$$\mathbf{d}_r = \mathbf{L}\mathbf{m}, \tag{20}$$

where

$$\mathbf{L} = \mathbf{K}_r^{\mathrm{T}}\mathbf{E}_{+}(\mathbf{E}_{+}\mathbf{K}_s\mathbf{d}_s)\nabla^2 = \mathbf{K}_r^{\mathrm{T}}\mathbf{E}_{+}\mathbf{u}_s\nabla^2. \tag{21}$$

The source term for this case is $\mathbf{u}_s\nabla^2\mathbf{m}$ and can explicitly be written for each point in space and time as

$$[\mathbf{u}_s\nabla^2\mathbf{m}](\mathbf{x},t) = \mathbf{u}_s(\mathbf{x},t)\nabla^2\mathbf{m}. \tag{22}$$

LSM and preconditioning with an approximation of the Hessian

A pair of linearized modeling and migration operators enables us to compute an image that minimizes the L_2 norm of the difference between observed and modeled data:

$$E(\mathbf{m}) = \frac{1}{2}\|\mathbf{L}\mathbf{m} - \mathbf{d}_r^{\mathrm{obs}}\|^2, \tag{23}$$

and such image is mathematically described by

$$\mathbf{m}^{\mathrm{LS}} = (\mathbf{L}^{\mathrm{T}}\mathbf{L})^{-1}\mathbf{L}^{\mathrm{T}}\mathbf{d}_r^{\mathrm{obs}}. \tag{24}$$

To find $\mathbf{m}^{\mathrm{LS}}$, one generally uses iterative procedures that exploit the direction of the gradient of the objective function in equation 23 at a given iteration i:

$$\mathbf{g}_i = \frac{\partial E(\mathbf{m}_i)}{\partial\mathbf{m}_i} = \mathbf{L}^{\mathrm{T}}(\mathbf{L}\mathbf{m}_i - \mathbf{d}_r^{\mathrm{obs}}), \tag{25}$$

and the model update at each iteration in steepest descent and conjugate gradient methods is a scaled version of the gradient

$$\mathbf{m}_{i+1} = \mathbf{m}_i - \alpha\mathbf{g}_i. \tag{26}$$

In equation 24, the term $\mathbf{L}^{\mathrm{T}}\mathbf{L}$ is known as the Hessian operator of $E(\mathbf{m})$. If the inverse of the Hessian is applied to the RTM image $\mathbf{L}^{\mathrm{T}}\mathbf{d}_r^{\mathrm{obs}}$, the least-squares solution is achieved. However, in practice, the Hessian operator cannot explicitly be computed for LSM problems. As suggested by Guitton (2004), we can obtain an approximation of the Hessian operator if we compute an image

$$\mathbf{m}_1 = \mathbf{L}^{\mathrm{T}}\mathbf{L}\mathbf{m}_0, \tag{27}$$

where $\mathbf{m}_0 = \mathbf{L}^{\mathrm{T}}\mathbf{d}_r^{\mathrm{obs}}$ is the standard RTM image. The operator $\mathbf{B}$ that minimizes

$$E(\mathbf{B}) = \|\mathbf{m}_0 - \mathbf{B}\mathbf{m}_1\|^2, \tag{28}$$

is a good approximation of the inverse of the Hessian $(\mathbf{L}^{\mathrm{T}}\mathbf{L})^{-1}$ according to equation 27. Here, we define the operator $\mathbf{B}$ as a multidimensional convolutional operator along all spatial axes (Rickett et al., 2001).

Once we obtain the approximation of the inverse of the Hessian, we can expedite convergence in least-squares inversion by preconditioning the gradient with the operator $\mathbf{B}$ before updating the model at each iteration. Incorporating the preconditioning, equation 26 becomes

$$\mathbf{m}_{i+1} = \mathbf{m}_i - \alpha\mathbf{B}\mathbf{g}_i. \tag{29}$$

EXAMPLES

To show how LSRTM with proper imaging operators can get faster convergence rates relative to conventional methods, we perform an LSRTM experiment using a 2D Gulf of Mexico (GOM) data set, used by many authors in the past as a benchmark data set (Dragoset, 1999; Guitton and Cambois, 1999; Hadidi et al., 1999; Lamont et al., 1999; Lokshtanov, 1999; Verschuur and Prein, 1999) and described by Guitton et al. (2012). Such a 2D data set is useful to test the capability of the energy imaging operators in attenuating low-wavenumber artifacts caused by strong velocity contrasts (such as a salt body) and to benchmark our inversion procedure before application to a larger 3D data set. Standard preprocessing is applied to the data set prior to LSRTM, such as surface-related multiple suppression. We use 71 of the original 1001 shot records, with the first source location at $x = 4000$ m and the last at $x = 22{,}670$ m, resulting in a source spacing of approximately 267 m. The shot decimation decreases the computational cost of the entire experiment and also introduces truncation artifacts, which are useful to test the effectiveness of our LSRTM in attenuating acquisition artifacts. Each shot record contains 180 traces with receiver spacing of 26.67 m and a maximum offset of 4874 m. For migration, we use a maximum frequency of 40 Hz and a spatial sampling of 6.67 m. A standard RTM migration (Figure 1a) exhibits low-wavenumber artifacts and poor illumination in the subsalt area. We test two alternative LSRTMs, one using modeling and migration with the Laplacian operator

Downloaded 03/21/18 to 138.67.129.87. Redistribution subject to SEG license or copyright; see Terms of Use at http://library.seg.org/

(equations 19 and 21), and the other based on the energy norm (equations 15 and 17). For this experiment, the final images from the Laplacian and energy LSRTMs are similar in quality and resolution, and we show the final energy LSRTM image in Figure 1b, which contains suppressed low-wavenumber artifacts above the salt body as compared with the conventional image in Figure 1a. In addition, the objective functions (Figure 1c) of both alternative LSRTMs decrease faster than the one from conventional LSRTM, with the one based on the energy norm decreasing slightly faster when compared with its Laplacian counterpart. Although the two alternative LSRTM implementations are theoretically equivalent far from the source and receiver locations, one computes the imaging condition in the image domain and the other in the wavefield domain, leading to slightly different numerical implementations and, therefore, convergence performances. The energy migration and modeling operators applied either in the wavefield domain (equation 18) or in the image domain as a Laplacian operator provide faster convergence rates toward the final reflectivity model.

We apply our method to a 3D ocean-bottom node data set from the GOM. We process the data set to obtain only the downgoing pressure component and perform mirror imaging (Godfrey et al., 1998; Ronen et al., 2005; Wong et al., 2010, 2015) of shallow geologic structures with a fine spatial sampling of 12.5 m. We use 37 node gathers with sources densely distributed at the water surface (Figure 2a), and the velocity model used is shown in Figure 2b. In total, as shown in Figure 2c, we perform four LSRTM experiments:

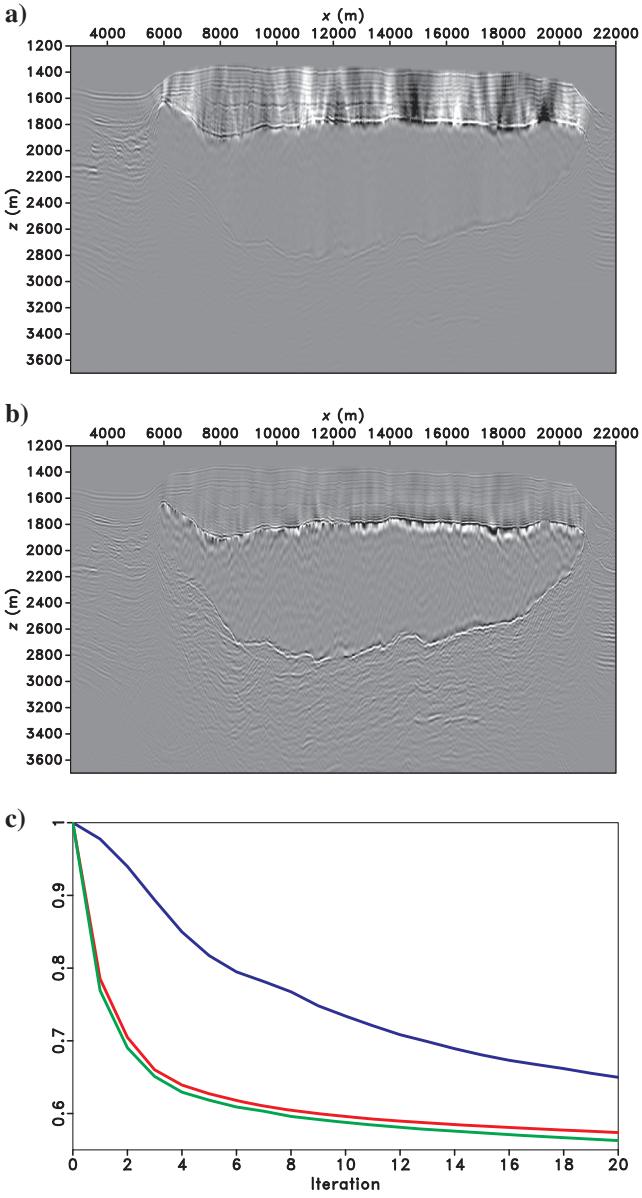


Figure 1. The GOM 2D data set: (a) conventional RTM image and (b) LSRTM image with the energy modeling and migration operators. (c) Normalized objective functions for conventional LSRTM (blue), LSRTM with Laplacian (red) and based on the energy norm (green). Note suppression of low-wavenumber artifacts above the salt and the convergence speed-up obtained by the alternative LSRTMs.

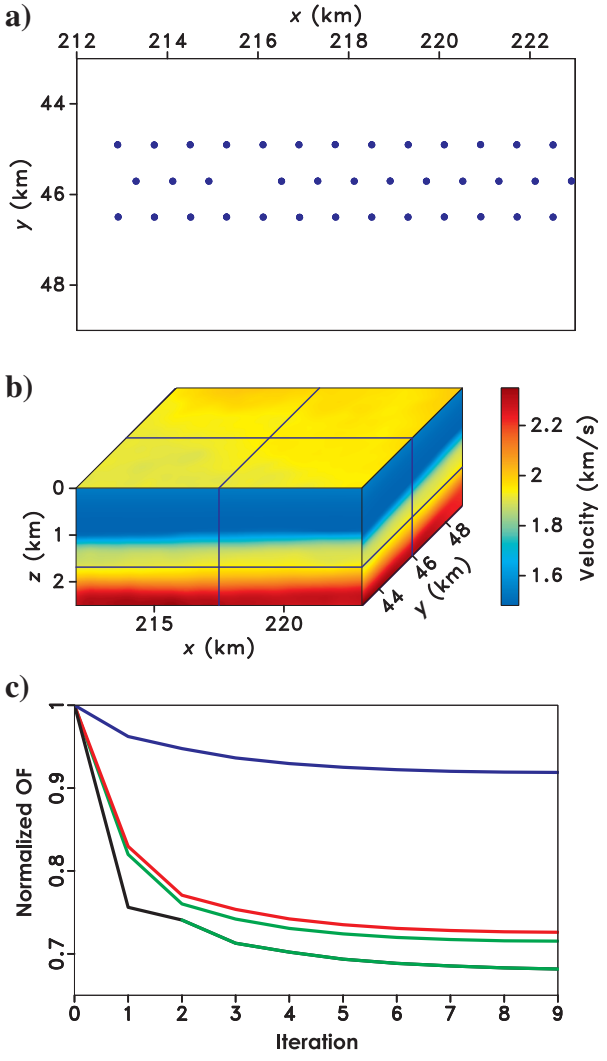


Figure 2. The GOM 3D data set: (a) 37 nodes spaced by approximately 800 m and sources densely distributed at the surface of the model. (b) Velocity model. (c) Normalized objective functions for conventional (blue), Laplacian-based (red), energy-based (green), and preconditioned energy-based (black and then green at iteration 2) LSRTMs. The energy LSRTM has the best performance if we apply preconditioning at the first two iterations.

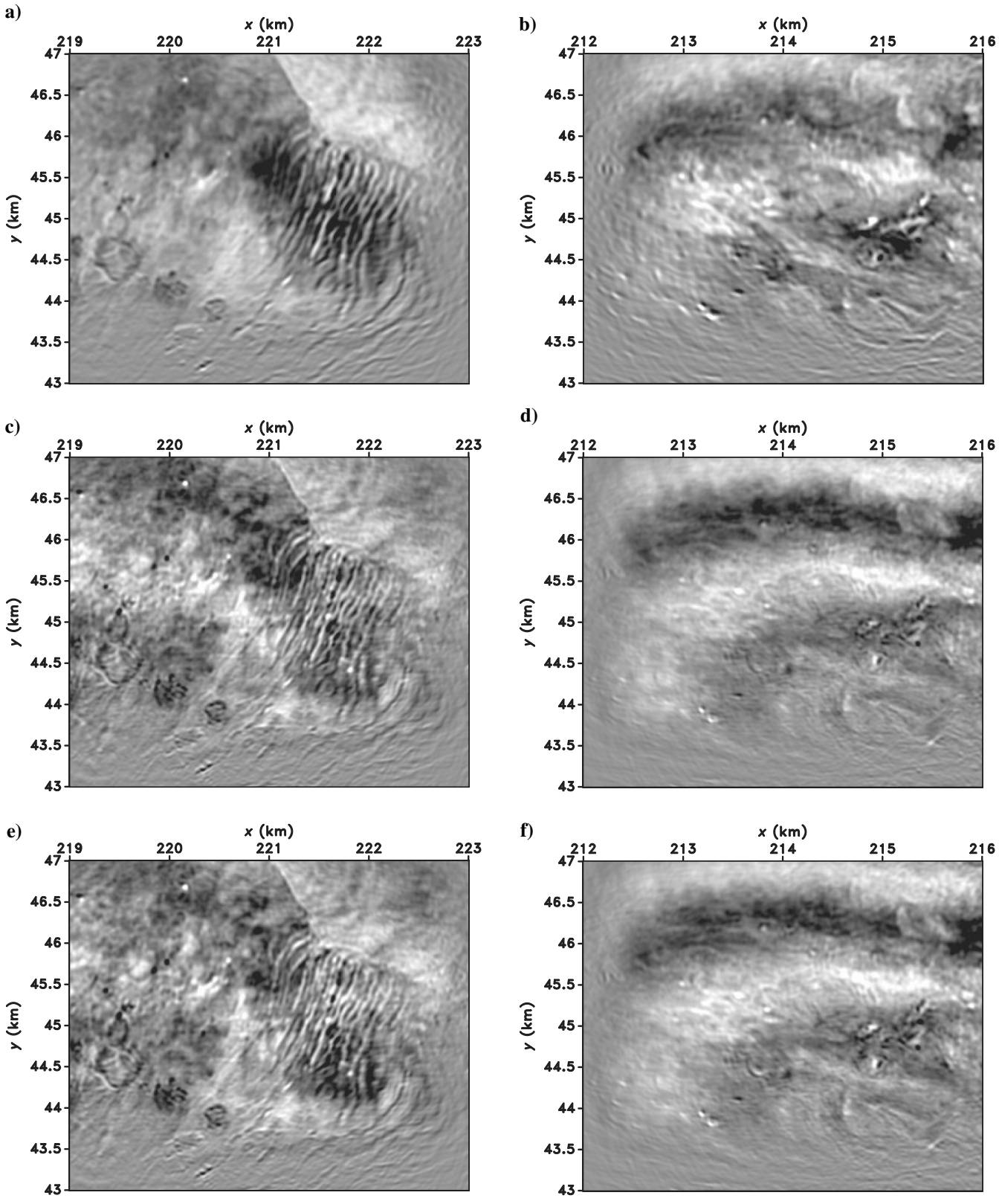


Figure 3. The GOM 3D data set: depth slices at $z = 1.55$ km and $z = 1.77$ km, respectively, for (a/b) RTM, (c/d) conventional LSRTM, and (e/f) energy LSRTM with preconditioning. In the LSRTM images, note the improvement in focusing of the diffractors and in delineation of the reflectors.

Downloaded 03/21/18 to 138.67.129.87. Redistribution subject to SEG license or copyright; see Terms of Use at http://library.seg.org/

conventional LSRTM, Laplacian-based LSRTM, and energy-norm based LSRTM with and without preconditioning. Similar to the preceding example, conventional LSRTM has a worse performance in terms of objective function decrease, and the energy LSRTM has a slightly smaller objective function value over iterations compared with its Laplacian counterpart. We compute the matching filter for the preconditioning using the energy RTM image ($\mathbf{m}_0 = \mathbf{L}^T\mathbf{d}_r^{obs}$) and an image generated by $\mathbf{m}_1 = \mathbf{L}^T\mathbf{L}\mathbf{m}_0$, where the modeling ($\mathbf{L}$) and migration ($\mathbf{L}^T$) operators are based on the energy norm. We use the preconditioning at the two first iterations and switch it off at the second iteration (solid black and green lines in Figure 2c) because the computed matching filter is related to the images at the first and second iterations ($\mathbf{m}_0$ and $\mathbf{m}_1$). The matching filter is biased toward the strong amplitude changes between such images at first iterations, and keeping preconditioning on does not give substantial benefit to the inversion at later iterations, thus only increasing computational cost. Although we apply the preconditioning only

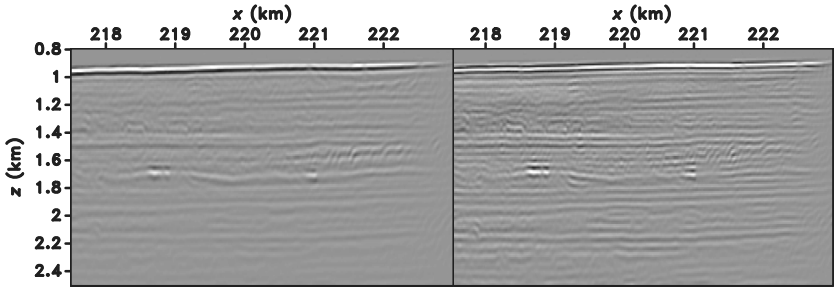


Figure 4. The GOM 3D data set: inline sections for energy RTM (left) and LSRTM with preconditioning (right) at $y = 44.9$ km. Note the increase in frequency content and illumination for the LSRTM image compared with RTM.

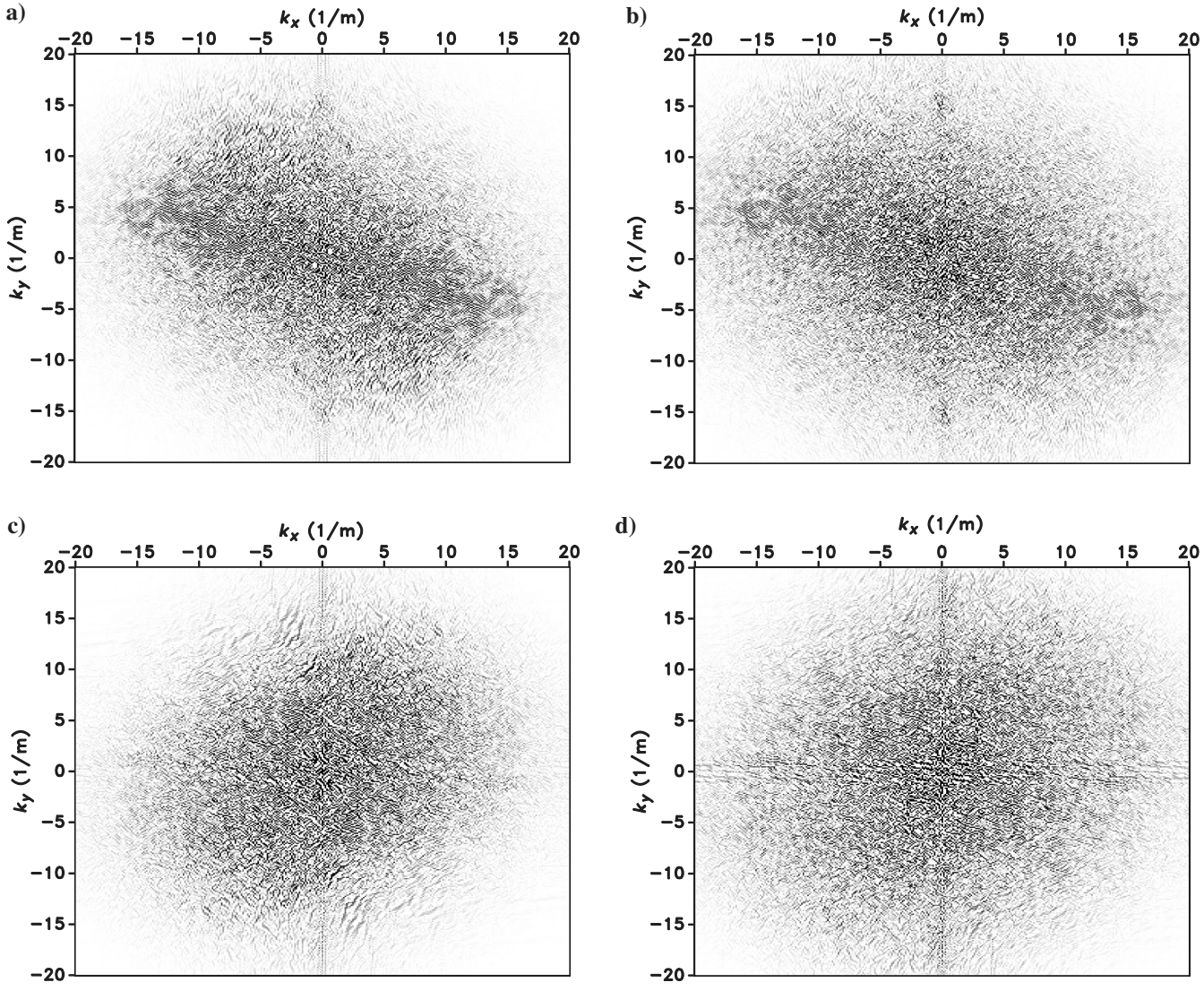


Figure 5. The GOM 3D data set: (a/c) wavenumber-domain RTM and (b/d) energy LSRTM images from Figure 3 (at $z = 1.55$ km and $z = 1.77$ km, respectively). Note the stronger amplitudes recovered by LSRTM at the high wavenumbers (near the corners of the plot).

Downloaded 03/21/18 to 138.67.129.87. Redistribution subject to SEG license or copyright; see Terms of Use at http://library.seg.org/

Downloaded 03/21/18 to 138.67.129.87. Redistribution subject to SEG license or copyright; see Terms of Use at http://library.seg.org/

at the first iterations, we obtain a significant speed-up in the convergence rate over all subsequent iterations.

For two different areas of the migration volume, the depth slices in Figure 3 compare RTM, conventional LSRTM, and our LSRTM that uses energy imaging operators and preconditioning. Although conventional LSRTM fits less data and has the worst convergence rate based on Figure 2c, it delivers images closer in quality to our LSRTM images (Figure 3e and 3f). In these depth slices, some imaged events that seem to disappear or become blurry after LSRTM might indicate that they do not correctly predict reflection events in the data domain, considering that our imaging operator is isotropic and acoustic. Figure 4 compares RTM and our LSRTM cross-sectional images side by side and shows the visual benefit due to the increase in illumination and frequency contents (especially toward the low frequencies) in the LSRTM image. To confirm the increase in bandwidth for the depth slices, we transform the images from Figure 3 to the wavenumber domain and obtain the spectral images in Figure 5. Note how LSRTM recovers high-resolution content by spreading the amplitudes toward the edges of the wavenumber domain. However, one needs to consider that some recovered high-wavenumber content might emerge from noise existent in the data.

As discussed in the "Theory" section, one of the benefits in LSRTM is to suppress the acquisition footprint. In some shallow areas of the migration volume, very close to the edges of the model and to the water bottom, imaging artifacts exist in the RTM image due to the large decimation of nodes for this experiment (from 924 to 37 nodes, causing nodes to be quite sparse as seen in Figure 2a) and the irregularity of source and receiver locations with respect to the regular extrapolation grid. Figure 6 shows an example: LSRTM (Figure 6b) attenuates some of these acquisition footprint artifacts at the water bottom from the RTM image (Figure 6a). Figure 7 compares observed and predicted data at the last iteration, and the difference (residual) between the two for a gather of traces sorted by increasing offset at a particular node location ($x = 214.5$ km, $y = 44.9$ km). Note that the main reflections at near and mid offsets are correctly predicted, but the far-offset amplitudes are not matched mainly because of elasticity and anisotropy, which are not accounted for by our acoustic imaging operators. As expected, the residual in Figure 7c still exhibits noise, such as a large dipping event that goes from $t = 2.2$s at trace 0 to $t = 3.0$s at trace 10,000. In summary, with several iterations corresponding to an order of magnitude of the standard RTM computational cost, we obtain LSRTM images that exhibit more focused diffractions and delineated structures, as shown by the depth slices in Figure 3.

DISCUSSION

The main advantage of our method when compared with conventional LSRTM is accelerated convergence to an optimal solution, as shown in Figures 1c and 2c. The energy imaging better recovers the reflector amplitudes and scattered wavefields, resulting in a closer match between the observed and simulated data. Second-order factors accountable for the decrease of the objective function include the mitigation of imaging artifacts and sharper focusing of events, and both of these improvements are similar between the energy (Figure 3e and 3f) and conventional (Figure 3c and 3d) LSRTM images.

In terms of computational cost, the energy LSRTM does not add significant burden. During extrapolation of source and receiver wavefields, the same finite-difference differential operators applied at each time step of wave extrapolation can also be applied to compute the terms required for the image (equation 10) or for the linearized source term (equation 18). In case these operators cannot be efficiently applied to the wavefields (e.g., insufficient memory storage for additional wavefield variables), the Laplacian-based imaging operators, although only adequate for image regions away from source and receiver locations, serve as good substitutes as seen by the convergence plots in Figures 1c and 2c.

We empirically observe that the preconditioning based on the matching filters benefits the first two iterations only (Figure 2c). This effect can be explained analytically by the fact that the least-squares solution of the filter estimation problem is

$$\hat{\mathbf{B}} = (\mathbf{M}_1^{\mathrm{T}}\mathbf{M}_1)^{-1}\mathbf{M}_1^{\mathrm{T}}\mathbf{M}_0, \tag{30}$$

where

$$\mathbf{M}_1 = \mathbf{L}^{\mathrm{T}}\mathbf{L}\mathbf{M}_0 \tag{31}$$

is the nonstationary convolutional operator using the entries of $\mathbf{m}_1$. Similarly, $\mathbf{M}_0$ is the corresponding convolutional operator for $\mathbf{m}_0$. Substituting equation 31 into equation 30 yields

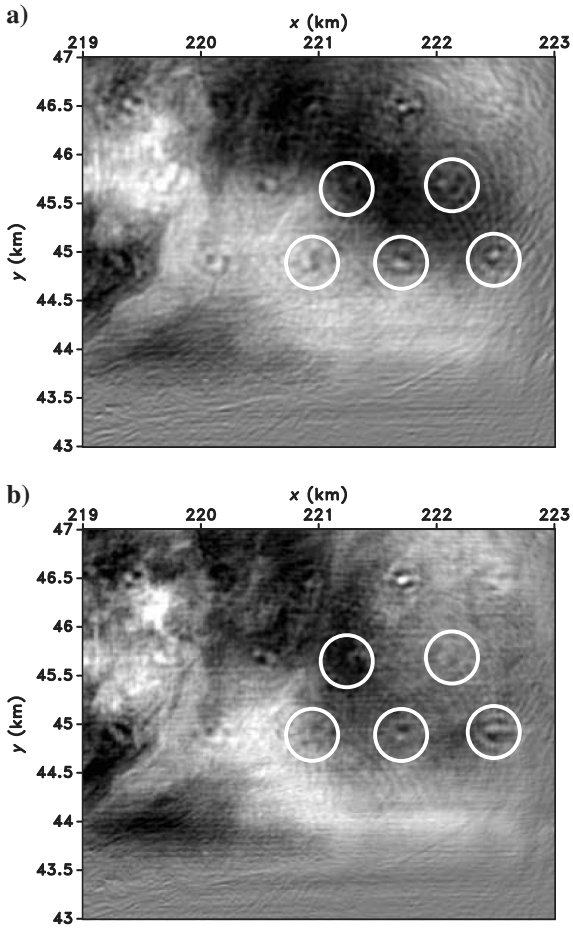


Figure 6. The GOM 3D data set: depth slices for (a) RTM and (b) LSRTM at the water bottom ($z = 1.20$ km). Note the artifact suppression in some locations indicated by the white circles.

$$\hat{\mathbf{B}} = (\mathbf{M}_0^{\mathrm{T}}\mathbf{L}^{\mathrm{T}}\mathbf{L}\mathbf{L}^{\mathrm{T}}\mathbf{L}\mathbf{M}_0)^{-1}\mathbf{M}_0^{\mathrm{T}}\mathbf{L}^{\mathrm{T}}\mathbf{L}\mathbf{M}_0, \quad (32)$$

which shows that the filters can be seen as a least-squares estimate of the inverse Hessian operator $(\mathbf{L}^{\mathrm{T}}\mathbf{L})^{-1}$, left and right multiplied by the convolutional operator $\mathbf{M}_0$. There is, therefore, an influence of the starting image $\mathbf{m}_0$ to the approximate inverse Hessian. In effect, the matching filters are a very good approximation of the operator taking us between the vectors $\mathbf{m}_0$ and $\mathbf{m}_1$ of the Krylov subspace generated by $\mathbf{L}^{\mathrm{T}}\mathbf{L}$ only (and used in the first two iterations of the conjugate gradient method). Beyond these two iterations, the matching filters do not provide any benefit anymore and switching to a nonpreconditioned inversion proves the most effective strategy (as illustrated in Figure 2c). Guitton (2017) also exemplifies this point on 3D synthetic and field data cases. As possible improvements, but not tested here, one could reestimate the matching filters as the iterations go on or find different reference images $\mathbf{m}_0$ and $\mathbf{m}_1$ for the filter estimation step.

Alternatively, our LSRTM preconditioning with matching filters can be implemented using point-spread functions (PSFs) (Fletcher et al., 2016). In our case, this approach involves creating a reference image $\mathbf{m}_0$ with spikes regularly spaced and then computing $\mathbf{m}_1$ after linearized modeling and migration of the reference image with spikes ($\mathbf{m}_0$). However, spike spacing parameterization involves an inherent trade-off between fine and coarse sampling. Too much fine sampling causes an overlap of the blurred events, thus undermining our ability to accurately extract the PSFs, and too much coarse sampling does not capture enough of the subsurface structure details. In addition, this approach requires a robust interpolation of the matching filter for image samples away from the spikes. Therefore, we do not use this alternative approach for the computation of the matching filters.

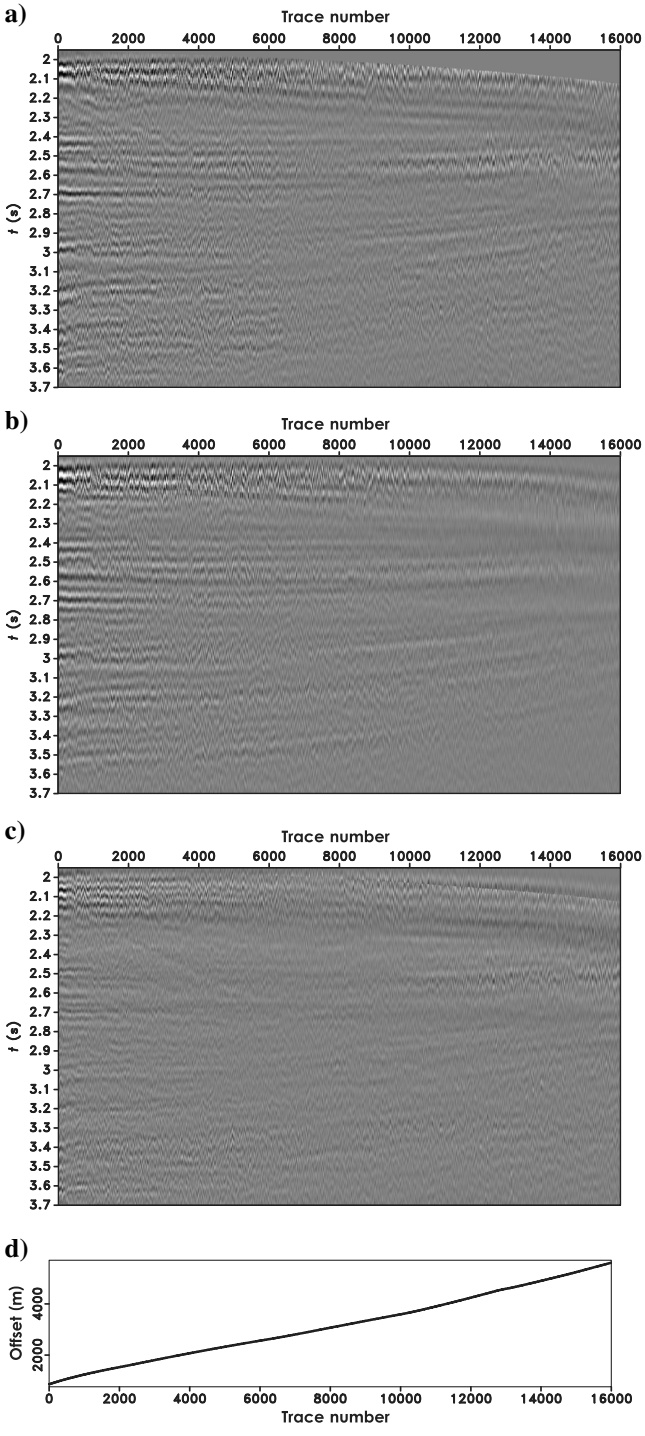


Figure 7. The GOM 3D data set: node gather at $x = 214.5$ km and $y = 44.9$ km sorted by increasing offset: (a) observed, final (b) predicted and (c) residual data. (d) Offset values for each trace. Note that the main reflection events are predicted and eliminated in the data residual, especially for events at near- and mid-offsets. Linear moveout is applied on these plots for display purposes.

CONCLUSION

We demonstrate that using proper linearized modeling and migration operators expedites LSRTM, which otherwise suffers from high computational cost. We test modeling and migration operators based on the energy norm, and we obtain faster convergence rates for LSRTM inversion because the energy operators attenuate artifacts that do not properly characterize subsurface reflectivity. In addition, a preconditioning operator that uses a multidimensional nonstationary matching filter decreases the objective function substantially at the first iterations, allowing a significant speed-up for the following iterations. Our field data examples show significant image quality improvement within less than 10 inversion iterations using an accelerated LSRTM compared with regular RTM.

ACKNOWLEDGMENTS

We thank the sponsors of the Center for Wave Phenomena, whose support made this research possible. We are grateful to the Shell Exploration and Production Company for sharing the 3D GOM data set with Colorado School of Mines and their permission to publish the results using this data set. The authors appreciate the constructive comments from the anonymous reviewers of this paper. The reproducible numeric examples in this paper use the Madagascar open-source software package (Fomel et al., 2013) available at http://www.ahay.org.

APPENDIX A

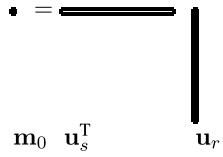
CONVENTIONAL MIGRATION AND MODELING OPERATORS

For source and receiver wavefields $\mathbf{u}_s(\mathbf{x}, t)$ and $\mathbf{u}_r(\mathbf{x}, t)$, an image $\mathbf{m}_0(\mathbf{x})$ is conventionally obtained by

Downloaded 03/21/18 to 138.67.129.87. Redistribution subject to SEG license or copyright; see Terms of Use at http://library.seg.org/

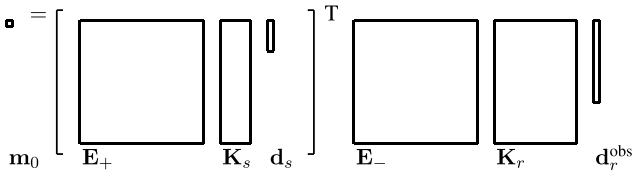
$$\mathbf{m}_0 = \mathbf{u}_s^{\mathrm{T}}\mathbf{u}_r. \tag{A-1}$$

We can represent equation A-1 and the following equations pictorially using matrices, which indicate the relative dimensions of operators and variables:



By using forward and backward extrapolator operators $\mathbf{E}_+(\mathbf{x}, t)$ and $\mathbf{E}_-(\mathbf{x}, t)$, and source and receiver injection operators $\mathbf{K_s}(\mathbf{x}, \mathbf{x}_s, t)$ and $\mathbf{K_r}(\mathbf{x}, \mathbf{x}_r, t)$, we have

$$\mathbf{m}_0 = (\mathbf{E}_+\mathbf{K}_s\mathbf{d}_s)^{\mathrm{T}}\mathbf{E}_-\mathbf{K}_r\mathbf{d}_r^{\mathrm{obs}}, \tag{A-2}$$



where $\mathbf{d}_s(\mathbf{x}_s, t)$ and $\mathbf{d}_r^{\mathrm{obs}}(\mathbf{x}_r, t)$ are the source and receiver data, respectively. In compact form, we can rewrite equation A-2 as

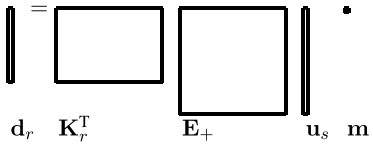
$$\mathbf{m}_0 = \mathbf{L}^{\mathrm{T}}\mathbf{d}_r^{\mathrm{obs}}. \tag{A-3}$$

Linearized modeling is defined as

$$\mathbf{d}_r = \mathbf{L}\mathbf{m}. \tag{A-4}$$

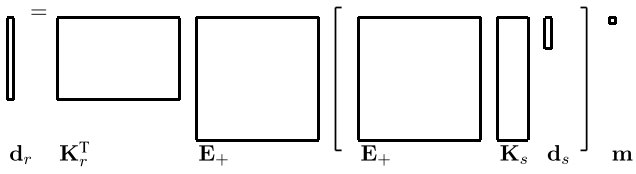
Based on equations A-2 and A-3, we can rewrite equation A-4 as

$$\mathbf{d}_r = \mathbf{K}_r^{\mathrm{T}}\mathbf{E}_-^{\mathrm{T}}\mathbf{u}_s\mathbf{m}, \tag{A-5}$$



or

$$\mathbf{d}_r = \mathbf{K}_r^{\mathrm{T}}\mathbf{E}_+(\mathbf{E}_+\mathbf{K}_s\mathbf{d}_s)\mathbf{m}, \tag{A-6}$$



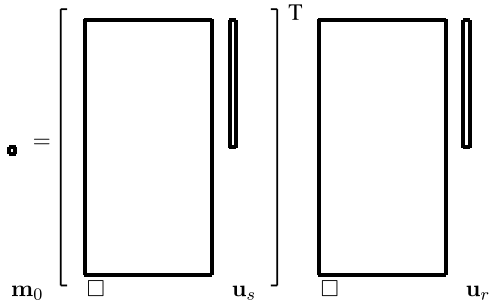
where $\mathbf{E}_-^{\mathrm{T}} = \mathbf{E}_+$, and $\mathbf{K}_r^{\mathrm{T}}$ represents the extraction at the receiver locations (adjoint operator of the injection operator).

APPENDIX B

ENERGY MIGRATION AND MODELING OPERATORS

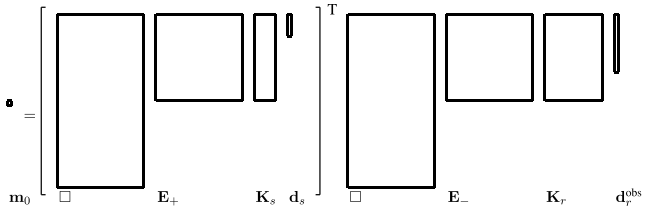
The energy image $\mathbf{m}_0(\mathbf{x})$ is obtained by

$$\mathbf{m}_0 = (\square\mathbf{u}_s)^{\mathrm{T}}\square\mathbf{u}_r. \tag{B-1}$$



The operator □ turns an acoustic wavefield into a vector field of which components are spatial and temporal. We represent this increase in dimensions pictorially by making the matrix of □ considerably larger than the wavefields matrices. Using extrapolator and injection operators, we have

$$\mathbf{m}_0 = (\square\mathbf{E}_+\mathbf{K}_s\mathbf{d}_s)^{\mathrm{T}}\square\mathbf{E}_-\mathbf{K}_r\mathbf{d}_r^{\mathrm{obs}}. \tag{B-2}$$



In compact form, we can rewrite equation B-2 as

$$\mathbf{m}_0 = \mathbf{L}^{\mathrm{T}}\mathbf{d}_r^{\mathrm{obs}}. \tag{B-3}$$

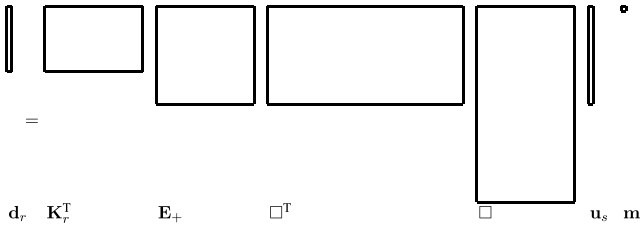
Linearized modeling is defined as

$$\mathbf{d}_r = \mathbf{L}\mathbf{m}. \tag{B-4}$$

Based on equations B-2 and B-3, we can rewrite equation B-4 as

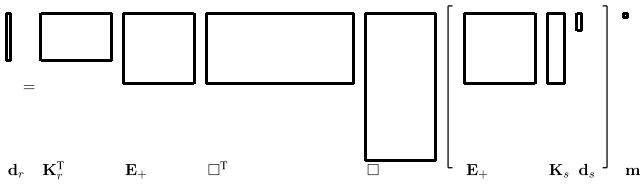
$$\mathbf{d}_r = \mathbf{K}_r^{\mathrm{T}}\mathbf{E}_-^{\mathrm{T}}\square^{\mathrm{T}}\square\mathbf{u}_s\mathbf{m}, \tag{B-5}$$

Downloaded 03/21/18 to 138.67.129.87. Redistribution subject to SEG license or copyright; see Terms of Use at http://library.seg.org/



or

$$\mathbf{d}_r = \mathbf{K}_r^{\mathrm{T}}\mathbf{E}_+\square^{\mathrm{T}}\square(\mathbf{E}_+\mathbf{K}_s\mathbf{d}_s)\mathbf{m}. \tag{B-6}$$



REFERENCES

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

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

Biondi, B., 2012, 3D seismic imaging: SEG.

Brandsberg-Dahl, S., N. Chemingui, D. Whitmore, S. Crawley, E. Klochikhina, and A. Valenciano, 2013, 3D RTM angle gathers using an inverse scattering imaging condition: 83rd Annual International Meeting, SEG, Expanded Abstracts, 3958–3962.

Chavent, G., and R.-E. Plessix, 1999, An optimal true-amplitude least-squares prestack depth-migration operator: Geophysics, **64**, 508–515, doi: 10.1190/1.1444557.

Claerbout, J. F., 1992, Earth soundings analysis: Processing versus inversion: Blackwell Scientific Publications.

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

Dai, W., and G. T. Schuster, 2012, Plane-wave least-squares reverse time migration: 82nd Annual International Meeting, SEG, Expanded Abstracts, doi: 10.1190/segam2012-0382.1.

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: 82nd Annual International Meeting, SEG, Expanded Abstracts, doi: 10.1190/segam2012-1488.1.

Douma, H., D. Yingst, I. Vasconcelos, and J. Tromp, 2010, On the connection between artifact filtering in reverse-time migration and adjoint tomography: Geophysics, **75**, no. 6, S219–S223, doi: 10.1190/1.3505124.

Dragoset, B., 1999, A practical approach to surface multiple attenuation: The Leading Edge, **18**, 104–108, doi: 10.1190/1.1438132.

Fletcher, R. P., D. Nichols, R. Bloor, and R. T. Coates, 2016, Least-squares migration: Data domain versus image domain using point spread functions: The Leading Edge, **35**, 157–162, doi: 10.1190/tle35020157.1.

Fomel, S., P. Sava, I. Vlad, Y. Liu, and V. Bashkardin, 2013, Madagascar: Open-source software project for multidimensional data analysis and reproducible computational experiments: Journal of Open Research Software, **1**, e8, doi: 10.5334/jors.ag.

Godfrey, R., P. Kristiansen, B. Armstrong, M. Copper, and E. Thorogood, 1998, Imaging the Foinaven ghost: 68th Annual International Meeting, SEG, Expanded Abstracts, 1333–1335.

Guitton, A., 2004, Amplitude and kinematic corrections of migrated images for nonunitary imaging operators: Geophysics, **69**, 1017–1024, doi: 10.1190/1.1778244.

Guitton, A., 2017, Fast 3D least-squares RTM by preconditioning with non-stationary matching filters: 87th Annual International Meeting, SEG, Expanded Abstracts, 4395–4399.

Guitton, A., G. Ayeni, and E. Daz, 2012, Constrained full-waveform inversion by model reparameterization: Geophysics, **77**, no. 2, R117–R127, doi: 10.1190/geo2011-0196.1.

Guitton, A., and G. Cambois, 1999, Multiple elimination using a pattern-recognition technique: The Leading Edge, **18**, 92–98, doi: 10.1190/1.1438166.

Hadidi, M. T., M. Sabih, D. E. Johnston, and C. Calderon-Macias, 1999, Mobil's results for the 1997 workshop on multiple attenuation: The Leading Edge, **18**, 100–103, doi: 10.1190/1.1513313.

Hestenes, M. R., and E. Stiefel, 1952, Methods of conjugate gradients for solving linear systems: Journal of Research of the National Bureau of Standards, **49**, 409–436, doi: 10.6028/jres.049.044.

Hou, J., and W. W. Symes, 2016, Accelerating extended least-squares migration with weighted conjugate gradient iteration: Geophysics, **81**, no. 4, S165–S179, doi: 10.1190/geo2015-0499.1.

Huang, W., P. Deng, and H.-W. Zhou, 2016, Least-squares reverse-time migration with hessian preconditioning: 86th Annual International Meeting, SEG, Expanded Abstracts, 1043–1046.

Kazemi, N., and M. Sacchi, 2014, Filter-based least squares wave equation shot profile migration: 76th Annual International Conference and Exhibition, EAGE, Extended Abstracts, doi: 10.3997/2214-4609.20141225.

Kazemi, N., and M. D. Sacchi, 2015, Block row recursive least-squares migration: Geophysics, **80**, no. 5, A95–A101, doi: 10.1190/geo2015-0070.1.

Kuehl, H., and M. D. Sacchi, 2003, Least-squares wave-equation migration for AVO/AVA inversion: Geophysics, **68**, 262–273, doi: 10.1190/1.1543212.

Lailly, P., 1983, The seismic inverse problem as a sequence of before stack migrations: Proceedings of the Conference on Inverse Scattering, Theory and Application, SIAM, 206–220.

Lamont, M. G., B. M. Hartley, and N. F. Uren, 1999, Multiple attenuation using the MMO and ISR preconditioning transforms: The Leading Edge, **18**, 110–114, doi: 10.1190/1.1438134.

Levin, S. A., 1984, Principle of reverse-time migration: Geophysics, **49**, 581–583, doi: 10.1190/1.1441693.

Lokshtanov, D., 1999, Multiple suppression by data-consistent deconvolution: The Leading Edge, **18**, 115–119, doi: 10.1190/1.1438136.

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

Pestana, R., A. W. G. dos Santos, and E. S. Araujo, 2013, RTM imaging condition using impedance sensitivity kernel combined with Poynting vector: Proceedings of the 13th International Congress of The Brazilian Geophysical Society.

Rickett, J., A. Guitton, and D. Gratwick, 2001, Adaptive multiple subtraction with non-stationary helical shaping filters: 63rd Annual International Conference and Exhibition, EAGE, Extended Abstracts, P167.

Rocha, D., N. Tanushev, and P. Sava, 2016, Acoustic wavefield imaging using the energy norm: Geophysics, **81**, no. 4, S151–S163, doi: 10.1190/geo2015-0486.1.

Ronen, S., L. Comeaux, and X.-G. Miao, 2005, Imaging downgoing waves from ocean bottom stations: 75th Annual International Meeting, SEG, Expanded Abstracts, 963–967.

Scales, J., 1987, Tomographic inversion via the conjugate gradient method: Geophysics, **52**, 179–185, doi: 10.1190/1.1442293.

Sun, J., Y. Zhang, S. Gray, C. Notfors, and J. Young, 2003, 3-D prestack depth migration by wave-field extrapolation methods: 16th Geophysical Conference, ASEG, Extended Abstracts, 1–4.

Tang, Y., 2009, Target-oriented wave-equation least-squares migration/inversion with phase-encoded Hessian: Geophysics, **74**, no. 6, WCA95–WCA107, doi: 10.1190/1.3204768.

Verschuur, D. J., and R. J. Prein, 1999, Multiple removal results from Delft University: The Leading Edge, **18**, 86–91, doi: 10.1190/1.1438164.

Whitmore, N. D., and S. Crawley, 2012, Application of RTM inverse scattering imaging conditions: 82nd Annual International Meeting, SEG, Expanded Abstracts, doi: 10.1190/segam2012-0779.1.

Wong, M., B. L. Biondi, and S. Ronen, 2010, Joint least-squares inversion of up- and down-going signal for ocean bottom data sets: 80th Annual International Meeting, SEG, Expanded Abstracts, 2752–2756.

Wong, M., B. L. Biondi, and S. Ronen, 2015, Imaging with primaries and free-surface multiples by joint least-squares reverse time migration: Geophysics, **80**, no. 6, S223–S235, doi: 10.1190/geo2015-0093.1.

Yao, G., and H. Jakubowicz, 2012, Least-squares reverse-time migration: 82nd Annual International Meeting, SEG, Expanded Abstracts, doi: 10.1190/segam2012-1425.1.

Zhang, Y., L. Duan, and Y. Xie, 2015, A stable and practical implementation of least-squares reverse time migration: Geophysics, **80**, no. 1, V23–V31, doi: 10.1190/geo2013-0461.1.

Zhang, Y., and J. Sun, 2009, Practical issues in reverse time migration: True amplitude gathers, noise removal and harmonic source encoding: First Break, **27**, 53–59, doi: 10.3997/1365-2397.2009002.

Downloaded 03/21/18 to 138.67.129.87. Redistribution subject to SEG license or copyright; see Terms of Use at http://library.seg.org/