GEOPHYSICS, VOL. 86, NO. 3 (MAY-JUNE 2021); P. S221–234, 22 FIGS.
10.1190/GEO2020-0629.1

Accelerating least-squares Kirchhoff time migration using beam methodology

Yubo Yue[1], Yujin Liu[2], and Samuel H. Gray[3]

ABSTRACT

Least-squares migration is an advanced imaging technique capable of producing images with improved spatial resolution, balanced illumination, and reduced migration artifacts; however, the prohibitive computational cost poses a great challenge for its practical application. We have incorporated the beam methodology into the implementation of Kirchhoff time modeling/migration and developed a fast common-offset least-squares Kirchhoff beam time migration (LSKBTM). Different from conventional Kirchhoff time modeling/migration in which the seismic data are modeled/migrated trace by trace, the mapping operation in Kirchhoff beam time modeling/migration is performed in terms of beam components and is performed only at sparsely sampled beam centers. Therefore, the computational cost of LSKBTM is significantly reduced in comparison with that of least-squares Kirchhoff time migration (LSKTM). In addition, based on the second-order Taylor expansion of the diffraction traveltime, we introduce a quadratic correction term into the inverse/forward local slant stacking, effectively enhancing the computational accuracy of LSKBTM. We used 2D synthetic and 3D field data examples to verify the effectiveness of our method. Our results indicate that LSKBTM can produce images comparable with those of LSKTM, but at considerably reduced computational cost.

INTRODUCTION

Seismic migration (Claerbout and Doherty, 1972; Gazdag, 1978; Schneider, 1978; Stolt, 1978; Baysal et al., 1983; Hill, 1990; Gray et al., 2001) is an important technique in seismic data processing to generate not only structural images for geologic interpretation but also material properties for reservoir characterization. Migration algorithms can be broadly categorized into depth- and time-migration methods, depending on whether they allow and fully honor spatial variant velocity models. Depth migration is capable of producing accurate subsurface images, even in areas with complex geologic structures (Etgen et al., 2009). However, its successful application strongly depends on an accurate interval velocity model, whose estimation can be time-consuming and still remains a challenge. For subsurface velocities with slight to moderate lateral variations, time migration is an efficient and robust alternative to depth migration (Alkhalifah, 2003, 2005). It is fast and less sensitive to velocity errors compared to depth migration, and its migration velocity model, the root-mean-square (rms) velocity model (Dix, 1955; Yilmaz, 1979), is simple to estimate. Even today, time migration is still an attractive and widely used imaging tool in the industry.

Suppose the recorded seismic data can be predicted by a linearized modeling operator, then conventional migration methods are generally formulated as adjoints rather than as the exact inverses (Tarantola, 1984; Claerbout, 1992). They sometimes fail to provide satisfactory images with high spatial resolution and amplitude fidelity because seismic data are acquired with finite acquisition aperture, band-limited source function, and sometimes poor source-receiver sampling. To mitigate these defects, least-squares migration (LSM) has been developed as an inverse problem based on the minimization of a least-squares misfit function (Tarantola, 1984; Claerbout, 1992). The main difficulty with LSM lies in the construction of the Hessian matrix (Plessix and Mulder, 2004), which represents the second derivatives of the misfit function with respect to the model parameters. Because it is prohibitively expensive to explicitly compute and store the full Hessian matrix, LSM is commonly solved with iterative gradient-based optimization algorithms (Nemeth et al., 1999), using either a Kirchhoff engine (Nemeth et al., 1999; Duquet et al., 2000) or wave-equation propagators (Kuehl and Sacchi, 2002; Kaplan et al., 2010; Huang and

---

Manuscript received by the Editor 27 August 2020; revised manuscript received 17 January 2021; published ahead of production 6 February 2021; published online 08 April 2021.

[1]R&D Center, Bureau of Geophysical Prospecting, Zhuozhou 072751, China. E-mail: geoyyb@163.com.
[2]Beijing Research Center, Aramco Asia, Beijing 100102, China. E-mail: yujin.liu@aramcoasia.com (corresponding author).
[3]CGG, 4708 Charles Ave SW, Calgary, Alberta T2S 1N5, Canada. E-mail: sam.gray-contractor@cgg.com.

Schuster, 2012; Dai and Schuster, 2013; Aldawood et al., 2015; Hou and Symes, 2015). Recently, Hu et al. (2016), Yang et al. (2018), and Yue et al. (2019, 2021) demonstrate the effectiveness of least-squares Gaussian beam migration, providing a competitive alternative to current LSM methods. Because each iteration of LSM requires one Born modeling and one migration, the total computational cost of all the iterations of LSM is at least one order of magnitude more than that of conventional migration. Therefore, despite the ongoing improvements in computing capacity, the expense of LSM still poses a great challenge for practical-sized applications.

To provide an efficient least-squares inversion for time-domain imaging, we develop a common-offset least-squares Kirchhoff beam time migration (LSKBTM) method based on the framework of beam depth migrations (Hill, 1990, 2001; Alkhalifah, 1995; Sun et al., 2000; Nowack et al., 2003; Gray, 2005; Gray and Bleistein, 2009; Liu and Palacharla, 2011; Hu et al., 2016; Yang et al., 2018; Yue et al., 2019, 2021). By incorporating the beam methodology into Kirchhoff time forward modeling/migration (KFM/KTM), we derive the Kirchhoff beam time forward modeling/migration (KBFM/KBTM) operator, whose computation consists of two major steps: (1) beam spraying/forming via inverse/forward local slant stacking and (2) forward/inverse mapping between the beam components and the image volume. Due to the application of a quadratic traveltime approximation in spraying/forming the beam components, KBFM/KBTM basically retains the accuracy of KFM/KTM. Its computational cost is, however, greatly reduced in comparison with that of KFM/KTM because the mapping operation in KBFM/KBTM is performed only at the beam centers (Hale, 1992; Sun et al., 2000; Hill, 2001), whose sampling is far sparser than the trace sampling in a common-offset record. With the constructed KBFM/KBTM operator, we formulate a LSKBTM scheme based on the minimization of the least-squares misfit function between the modeled and the recorded data, and we use the conjugate gradient to iteratively solve for the optimal image.

The outline of this paper is as follows. First, we introduce the quadratic traveltime approximation and verify its improved accuracy over the conventional linear approximation used in conventional beam migration. Then, we present the derivations of the common-offset KBFM and KBTM operators and discuss in detail their implementation. Finally, we use 2D synthetic and 3D field data examples to test our proposed LSKBTM, and we demonstrate its comparable migration accuracy and considerably improved computational efficiency by comparing it to conventional LSKTM.

METHOD

The quadratic traveltime approximation

In KTM, the two-way diffraction traveltime t_{sd} of a seismic wave traveling from a source at $\mathbf{r}_s = (x_s, y_s, 0)$, through a subsurface image point at $\mathbf{r} = (x_0, y_0, t_0)$ and to a receiver at $\mathbf{r}_d = (x_d, y_d, 0)$, is commonly evaluated by a double-square-root (DSR) equation based on the straight ray assumption (Yilmaz and Claerbout, 1980; Alkhalifah, 2006):

$$t_{sd} = \sqrt{\frac{t_0^2}{4} + \frac{(x_s - x_0)^2 + (y_s - y_0)^2}{V^2}} + \sqrt{\frac{t_0^2}{4} + \frac{(x_d - x_0)^2 + (y_d - y_0)^2}{V^2}}, \tag{1}$$

where the first and the second terms on the right side of equation 1 are the one-way traveltimes t_s and t_d from $\mathbf{r}_s$ and $\mathbf{r}_d$ to $\mathbf{r}$, respectively; V is the corresponding rms velocity at $\mathbf{r}$; and t_0 is the two-way vertical traveltime, representing the vertical axis of the image volume. Assume that there is another slightly perturbed source-receiver pair $(\mathbf{r}'_s, \mathbf{r}'_d)$ in the vicinity of $(\mathbf{r}_s, \mathbf{r}_d)$ with identical source and receiver coordinate shifts of $\Delta\mathbf{r}' = \mathbf{r}'_s - \mathbf{r}_s = \mathbf{r}'_d - \mathbf{r}_d$, then the traveltime t'_{sd} along the raypath connecting $\mathbf{r}'_s$, $\mathbf{r}$, and $\mathbf{r}'_d$ can be approximated by the Taylor expansion of t_{sd}. Instead of keeping only the linear terms in the Taylor expansion, we incorporate the second-order terms to improve accuracy and arrive at a quadratic traveltime approximation formula (see Appendix A):

$$t'_{sd} \approx t_s + t_d + \mathbf{p}_m \cdot \Delta\mathbf{r}' + \frac{1}{2}\Delta\mathbf{r}' \cdot (\mathbf{H} \cdot \Delta\mathbf{r}'), \tag{2}$$

where $\mathbf{p}_m$ represents the gradient vector of t_{sd} and is also referred to as the midpoint ray parameter

$$\mathbf{p}_m = (p_{mx}, p_{my}) = \left(\frac{\partial t_s}{\partial x_s} + \frac{\partial t_d}{\partial x_d}, \frac{\partial t_s}{\partial y_s} + \frac{\partial t_d}{\partial y_d}\right), \tag{3}$$

and $\mathbf{H}$ is a square matrix of second-order partial derivatives of traveltimes, also referred to as the Hessian matrix:

$$\mathbf{H} = \begin{pmatrix} \frac{\partial^2 t_s}{\partial x_s^2} + \frac{\partial^2 t_d}{\partial x_d^2} & \frac{\partial^2 t_s}{\partial x_s y_s} + \frac{\partial^2 t_d}{\partial x_d \partial y_d} \\ \frac{\partial^2 t_s}{\partial x_s y_s} + \frac{\partial^2 t_d}{\partial x_d \partial y_d} & \frac{\partial^2 t_s}{\partial y_s^2} + \frac{\partial^2 t_d}{\partial y_d^2} \end{pmatrix}, \tag{4}$$

whose elements are derived in Appendix A. Note that this Hessian matrix is different from the LSM Hessian mentioned previously.

We use a simple 2D model with a constant velocity of 2000 m/s to verify the improvements of the quadratic traveltime approximation over the linear one. We first compute a reference traveltime map of a given source-receiver pair using equation 1, and then we use the Taylor expansion of the reference traveltimes to approximate the traveltime map of a nearby source-receiver pair with a horizontal shift of 100 m. Figure 1 displays the absolute errors between the reference and the linearly and quadratically approximated traveltime maps for the source-receiver pairs with offset = 0.0 km (Figure 1a and 1b) and offset = 1.6 km (Figure 1c and 1d), respectively. It is evident that the quadratic approximation significantly improves the accuracy of the approximated traveltimes in comparison with the linear approximation. Meanwhile, it is notable that both approximation methods have relatively larger errors in the near source and receiver regions where the wavefronts have high curvatures. To quantify the magnitude of the errors, we extract the horizontal traces of the error maps at $t_0 = 100$ ms and compare them in Figure 2a and 2b. As can be seen, the maximum error of the quadratic approximation (the blue line) is only approximately 3 ms, which is acceptable at such a shallow depth and far less than that of the linear approximation (the red line). We further compare the vertical traces of the error maps at distance = 2.0 km in Figure 2c and 2d. It is clear that the quadratic approximation has not only smaller but also more rapidly decreasing errors compared to the linear approximation. These comparisons demonstrated the effectiveness of the quadratic traveltime approximation, which will be used in the following sections to derive the forward modeling (KBFM) and migration (KBTM) operators for LSKBTM.

Least-squares Kirchhoff time migration

LSM aims to find an optimal subsurface reflectivity model that best predicts the recorded seismic data. It is commonly implemented with gradient-based optimization methods (Nemeth et al., 1999; Kaplan et al., 2010) to iteratively minimize the L2 norm of the residual between the modeled and the recorded data:

$$J = \frac{1}{2}\|\mathbf{Lm} - \mathbf{d}\|^2, \tag{5}$$

where **d** is the recorded data, **m** is the reflectivity, and **L** is a forward-modeling operator whose adjoint is denoted as $\mathbf{L}^T$. In this work, we are using the Kirchhoff-type modeling/migration operator to estimate the time-domain reflectivity model.

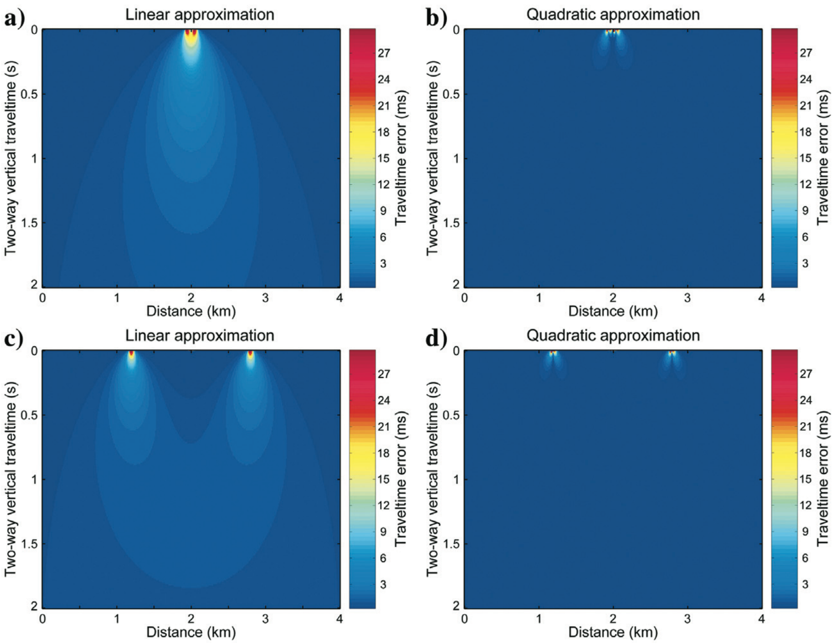

Figure 1. Absolute errors between the reference traveltime map of a given source-receiver pair and the approximated traveltime map of a slightly perturbed source-receiver pair in a constant-velocity model (the maps are aligned to make the source-receiver pairs coincide), where the distance label in the figures denotes the horizontal distances from the image points to the origin. Images (a and b) are for the source-receiver pair with offset = 0.0 km, whereas images (c and d) are for the source-receiver pair with offset = 1.6 km. Note the reduced errors of the quadratic traveltime approximation in (b and d) in comparison with the linear approximation in (a and c).

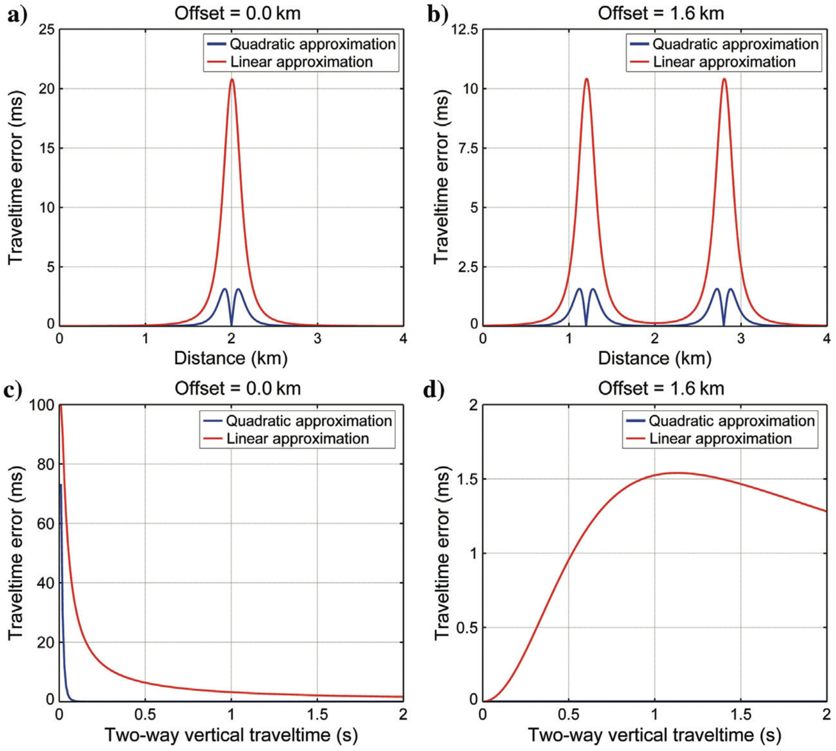

Figure 2. The extracted horizontal traces of the error maps at $t_0 = 100$ms in (a and b), and the extracted vertical traces of the error maps at distance = 2.0 km in (c and d), where images (a and c) are for the source-receiver pairs with offset = 0.0 km, and images (b and d) are for the source-receiver pairs with offset = 1.6 km, respectively. It is clear that the quadratic traveltime approximation (the blue line) not only has smaller errors, but it also has higher decay rates than the linear approximation (the red line). Also note in (d) that the blue line visually coincides with the horizontal axis because the traveltime errors of the quadratic approximation are very close to zero.

Following Bleistein et al. (2001), the common-offset KFM operator can be expressed as a volume integral:

$$D(\mathbf{h},\mathbf{r}_m,\omega) = -i\omega\int_{\Omega_0} d\mathbf{r}_0 R(\mathbf{r}_0)F(\omega)G(\mathbf{r}_0,\mathbf{r}_m-\mathbf{h},\omega) \times G(\mathbf{r}_0,\mathbf{r}_m+\mathbf{h},\omega), \tag{6}$$

where Ω_0 represents the volume of subsurface scattering points; $D(\mathbf{h},\mathbf{r}_m,\omega)$ is the modeled common-offset record at midpoint $\mathbf{r}_m=(\mathbf{r}_s+\mathbf{r}_d)/2$ and with offset $\mathbf{h}=(\mathbf{r}_d-\mathbf{r}_s)/2$; $F(\omega)$ is the source spectrum; $R(\mathbf{r}_0)$ is the reflectivity function at scattering point $\mathbf{r}_0$ (we omit its angle dependency for simplicity); and $G(\mathbf{r}_0,\mathbf{r}',\omega)$ is the Green's function corresponding to a source at $\mathbf{r}'$ and an observation point at $\mathbf{r}_0$, and its high-frequency asymptotic representation (Červený, 2001) is given by

$$G(\mathbf{r}_0,\mathbf{r}',\omega) = A(\mathbf{r}_0,\mathbf{r}')\exp[i\omega T(\mathbf{r}_0,\mathbf{r}')], \tag{7}$$

where $T(\mathbf{r}_0,\mathbf{r}')$ is the one-way traveltime defined in equation 1 and $A(\mathbf{r}_0,\mathbf{r}')$ is the amplitude, whose expression for time migration is given by Zhang et al. (2000). By inserting equation 7 into 6, we have the KFM operator expressed as

$$D(\mathbf{h},\mathbf{r}_m,\omega) = -i\omega\int_{\Omega_0} d\mathbf{r}_0 R(\mathbf{r}_0)F(\omega)A(\mathbf{r}_0,\mathbf{r}_m+\mathbf{h})A(\mathbf{r}_0,\mathbf{r}_m-\mathbf{h}) \times \exp\{i\omega[T(\mathbf{r}_0,\mathbf{r}_m+\mathbf{h})+T(\mathbf{r}_0,\mathbf{r}_m-\mathbf{h})]\}, \tag{8}$$

and the adjoint KTM operator expressed as

$$\hat{R}(\mathbf{r}) = 4\int d\omega\int d\mathbf{h}\int d\mathbf{r}_m i\omega F^*(\omega)A^*(\mathbf{r},\mathbf{r}_m+\mathbf{h})A^*(\mathbf{r},\mathbf{r}_m-\mathbf{h}) \times \exp\{-i\omega[T(\mathbf{r},\mathbf{r}_m+\mathbf{h})+T(\mathbf{r},\mathbf{r}_m-\mathbf{h})]\}\hat{D}(\mathbf{h},\mathbf{r}_m,\omega), \tag{9}$$

where $\hat{R}(\mathbf{r})$ is the migrated image at the image point $\mathbf{r}$ and the symbol $*$ denotes complex conjugation. It is notable that the amplitudes and traveltimes in equations 8 and 9 are real quantities in KFM and KTM.

It is obvious that the computational cost of KFM/KTM is proportional to the number of seismic traces. In the following sections, we will incorporate the beam methodology into the computation of KFM/KTM and derive the KBFM/KBTM operator to accelerate LSKTM.

The KBFM operator

To derive the KBFM operator, we first apply a partition of unity (Hill, 2001; Gray and Bleistein, 2009) to the KFM formula in equation 8 to divide the common-offset record into small subsets, and then we use the quadratic Taylor expansion of the diffraction traveltimes of the trace at each subset center (also known as the beam center) to approximate the traveltimes of nearby traces. After the derivations in Appendix B, we obtain the KBFM formula expressed as a modified inverse local slant stacking (ILSS) with a quadratic correction term in equation 2:

$$d(\mathbf{h},\mathbf{r}_m,t) = -\Phi\sum_{\mathbf{L}}\int d\mathbf{p}_m\gamma(\Delta\mathbf{r}_m,t) \otimes s\left(\mathbf{L},\mathbf{p}_m,t-\mathbf{p}_m\cdot\Delta\mathbf{r}_m-\frac{1}{2}\Delta\mathbf{r}_m\cdot(\mathbf{H}\cdot\Delta\mathbf{r}_m)\right), \tag{10}$$

where $d(\mathbf{h},\mathbf{r}_m,t)$ is the modeled time-domain common-offset record, $\otimes$ denotes the temporal convolution, Φ is a constant term defined in Appendix B, $\mathbf{L}$ denotes the sparsely sampled beam centers with spacing of $\Delta\mathbf{L}=(\Delta L_x,\Delta L_y)$, $\Delta\mathbf{r}_m=(\mathbf{r}_m-\mathbf{L})$ is the vector from the beam center $\mathbf{L}$ to the midpoint $\mathbf{r}_m$, $\gamma(\Delta\mathbf{r}_m,t)$ is the time-domain response of the Gaussian taper function in equation B-1, and $s(\mathbf{L},\mathbf{p}_m,t)$ is the set of beam components synthesized at each beam center:

$$s(\mathbf{L},\mathbf{p}_m,t) = f(t)\otimes\int_{\Omega_0'(\mathbf{p}_m)} d\mathbf{r}_0 R(\mathbf{r}_0)A(\mathbf{r}_0,\mathbf{L}+\mathbf{h})A(\mathbf{r}_0,\mathbf{L}-\mathbf{h}) \times\delta(t-T(\mathbf{r}_0,\mathbf{L}+\mathbf{h})-T(\mathbf{r}_0,\mathbf{L}-\mathbf{h})), \tag{11}$$

where δ denotes the Dirac delta function, $\Omega_0'(\mathbf{p}_m)$ is the set of scattering points struck by the ray pairs with midpoint ray parameter $\mathbf{p}_m$, and $f(t)$ is a time-domain wavelet defined as

$$f(t) = \int i\omega|\omega|F(\omega)\exp(i\omega t)d\omega. \tag{12}$$

The computational steps of KBFM in simulating a specific common-offset record can be summarized as

1) Choose the necessary modeling parameters including the beam center spacing $\Delta\mathbf{L}$, initial beam width w_0, and ray parameter spacing $\Delta\mathbf{p}_m$.
2) For each beam center $\mathbf{L}$, loop over subsurface scattering points and
    - calculate the amplitude and traveltime terms in equation 11, as well as the midpoint ray parameter $\mathbf{p}_m$
    - distribute the subsurface reflectivities onto the corresponding tau-*p* panels, using the traveltimes and amplitudes associated with each ray parameter $\mathbf{p}_m$.
3) Repeat step 2 for all the beam centers, and then convolve the tau-*p* panels with the wavelet $f(t)$ to produce the beam components $s(\mathbf{L},\mathbf{p}_m,t)$.
4) Smear the produced beam components at each beam center onto nearby traces within the Gaussian taper, along the trajectory defined by the quadratic function in equation 10, to produce the common-offset record.

The above implementation consists of two major operations: (1) beam mapping from the scattering volume to the beam components and (2) beam spraying from the beam components to the receiver wavefields. Because these two operations are performed only at sparsely sampled beam centers, the computational cost of KBFM is significantly reduced, compared to that of KFM. Appendix C presents an estimation of the speedup of KBFM over KFM given by

$$\text{Speedup} \approx \frac{\Delta L_x \Delta L_y}{2\Delta x \Delta y}, \tag{13}$$

where Δx and Δy are the inline and crossline trace spacings, respectively, and ΔL_x and ΔL_y are the inline and crossline beam center spacings, respectively. Although equation 13 is a rough estimation of the speedup factor, it is well consistent with our observations. The disk or network I/O cost may pose a bottleneck for large-scale Kirchhoff-type modeling/migration parallelly running on a PC cluster and degrade the speedup of KBFM. To solve this problem, one could refer to the multilevel decomposition algorithm proposed by Zhao et al. (2015), which makes the full-volume parallel computation of KTM highly scalable.

Adjoint KBTM

By applying similar traveltime approximation and partition strategies into the computation of KTM, we obtain the adjoint KBTM operator expressed as

$$\hat{R}(\mathbf{r}) = 4\int d\mathbf{h}\sum_{\mathbf{L}} A^*(\mathbf{r},\mathbf{L}+\mathbf{h})A^*(\mathbf{r},\mathbf{L}-\mathbf{h}) \times \hat{s}(\mathbf{L},\mathbf{p}_m,T(\mathbf{r},\mathbf{L}+\mathbf{h})+T(\mathbf{r},\mathbf{L}-\mathbf{h})), \tag{14}$$

where $\hat{R}(\mathbf{r})$ represents the stacked migration image and $\hat{s}(\mathbf{L},\mathbf{p}_m,t)$ is a modified local slant stacking with the summation path defined by the quadratic function in equation 2:

$$\hat{s}(\mathbf{L},\mathbf{p}_m,t) = \Phi\int d\mathbf{r}_m \hat{d}\left(\mathbf{h},\mathbf{r}_m,t+\mathbf{p}_m\cdot\Delta\mathbf{r}_m+\frac{1}{2}\Delta\mathbf{r}_m\cdot(\mathbf{H}\cdot\Delta\mathbf{r}_m)\right), \tag{15}$$

where $\hat{d}(\mathbf{h},\mathbf{r}_m,t)$ is a filtered version of the input common-offset record:

$$\hat{d}(\mathbf{h},\mathbf{r}_m,t) = \int d\omega i\omega^2 \mathrm{sgn}(\omega)F^*(\omega)\hat{D}(\mathbf{h},\mathbf{r}_m,\omega) \times \exp\left[-i\omega t - \left|\frac{\omega}{\omega_r}\right|\frac{|\Delta\mathbf{r}_m|^2}{2w_0^2}\right], \tag{16}$$

ω_r is the reference frequency, and $\hat{D}(\mathbf{h},\mathbf{r}_m,\omega)$ is the Fourier-domain representation of the input data.

In the implementation of KBTM, we first decompose the input data into beam components at each beam center using the modified local slant stacking in equation 15 and then we project the beam components into the image volume according to the two-way traveltimes and midpoint ray parameters. KBTM has a computational cost similar to KBFM, and its speedup over KTM is close to that of KBFM over KFM. Therefore, we can conclude that the speedup factor given in equation 13 is suitable to evaluate the speedup of LSKBTM over LSKTM.

Computational aspects

Incorporating the quadratic correction terms into the inverse/forward local slant stacking is important for the computational accuracy of KBFM/KBTM, especially at early times and shallow depths, but evaluation of these terms requires the locations of the subsurface scattering points that are unknown quantities while implementing local slant stacking. In Appendix D, we propose a method to construct an approximate table of the subsurface scattering points $\mathbf{r}(\bar{\mathbf{p}}_m,\bar{t})$, which is a function of the discrete midpoint ray parameter $\bar{\mathbf{p}}_m$ and two-way traveltime $\bar{t}$ and can be stored for repeated use in the implementation of LSKBTM.

Because the scattering points only affect the second-order term of the quadratic traveltime expansion, the above method does not result in traveltime errors that are large enough to affect the imaging accuracy, especially considering that the variations of the rms velocities are generally mild. To verify this statement, we recompute the traveltime errors for the test in Figure 1d, but using velocities deviating ±15% from the true velocities (2000 m/s) during the evaluation of the scattering points. The resultant traveltime errors for the deviated velocities of 1700 and 2300 m/s are shown in Figure 3a and 3b, respectively, and the extracted horizontal traces at $t_0 = 100$ ms and $t_0 = 200$ ms of these error maps and the error map in Figure 1d are further compared in Figure 4a and 4b, respectively. As can be seen, although slightly larger than the traveltime errors produced using the true velocities (the blue line), the traveltime errors produced using the deviated velocities (the red and green lines) are acceptable at such shallow depths (less than 1.2 ms at $t_0 = 200$ ms) and are dominant only in the near source and receiver regions that are typically rejected by angular muting during migration.

As in conventional depth-domain beam migrations, the beam width w_0, beam center spacing $\Delta\mathbf{L}$, and ray parameter spacing $\Delta\mathbf{p}_m$ are the key parameters that affect the performance of our proposed KBFM/KBTM. We basically adopt the parameter selections from Gaussian beam migration (Hill, 1990, 2001) and make some empirical modifications. For example, to match the rms velocities used in time migration, we choose the beam width w_0 as

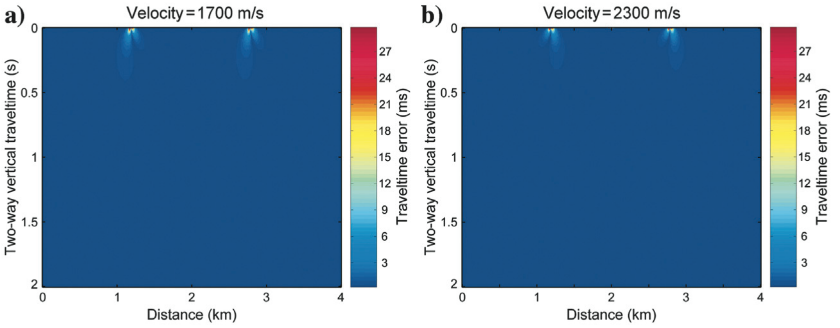

Figure 3. Absolute traveltime errors for the test in Figure 1d, but produced using velocities of (a) 1700 and (b) 2300 m/s while evaluating the scattering points. Note that the errors are dominant only in the near source (distance = 1.2 km) and receiver (distance = 2.8 km) regions.

$$w_0 = \frac{3\pi V_a}{\omega_r}, \tag{17}$$

where V_a is an average of the rms velocities of the shallow part of the velocity model instead of the whole model used in Gaussian beam migration and ω_r is chosen as the lower end of the seismic data bandwidth. In addition, we choose a slightly larger ray parameter spacing than the Hill (2001) choice to improve the computational efficiency:

$$\Delta p_x = \Delta p_x = \frac{1.5}{w_0\sqrt{\omega_r \omega_h}}, \tag{18}$$

where ω_h is the higher end of the data bandwidth.

NUMERICAL EXAMPLES

In this section, we use synthetic and field data examples to verify the effectiveness of our proposed LSKBTM. We first use a 2D layered model to test the KBFM/KBTM operator, and we demonstrate the importance of the quadratic correction term used in KBFM as

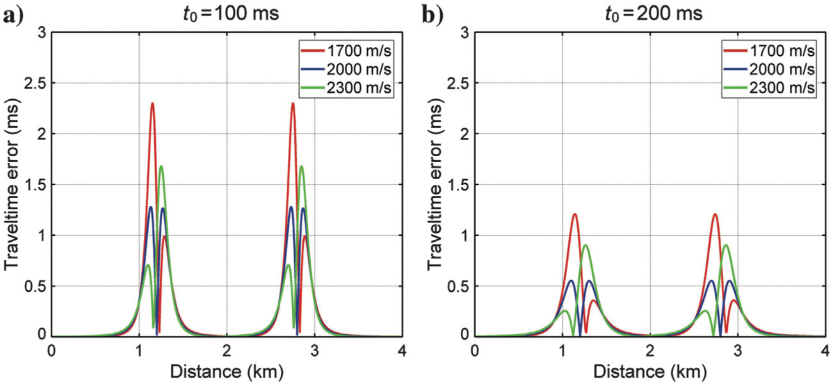

Figure 4. The extracted horizontal traces at (a) $t_0 = 100$ ms and (b) $t_0 = 200$ ms of the error maps shown in Figure 1d (the blue line), Figure 3a (the red line), and Figure 3b (the green line). Note that the errors produced using the deviated velocities (the red and green lines) are acceptable at such shallow depths and dominant only in the near source (distance = 1.2 km) and receiver (distance = 2.8 km) regions.

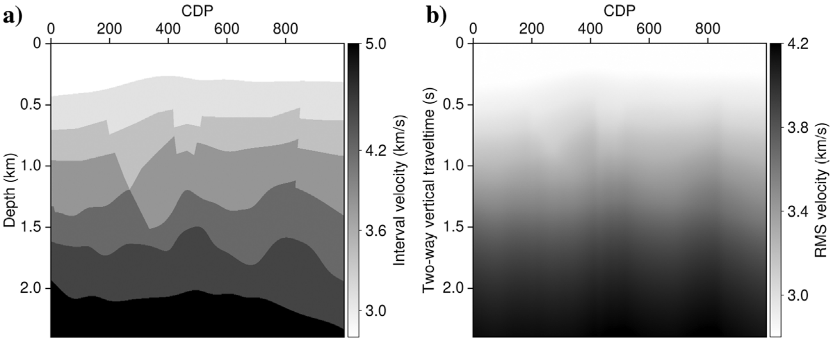

Figure 5. The 2D layered model: (a) the interval velocity model and (b) the rms velocity model.

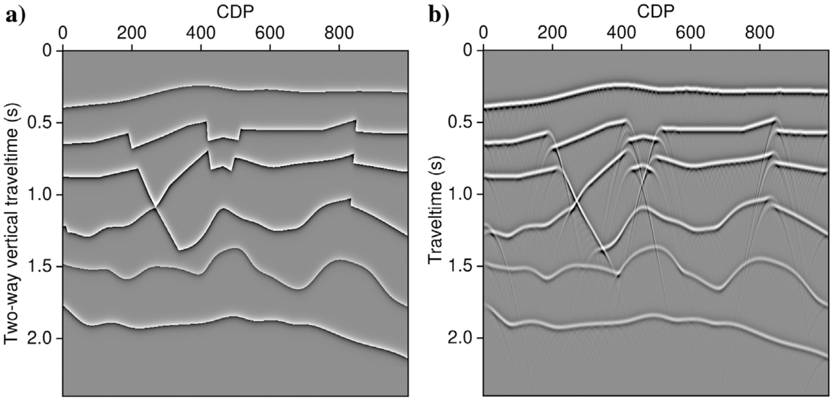

Figure 6. (a) The time-domain reflectivity model (to facilitate the image comparison, we directly use the depth-domain reflectivity model computed using the interval velocity model as the time-domain reflectivities) and (b) the reference zero-offset record simulated using KFM.

well as the enhanced efficiency of KBTM over KTM. Then, we use a modified 2D SEG/EAGE overthrust model and a 3D field data set to test LSKBTM and compare its performance with conventional LSKTM.

A 2D layered model

The interval velocity model is shown in Figure 5a, which has a grid size of 1000 × 550, with lateral and vertical spacings of 15 and 5 m, respectively. Based on the interval velocity model, we compute the rms velocity model (Figure 5b) and the subsurface reflectivities (Figure 6a) and use them as the inputs for KFM to simulate a reference zero-offset record shown in Figure 6b, whose source function is a Ricker wavelet with 20 Hz peak frequency.

To simulate the zero-offset record using KBFM, we first smear the reflectivities onto tau-*p* panels and then convolve the panels with the source function to produce the beam components. Figure 7 displays the produced beam components at six consecutive beam centers ranging from common-depth point (CDP) = 372 to CDP = 452 with spacing of 13 CDPs (195 m); each beam component has 30 dips corresponding to emergence angles between ±60°. By smearing the beam components onto nearby traces, we generate two kinds of zero-offset records, one of which is computed with the modified ILSS using the quadratic traveltime approximation and is shown in Figure 8a, and the other of which is computed with conventional ILSS using the linear traveltime approximation and is shown in Figure 8b. We can see that the record computed with the modified ILSS is comparable to the reference KFM record shown in Figure 6b, whereas the record computed with conventional ILSS is contaminated with unwanted oscillatory noise. We further compare the traces at CDP = 500 among the three records in Figure 9. As can be seen, the trace computed with the modified ILSS (the green line) is almost identical to the reference one (the red line, basically covered by the green line), whereas the trace computed with conventional ILSS (the blue line) has different amplitudes and phases, especially for the early arrivals. It is obvious that using conventional ILSS in KBFM is insufficient for accurate modeling; however, after including the quadratic correction term into conventional ILSS, the modeling accuracy is dramatically improved at early times.

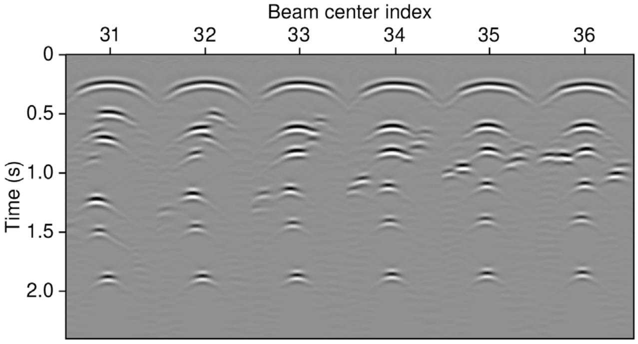

Figure 7. Beam components produced at six consecutive beam centers ranging from CDP = 362 to CDP = 452, with spacing of 13 CDPs (195 m).

We use KTM and KBTM to migrate the reference zero-offset record and show the produced images in Figure 10a and 10b, respectively. We can see that both migrations produce comparable images that accurately recover the subsurface structures, thus demonstrating the computational accuracy of KBTM. We also compare the computational costs between the two methods. Whereas it takes 16.3 s for KTM to migrate the reference zero-offset record with 1000 traces, only 2.7 s are required for KBTM to decompose and map the beam components at 82 beam centers. Therefore, KBTM achieves an approximate speedup of 6.0 in this 2D model test.

A modified SEG/EAGE overthrust model

The interval velocity model is shown in Figure 11a, which has a grid size of 640 × 374, with lateral and vertical sampling intervals of 15 and 7.5 m, respectively. Based on the interval velocity model,

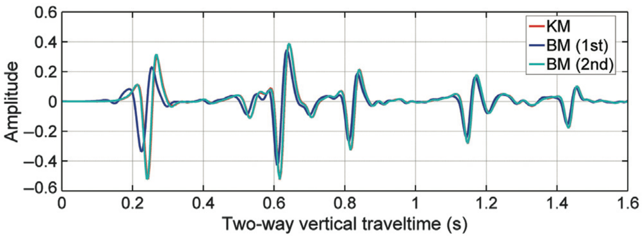

Figure 9. Trace comparison (at CDP = 500) among the reference record (the red line), the KBFM records computed with the modified ILSS (the green line), and the conventional ILSS (the blue line).

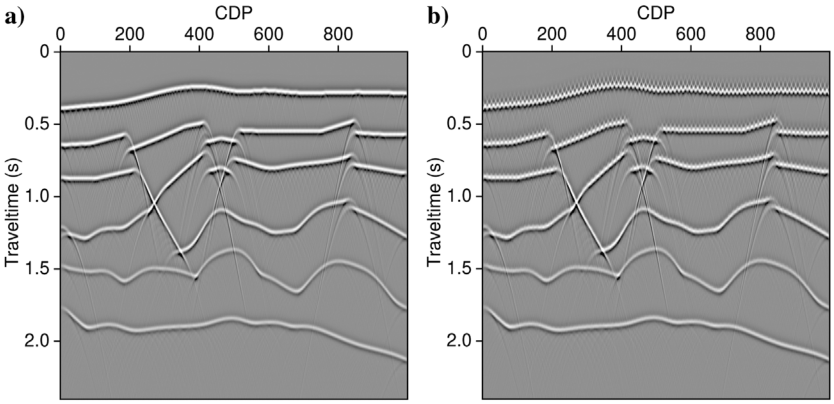

Figure 8. Zero-offset records simulated using KBFM with (a) the modified ILSS and (b) the conventional ILSS.

we first compute the rms velocity model (Figure 11b) as the imaging velocity using the Dix formula (Dix, 1955), and then we use finite-difference modeling to generate 200 shot records (Figure 12a) with shot spacing of 30 m. Each shot record is recorded by 201 uniformly spaced receivers within a 6.0 km split-spread aperture, and the source function is again a Ricker wavelet with 15 Hz peak frequency. The simulated shot records are first sorted into 101 common-offset records by using an offset bin size of 60 m, and then a partial normal moveout (NMO) correction is applied to regularize the offset values within each common-offset record. Figure 12b displays the produced common-offset record with an offset of −60 m. In this test, the beam center spacing is 210 m and the ray parameter sampling number is 37.

We migrate the common-offset records using LSKTM and LSKBTM, and then we compare their imaging difference and computational efficiency. The KTM and KBTM images after a single iteration of least-squares inversion are, respectively, shown in Figure 13a and 13b, and the traces of the produced images (the blue line for KTM and the green line for KBTM) at CDP = 380 are compared in Figure 14. KTM and KBTM produce similar images, which are hard to distinguish from each other, even in terms of the detailed waveform comparison. However, the deficiencies of adjoint migrations are also evident in both images, which not only show blurry structural imaging but also have poor amplitude fidelity. The final LSKTM and LSKBTM images after 10 iterations, as well as the trace comparison at CDP = 380, are shown in Figures 15 and 16. The two LSM images not only have comparable imaging accuracy, but they also have better resolution and more balanced amplitudes than the adjoint migration images. Figure 17 compares the convergence curves between LSKTM and LSKBTM; we can see that both LSM methods effectively minimize the data residual at similar convergence rates. However, their computational costs are

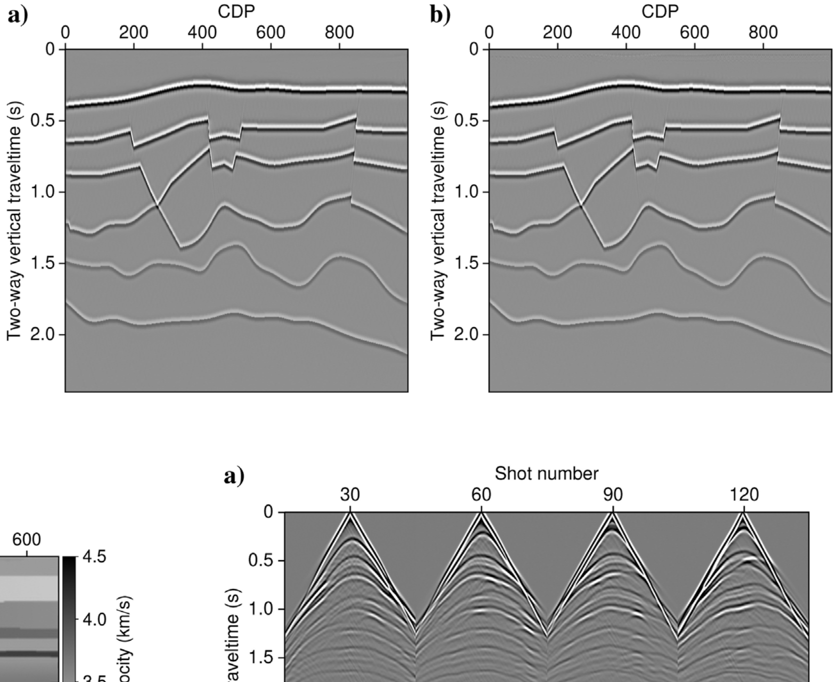

Figure 10. Migrated images using (a) adjoint KTM and (b) adjoint KBTM.

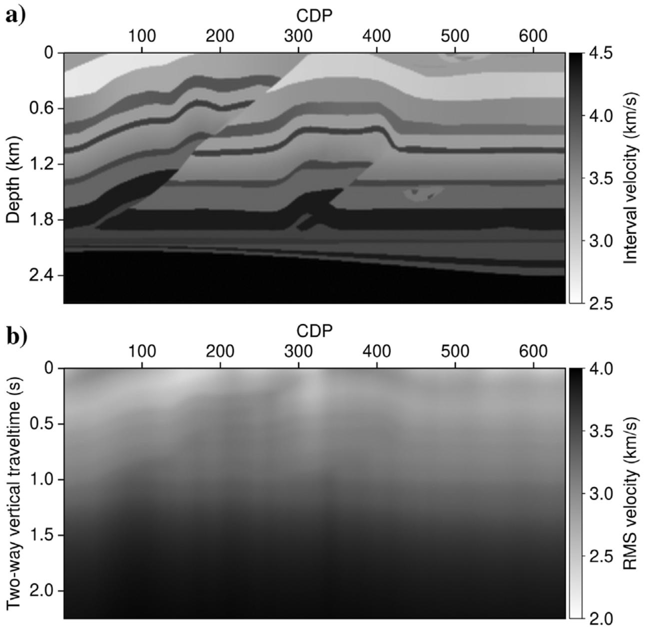

Figure 11. A modified SEG/EAGE overthrust model: (a) the interval velocity model and (b) the converted rms velocity model.

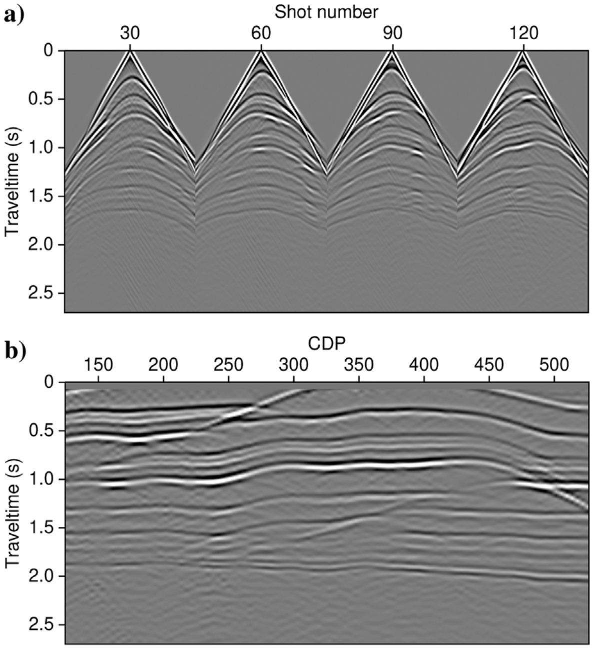

Figure 12. (a) Simulated shot gathers using finite-difference modeling and (b) regularized common-offset records with offset of −40 m after partial NMO.

quite different. It takes 383.1 s for one central processing unit (CPU: Intel i7-7700HQ, four cores and eight threads) to complete all 10 iterations of LSKTM, whereas for LSKBTM, it takes only 67.4 s. Therefore, a rough speedup of 5.7 is achieved with LSKBTM in this test.

3D field data

We finally test LSKBTM using a 3D field data example, whose rms velocity model used for imaging is shown in Figure 18a. The input data for LSKBTM are a poststack data cube after preprocessing including static correction, denoising, and NMO + stacking (Figure 18b), which is used as a zero-offset volume and has 700 inline and 400 crossline samples, with spacing of 12.5 and 25 m, respectively. In this test, the inline and crossline beam center spacings are both 230 m, the inline and crossline ray parameter sampling numbers are both 31, and the source function is a Ricker wavelet with 28 Hz peak frequency.

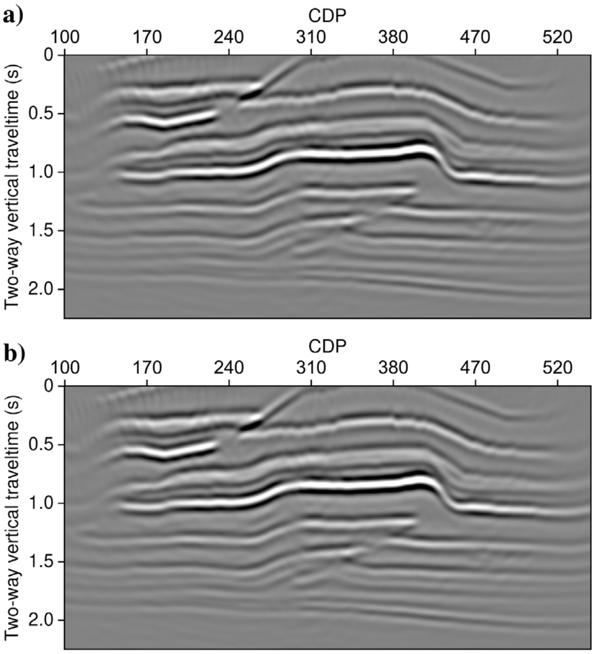

Figure 13. Migrated images using (a) the adjoint KTM and (b) the adjoint KBTM.

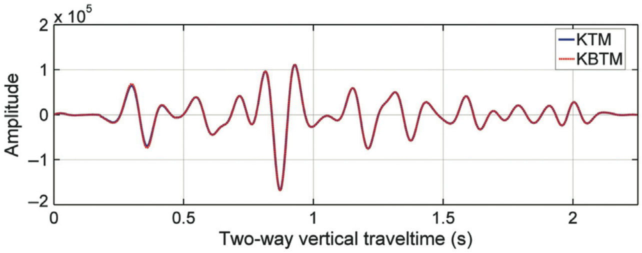

Figure 14. Trace comparison at CDP = 380 between the KTM image (the blue line) and the KBTM image (the red line).

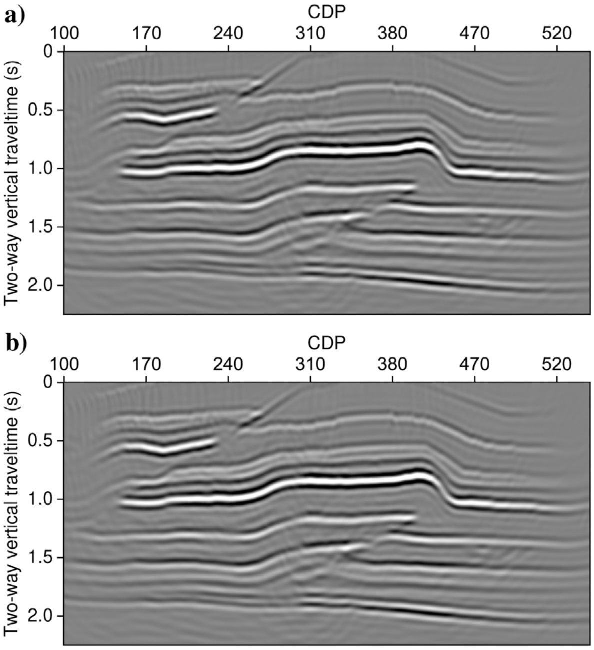

Figure 15. Migrated images using (a) LSKTM and (b) LSKBTM.

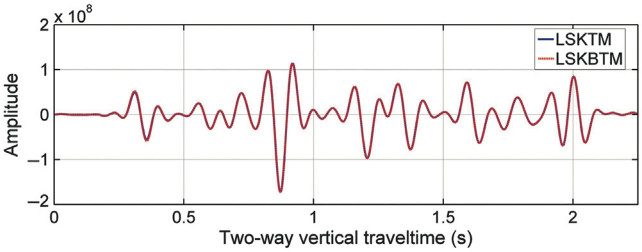

Figure 16. Trace comparison at CDP = 380 between the LSKTM image (the blue line) and the LSKBTM image (the red line).

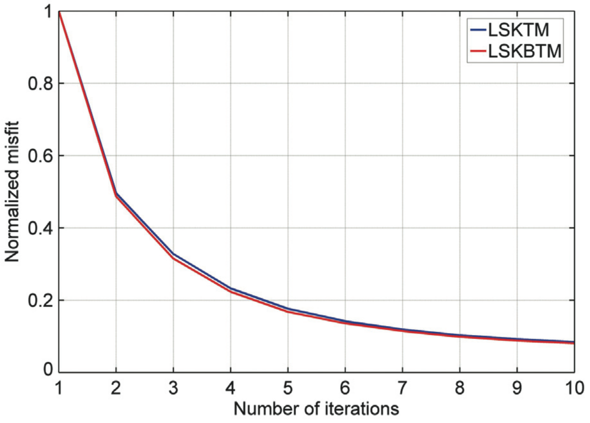

Figure 17. A comparison of convergence rates between LSKTM (the blue line) and LSKBTM (the red line).

Ten iterations are needed to achieve an effective convergence of the least-squares inversion, with the data residual decreasing by approximately 84%. The produced KBTM and LSKBTM image cubes are displayed in Figure 19a and 19b, respectively. The LSKBTM image has overall better focused and resolved events when compared to the KBTM image. Figure 20 further compares the magnified sections of crossline 240 (Figure 20a and 20b) and inline 180 (Figure 20c and 20d), as well as the time slices at 1.4 s, between the KBTM and LSKBTM images. Clear improvements of imaging quality can be observed in the sections of LSKBTM, in which not only are the fine layers and faults (the white arrows) better resolved, but the event continuity (the black arrows) is improved as well. It is also notable that the structural details are better revealed in the time slice of LSKBTM (Figure 20f) in comparison with that of KBTM (Figure 20e). The improved resolution is further verified by a spectrum comparison shown in Figure 21. It is clear that, compared with KBTM (the red line), LSKBTM (the blue line) not only broadens the image bandwidth but also increases the dominant frequency.

We also migrate the input data cube using KTM, which produces a migrated image (Figure 22) almost identical to the KBTM image shown in Figure 19a, but at a significantly increased computational cost. By using the same CPU as in the previous test, the computational time of KBTM and KTM is, respectively, 6.1 min and 8.9 h. Therefore, we can conclude that the speedup of KLSBTM over LSKTM in this 3D test is as high as 87. It is also worth mentioning that the computational time of KLSBTM is 2.1 h after all 10 iterations, only approximately one fourth of that of the adjoint KTM.

DISCUSSION

In the previous sections, we have extended beam migration to time-domain imaging and developed a fast LSKBTM for the estimation of subsurface reflectivity. Different from conventional

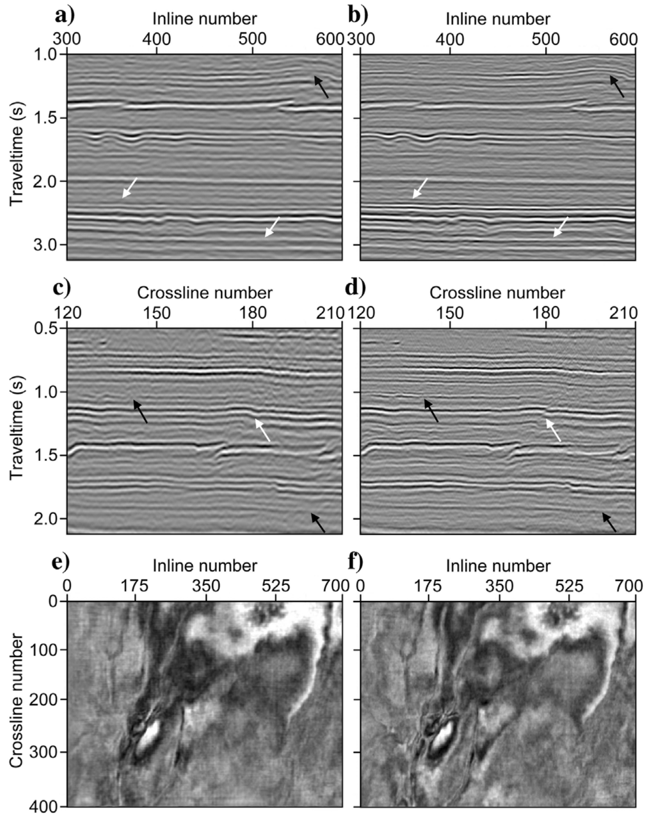

Figure 20. Detailed comparisons between the KBTM (the left column) and LSKBTM (the right column) images. (a and b) Magnified sections of crossline 240, (c and d) magnified sections of inline 180, and (e and f) time slices extracted at 1.4 s. Note the enhanced resolution (the white arrows) and event continuity (the black arrows) in the sections of LSKBTM, as well as the better revealed structural details in the time slice of LSKBTM.

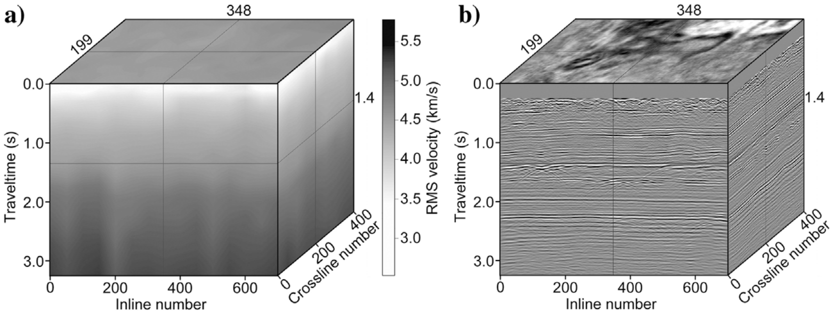

Figure 18. The 3D field data example: (a) the rms velocity model and (b) poststack data cube after NMO + stacking.

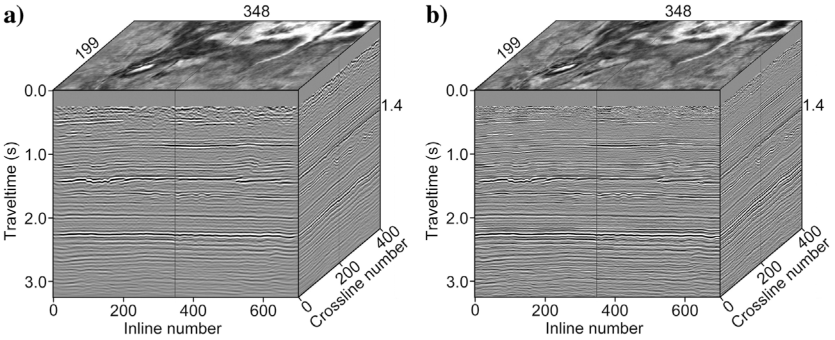

Figure 19. Migrated images using (a) KBTM and (b) LSKBTM.

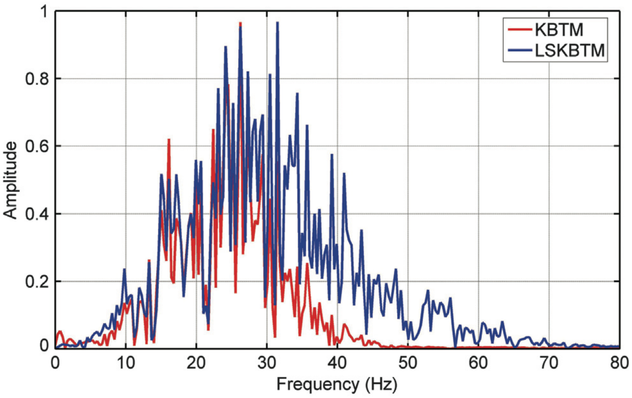

Figure 21. Spectrum comparison between KBTM (the red line) and LSKBTM (the blue line).

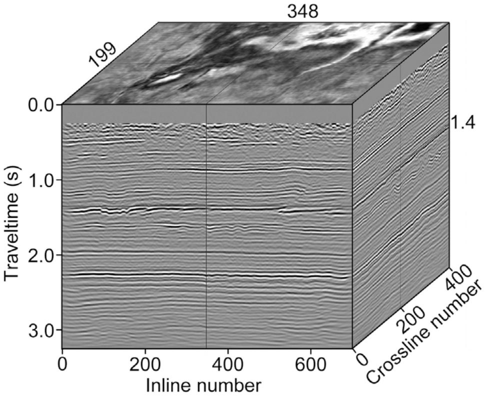

Figure 22. The KTM image.

depth-domain beam migrations (Hill, 1990, 2001; Nowack et al., 2003; Gray, 2005; Gray and Bleistein, 2009; Hu et al., 2016; Yue et al., 2019, 2021) that mainly focus on imaging multipath arrivals, our method emphasizes the ability to improve computational efficiency, which is an important requirement for iterative least-squares inversion. Although the derivations are based on the simplified DSR equation based on straight ray assumption, they can be readily modified to include high-order terms to account for ray-bending effects (Sun and Martinez, 2002; Lee et al., 2004) and anisotropy (Alkhalifah, 2006; Hao et al., 2016).

We use the quadratic traveltime approximation instead of the conventional linear approximation to derive the KBFM/KBTM operator. This leads to the quadratic correction term that is applied as part of the inverse/forward slant stacking and is vital to the computational accuracy of LSKBTM, especially for the near surface. The evaluation of the quadratic correction term in the slant stacking can be cumbersome and could potentially result in a significant increase in runtime. However, with the approximate workaround presented in Appendix D, the additional computational cost is negligible.

Because LSKBTM is formulated in the common-offset domain, the input data should ideally be single-fold and single-offset common offset volumes, with one trace per bin center in each panel. For the marine data recorded by narrow-azimuth steamers, a simple partial NMO might be sufficient to regularize the offset values within an approximate common-offset panel. However, for the land data acquired with wide-azimuth acquisition geometry, the sorted offset-vector tiles (OVTs) usually contain a range of offsets and azimuths that depend on the source and receiver line spacing. Therefore, advanced data regularization techniques such as 5D interpolation (Trad, 2009; Carozzi and Sacchi, 2019) are more suitable to regularize the OVT volumes to allow for accurate local slant stacking.

CONCLUSION

We have demonstrated the feasibility of KBFM/KBTM, derived by incorporating the beam methodology into the computation of KFM/KTM, and we developed a fast least-squares time migration scheme, referred to as LSKBTM, to provide time-migrated images with higher resolution and better amplitude fidelity than conventional KTM. Different from KFM/KTM that operates trace by trace, the mapping operation in KBFM/KBTM is performed only at sparsely sampled beam centers; therefore, the computational cost of LSKBTM is significantly reduced in comparison to that of LSKTM. Meanwhile, by including a quadratic correction term into the inverse/forward local slant stacking embedded in the KBFM/KBTM operator, LSKBTM also achieves imaging accuracy comparable to LSKTM. We have tested our proposed method using 2D synthetic and 3D field data sets, and the results show that LSKBTM is an efficient alternative to LSKTM.

ACKNOWLEDGMENTS

We are grateful to J. Shragge (editor in chief), J. Schleicher (associate editor), C. Torres-Verdin (assistant editor), T. Alkhalifah, and three anonymous reviewers for their insightful suggestions and comments. This research was supported by the National Science Foundation of China (grant no. 42074173).

DATA AND MATERIALS AVAILABILITY

Data associated with this research are available and can be obtained by contacting the corresponding author.

APPENDIX A

QUADRATIC TAYLOR EXPANSION OF THE DIFFRACTION TRAVELTIME

Based on the two-way diffraction traveltime t_{sd} of the source-receiver pair $(\mathbf{r}_s, \mathbf{r}_d)$ expressed by the DSR formula in equation 1, we can approximate the traveltime t'_{sd} of a neighboring source-receiver pair $(\mathbf{r}'_s, \mathbf{r}'_d)$ using the second-order Taylor expansion of t_{sd}, written as

$$t \approx t_s + t_d + \left(\frac{\partial t_s}{\partial x_s} + \frac{\partial t_d}{\partial x_d}\right)\Delta x' + \left(\frac{\partial t_s}{\partial y_s} + \frac{\partial t_d}{\partial y_d}\right)\Delta y' + \frac{1}{2}\left(\frac{\partial^2 t_s}{\partial x_s^2} + \frac{\partial^2 t_d}{\partial x_d^2}\right)\Delta x'^2 + \frac{1}{2}\left(\frac{\partial^2 t_s}{\partial y_s^2} + \frac{\partial^2 t_d}{\partial y_d^2}\right)\Delta y'^2 + \left(\frac{\partial^2 t_s}{\partial x_s y_s} + \frac{\partial^2 t_d}{\partial x_d \partial y_d}\right)\Delta x' \Delta \mathrm{y}', \tag{A-1}$$

where $\Delta \mathbf{r}' = (\Delta x', \Delta y')$ is a vector representing the coordinate shift between the source-receiver pairs $(\mathbf{r}_s, \mathbf{r}_d)$ and $(\mathbf{r}'_s, \mathbf{r}'_d)$ and the first-order partial derivatives are defined as

$$\frac{\partial t_s}{\partial x_s} = \frac{(x_s - x_0)}{V^2 t_s}, \; \frac{\partial t_s}{\partial y_s} = \frac{(y_s - y_0)}{V^2 t_s}, \\ \frac{\partial t_d}{\partial x_d} = \frac{(x_d - x_0)}{V^2 t_d}, \; \frac{\partial t_d}{\partial y_d} = \frac{(y_d - y_0)}{V^2 t_d}, \tag{A-2}$$

as well as the corresponding second-order partial derivatives defined as

$$\frac{\partial^2 t_s}{\partial x_s^2} = \frac{1}{V^2 t_s} - \frac{(x_s - x_0)^2}{V^4 t_s^3} = \frac{1}{V^2 t_s} - \frac{p_{sx}^2}{t_s}, \frac{\partial^2 t_s}{\partial y_s^2} = \frac{1}{V^2 t_s} - \frac{(y_s - y_0)^2}{V^4 t_s^3} = \frac{1}{V^2 t_s} - \frac{p_{sy}^2}{t_s}, \\ \frac{\partial^2 t_d}{\partial x_d^2} = \frac{1}{V^2 t_d} - \frac{(x_d - x_0)^2}{V^4 t_d^3} = \frac{1}{V^2 t_d} - \frac{p_{dx}^2}{t_d}, \frac{\partial^2 t_d}{\partial y_d^2} = \frac{1}{V^2 t_d} - \frac{(y_d - y_0)^2}{V^4 t_d^3} = \frac{1}{V^2 t_d} - \frac{p_{dy}^2}{t_d}, \\ \frac{\partial^2 t_s}{\partial x_s \partial y_s} = -\frac{(x_s - x_0)(y_s - y_0)}{V^4 t_s^3} = -\frac{p_{sx} p_{sy}}{t_s}, \frac{\partial^2 t_d}{\partial x_d \partial y_d} = -\frac{(x_d - x_0)(y_d - y_0)}{V^4 t_d^3} = -\frac{p_{dx} p_{dy}}{t_d}. \tag{A-3}$$

Note the first-order derivatives in equation A-1 are the components of the midpoint ray parameter

$$\mathbf{p}_m = (p_{mx}, p_{my}) = (p_{sx} + p_{dx}, p_{sy} + p_{dy}) \\ = \left(\frac{\partial t_s}{\partial x_s} + \frac{\partial t_d}{\partial x_d}, \frac{\partial t_s}{\partial y_s} + \frac{\partial t_d}{\partial y_d} \right), \tag{A-4}$$

and the second-order derivatives are the elements of the Hessian matrix:

$$\mathbf{H} = \begin{pmatrix} \frac{\partial^2 t_s}{\partial x_s^2} + \frac{\partial^2 t_d}{\partial x_d^2} & \frac{\partial^2 t_s}{\partial x_s \partial y_s} + \frac{\partial^2 t_d}{\partial x_d \partial y_d} \\ \frac{\partial^2 t_s}{\partial x_s \partial y_s} + \frac{\partial^2 t_d}{\partial x_d \partial y_d} & \frac{\partial^2 t_s}{\partial y_s^2} + \frac{\partial^2 t_d}{\partial y_d^2} \end{pmatrix}. \tag{A-5}$$

Hence, we can rewrite equation A-1 in matrix form as

$$t'_{sd} \approx t_s + t_d + \mathbf{p}_m \cdot \Delta \mathbf{r}' + \frac{1}{2} \Delta \mathbf{r}' \cdot (\mathbf{H} \cdot \Delta \mathbf{r}'). \tag{A-6}$$

APPENDIX B

THE KBFM OPERATOR

To derive the time-domain KBFM operator in equation 9, we first apply a partition of unity in terms of Gaussian tapers (Hill, 2001; Gray and Bleistein, 2009),

$$\frac{\sqrt{3}}{4\pi} \left| \frac{\omega}{\omega_r} \right| \left(\frac{\Delta L_x \Delta L_y}{w_0^2} \right) \sum_{\mathbf{L}} \exp\left[-\left| \frac{\omega}{\omega_r} \right| \frac{|\Delta \mathbf{r}_m|^2}{2 w_0^2} \right] \approx 1, \tag{B-1}$$

into the KFM formula expressed in equation 7, leading to

$$D(\mathbf{h}, \mathbf{r}_m, \omega) = -\Phi \sum_{\mathbf{L}} \int_{\Omega_0} d\mathbf{r}_0 R(\mathbf{r}_0) i\omega |\omega| F(\omega) A(\mathbf{r}_0, \mathbf{r}_m + \mathbf{h}) A(\mathbf{r}_0, \mathbf{r}_m - \mathbf{h}) \\ \times \exp\left\{ i\omega [T(\mathbf{r}_0, \mathbf{r}_m + \mathbf{h}) + T(\mathbf{r}_0, \mathbf{r}_m - \mathbf{h})] - \left| \frac{\omega}{\omega_r} \right| \frac{|\Delta \mathbf{r}_m|^2}{2 w_0^2} \right\}, \tag{B-2}$$

where $\Delta \mathbf{r}_m = (\mathbf{r}_m - \mathbf{L})$ denotes the spatial vector from beam center $\mathbf{L}$ to midpoint $\mathbf{r}_m$, $\Delta \mathbf{L} = (\Delta L_x, \Delta L_y)$ is the beam center spacing, ω_r is the reference frequency, w_0 is the beam width, and $\Phi = \sqrt{3}/4\pi(\Delta L_x \Delta L_x / \omega_r w_0^2)$.

Next, we approximate the amplitude and traveltime terms in equation B-2 by

$$A(\mathbf{r}_0, \mathbf{r}_m + \mathbf{h}) A(\mathbf{r}_0, \mathbf{r}_m - \mathbf{h}) \approx A(\mathbf{r}_0, \mathbf{L} + \mathbf{h}) A(\mathbf{r}_0, \mathbf{L} - \mathbf{h}), \\ T(\mathbf{r}_0, \mathbf{r}_m + \mathbf{h}) + T(\mathbf{r}_0, \mathbf{r}_m - \mathbf{h}) \approx T(\mathbf{r}_0, \mathbf{L} + \mathbf{h}) + T(\mathbf{r}_0, \mathbf{L} - \mathbf{h}) \\ + \mathbf{p}_m \cdot \Delta \mathbf{r}_m + \frac{1}{2} \Delta \mathbf{r}_m \cdot (\mathbf{H} \cdot \Delta \mathbf{r}_m), \tag{B-3}$$

and then arrive at

$$D(\mathbf{h}, \mathbf{r}_m, \omega) \approx -\Phi \sum_{\mathbf{L}} \int_{\Omega_0} d\mathbf{r}_0 R(\mathbf{r}_0) i\omega |\omega| F(\omega) A(\mathbf{r}_0, \mathbf{L} + \mathbf{h}) A(\mathbf{r}_0, \mathbf{L} - \mathbf{h}) \\ \times \exp\{ i\omega [T(\mathbf{r}_0, \mathbf{L} + \mathbf{h}) + T(\mathbf{r}_0, \mathbf{L} - \mathbf{h})] \} \\ \times \exp\left[i\omega \left(\mathbf{p}_m \cdot \Delta \mathbf{r}_m + \frac{1}{2} \Delta \mathbf{r}_m \cdot (\mathbf{H} \cdot \Delta \mathbf{r}_m) \right) - \left| \frac{\omega}{\omega_r} \right| \frac{|\Delta \mathbf{r}_m|^2}{2 w_0^2} \right]. \tag{B-4}$$

If we denote the set of subsurface scattering points struck by the ray pairs with midpoint ray parameter of $\mathbf{p}_m$ as $\Omega'_0(\mathbf{p}_m)$, then equation B-4 can be rewritten as

$$D(\mathbf{h}, \mathbf{r}_m, \omega) = -\Phi \sum_{\mathbf{L}} \int d\mathbf{p}_m S(\mathbf{L}, \mathbf{p}_m, \omega) \\ \times \exp\left[i\omega \left(\mathbf{p}_m \cdot \Delta \mathbf{r}_m + \frac{1}{2} \Delta \mathbf{r}_m \cdot (\mathbf{H} \cdot \Delta \mathbf{r}_m) \right) - \left| \frac{\omega}{\omega_r} \right| \frac{|\Delta \mathbf{r}_m|^2}{2 w_0^2} \right], \tag{B-5}$$

where

$$S(\mathbf{L}, \mathbf{p}_m, \omega) = \int_{\Omega'_0(\mathbf{p}_m)} d\mathbf{r}_0 i\omega |\omega| F(\omega) R(\mathbf{r}_0) A(\mathbf{r}_0, \mathbf{L} + \mathbf{h}) A(\mathbf{r}_0, \mathbf{L} - \mathbf{h}) \\ \times \exp\{ i\omega [T(\mathbf{r}_0, \mathbf{L} + \mathbf{h}) + T(\mathbf{r}_0, \mathbf{L} - \mathbf{h})] \}. \tag{B-6}$$

After applying the inverse Fourier transform to equations B-5 and B-6, we obtain the KBFM formulas expressed in equations 9 and 10, respectively.

APPENDIX C

ESTIMATING THE SPEEDUP OF KBFM OVER KFM

Suppose N_x, N_y, and N_τ are the grid sizes of a 3D scattering volume (modeling aperture) to compute an output trace in KFM, then the total cost to simulate a common-offset record with $N_{rx} \times N_{ry}$ traces can be approximately evaluated as

$$C_{\mathrm{KFM}} = O_{\mathrm{KM}} N_{rx} N_{ry} N_x N_y N_\tau, \tag{C-1}$$

where O_{KM} denotes the cost of an atomic level operation in KFM (projecting an input sample to an output sample).

With the same modeling configuration, the computation cost of KBFM can be evaluated as

$$\begin{aligned} C_{\mathrm{KBFM}} &= O_{\mathrm{BM}} N_{Lx} N_{Ly} N_x N_y N_\tau \\ &\quad + O_{\mathrm{BS}} N_{Lx} N_{Ly} N_g N_{px} N_{py} N_\tau, \end{aligned} \tag{C-2}$$

where the first and second terms on the right side of the equation represent the computational costs of beam mapping and beam spraying in KBFM, whose atomic level costs are denoted as O_{BM} and O_{BS}, respectively; N_{Lx} and N_{Ly} are the inline and crossline beam center sampling numbers, respectively; N_{px} and N_{py} are the ray parameter sampling numbers, whose spacings are given in equation 14; and N_g is the number of output traces corresponding to one beam component (within the Gaussian taper in equation B-1), and can be estimated as

$$N_g \approx \frac{\pi(2\Delta L_x)(2\Delta L_y)}{\Delta x \Delta y}, \tag{C-3}$$

where Δx and Δy are the inline and crossline trace spacings, respectively, and ΔL_x and ΔL_y are the inline and crossline beam center spacings, respectively.

By substituting equation C-3 into C-2 and applying the following empirical relationship between O_{KM}, O_{BM}, and O_{BS} observed in our experiments,

$$O_{\mathrm{KM}} \approx O_{\mathrm{BM}} \approx 3.5 O_{\mathrm{BS}}, \tag{C-4}$$

we express the speedup as

$$\mathrm{Speedup} = \frac{C_{\mathrm{KFM}}}{C_{\mathrm{KBFM}}} \approx \frac{N_{rx} N_{ry}}{N_{Lx} N_{Ly} + 3 N_{rx} N_{ry} N_{px} N_{py} / N_x N_y}, \tag{C-5}$$

which can be further simplified by applying the following substitution:

$$\frac{N_{rx} N_{ry}}{N_{Lx} N_{Ly}} = \frac{\Delta L_x \Delta L_y}{\Delta x \Delta y}, \tag{C-6}$$

and then it becomes

$$\mathrm{Speedup} \approx \frac{\Delta L_x \Delta L_y}{\Delta x \Delta y} \frac{1}{1+\varphi}, \tag{C-7}$$

where

$$\varphi = \frac{3 \Delta L_x \Delta L_y N_{px} N_{py}}{N_x N_y \Delta x \Delta y}, \tag{C-8}$$

and the terms $N_x \Delta x$ and $N_y \Delta y$ in the denominator are, respectively, the inline and crossline migration apertures, which are typically chosen at approximately 10 km in practice, and the typical values for N_{px} and N_{py} in the numerator are between 30 and 40 and for ΔL_x and ΔL_y are approximately 200 m. Therefore, the values of the numerator and denominator in equation C-7 are approximately the same and we can roughly rewrite the speedup as

$$\mathrm{Speedup} \approx \frac{\Delta L_x \Delta L_y}{2 \Delta x \Delta y}. \tag{C-9}$$

APPENDIX D

CONSTRUCTING THE TABLE $\mathbf{r}(\mathbf{p}_m, t)$

We present here an approximate method to construct the table of subsurface scattering points $\mathbf{r}(\bar{\mathbf{p}}_m, \bar{t})$ as a function of the discrete samples of midpoint ray parameter $\bar{\mathbf{p}}_m$ and two-way traveltime $\bar{t}$.

For a source-receiver pair with the midpoint at a beam center, we first average the rms velocities along the horizontal planes to create a laterally invariant velocity cube $V(\mathbf{r})$, whose spatial coverage is determined by the maximum migration (or modeling) aperture.

Next, for each sample of the velocity cube, we calculate the two-way traveltime table $t(\mathbf{r})$ using equation 1 as well as the ray parameter components $\mathbf{p_m}(\mathbf{r}) = (p_{mx}, p_{my})$ using

$$p_{mx} = \frac{\partial t_s}{\partial x_s} + \frac{\partial t_d}{\partial x_d}, \qquad p_{my} = \frac{\partial t_s}{\partial y_s} + \frac{\partial t_d}{\partial y_d}. \tag{D-1}$$

Then, by looping over the samples of the velocity cube, we find the active samples of $\mathbf{r}(\bar{\mathbf{p}}_m, \bar{t})$ by evaluating

$$\bar{\mathbf{p}}_m = \mathrm{NINT}\left(\frac{\mathbf{p}_m(\mathbf{r})}{d\bar{\mathbf{p}}_m}\right) d\bar{\mathbf{p}}_m, \qquad \bar{t} = \mathrm{NINT}\left(\frac{t(\mathbf{r})}{d\bar{t}}\right) d\bar{t}, \tag{D-2}$$

where NINT is the nearest integer function and $d\bar{\mathbf{p}}_m$ and $d\bar{t}$ are the sampling intervals of $\bar{\mathbf{p}}_m$ and $\bar{t}$, respectively, and then we assign $\mathbf{r}$ to $\mathbf{r}(\bar{\mathbf{p}}_m, \bar{t})$.

At last, we fill in the holes of $\mathbf{r}(\bar{\mathbf{p}}_m, \bar{t})$ using interpolation, followed by a spatial smoothing, to obtain the final table. Due to the lateral invariance of the velocity cube, once we construct a table of subsurface scattering points for one beam center, it can be repeatedly used for other beam centers as well.

REFERENCES

Aldawood, A., I. Hoteit, M. Zuberi, G. Turkiyyah, and T. Alkhalifah, 2015, The possibilities of least-squares migration of internally scattered seismic energy: Geophysics, **80**, no. 4, S93–S101, doi: 10.1190/geo2014-0436.1.

Alkhalifah, T., 1995, Gaussian beam depth migration for anisotropic media: Geophysics, **60**, 1474–1484, doi: 10.1190/1.1443881.

Alkhalifah, T., 2003, Tau migration and velocity analysis: Theory and synthetic examples: Geophysics, **68**, 1331–1339, doi: 10.1190/1.1598126.

Alkhalifah, T., 2005, τ-migration and velocity analysis: Application to data from the Red Sea: Geophysical Prospecting, **53**, 643–653, doi: 10.1111/j.1365-2478.2005.00493.x.

Alkhalifah, T., 2006, Kirchhoff time migration for transversely isotropic media: An application to Trinidad data: Geophysics, **71**, no. 1, S29–S35, doi: 10.1190/1.2159058.

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

Bleistein, N., J. K. Cohen, and J. W. John, 2001, Mathematics of multidimensional seismic imaging, migration, and inversion: Springer.

Carozzi, F., and M. D. Sacchi, 2019, Robust tensor-completion algorithm for 5D seismic-data reconstruction: Geophysics, **84**, no. 2, V97–V109, doi: 10.1190/geo2018-0109.1.

Červený, V., 2001, Seismic ray theory: Cambridge University Press.

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

Claerbout, J. F., and S. M. Doherty, 1972, Downward continuation of moveout-corrected seismograms: Geophysics, **37**, 741–768, doi: 10.1190/1.1440298.

Dai, W., and G. T. Schuster, 2013, Plane-wave least-squares reverse-time migration: Geophysics, **78**, no. 4, S165–S177, doi: 10.1190/geo2012-0377.1.

Dix, C. H., 1955, Seismic velocities from surface measurements: Geophysics, **20**, 68–86, doi: 10.1190/1.1438126.

Duquet, B., K. J. Marfurt, and J. A. Dellinger, 2000, Kirchhoff modeling, inversion for reflectivity, and subsurface illumination: Geophysics, **65**, 1195–1209, doi: 10.1190/1.1444812.

Etgen, J., S. H. Gray, and Y. Zhang, 2009, An overview of depth imaging in exploration geophysics: Geophysics, **74**, no. 6, WCA5–WCA7, doi: 10.1190/1.3223188.

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

Gray, S. H., 2005, Gaussian beam migration of common-shot records: Geophysics, **70**, no. 4, S71–S77, doi: 10.1190/1.1988186.

Gray, S. H., and N. Bleistein, 2009, True-amplitude Gaussian-beam migration: Geophysics, **74**, no. 2, S11–S23, doi: 10.1190/1.3052116.

Gray, S. H., J. Etgen, J. Dellinger, and D. Whitmore, 2001, Seismic migration problems and solutions: Geophysics, **66**, 1622–1640, doi: 10.1190/1.1487107.

Hale, D., 1992, Migration by the Kirchhoff, slant stack, and Gaussian beam methods: CWP Report 121, Colorado School of Mines.

Hao, Q., A. Stovas, and T. Alkhalifah, 2016, The offset-midpoint traveltime pyramid of P-waves in homogeneous orthorhombic media: Geophysics, **81**, no. 5, C151–C162, doi: 10.1190/geo2015-0352.1.

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

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

Hou, J., and W. W. Symes, 2015, An approximate inverse to the extended Born modeling operator: Geophysics, **80**, no. 6, R331–R349, doi: 10.1190/geo2014-0592.1.

Hu, H., Y. Liu, Y. Zheng, X. Liu, and H. Lu, 2016, Least-squares Gaussian beam migration: Geophysics, **81**, no. 3, S87–S100, doi: 10.1190/geo2015-0328.1.

Huang, Y., and G. T. Schuster, 2012, Multisource least-squares migration of marine streamer and land data with frequency-division encoding: Geophysical Prospecting, **60**, 663–680, doi: 10.1111/j.1365-2478.2012.01086.x.

Kaplan, S. T., P. S. Routh, and M. D. Sacchi, 2010, Derivation of forward and adjoint operators for least-squares shot-profile split-step migration: Geophysics, **75**, no. 6, S225–S235, doi: 10.1190/1.3506146.

Kuehl, H., and M. Sacchi, 2002, Robust AVP estimation using least-squares wave-equation migration: 72nd Annual International Meeting, SEG, Expanded Abstracts, 281–284, doi: 10.1190/1.1817231.

Lee, S., S. Lee, S. Lin, and J. Willis, 2004, Using 3D Kirchhoff time demigration as part of a cost effective depth migration process: 74th Annual International Meeting, SEG, Expanded Abstracts, 1009–1012, doi: 10.1190/1.1851061.

Liu, J., and G. Palacharla, 2011, Multiarrival Kirchhoff beam migration: Geophysics, **76**, no. 5, WB109–WB118, doi: 10.1190/geo2010-0403.1.

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.

Nowack, R. L., M. K. Sen, and P. L. Stoffa, 2003, Gaussian beam migration for sparse common-shot and common-receiver data: 73rd Annual International Meeting, SEG, Expanded Abstracts, 1114–1117, doi: 10.1190/1.1817470.

Plessix, R.-E., and W. A. Mulder, 2004, Frequency-domain finite-difference amplitude-preserving migration: Geophysical Journal International, **157**, 975–987, doi: 10.1111/j.1365-246X.2004.02282.x.

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

Stolt, R. H., 1978, Migration by Fourier transform: Geophysics, **43**, 23–48, doi: 10.1190/1.1440826.

Sun, C., and R. D. Martinez, 2002, Amplitude preserving 3D pre-stack Kirchhoff time migration for V(z) and VTI media: 72nd Annual International Meeting, SEG, Expanded Abstracts, 1224–1227, doi: 10.1190/1.1816873.

Sun, Y., F. Qin, S. Checkles, and J. P. Leveille, 2000, 3-D prestack Kirchhoff beam migration for depth imaging: Geophysics, **65**, 1592–1603, doi: 10.1190/1.1444847.

Tarantola, A., 1984, Inversion of seismic reflection data in the acoustic approximation: Geophysics, **49**, 1259–1266, doi: 10.1190/1.1441754.

Trad, D., 2009, Five-dimensional interpolation: Recovering from acquisition constraints: Geophysics, **74**, no. 6, V123–V132, doi: 10.1190/1.3245216.

Yang, J., H. Zhu, G. McMechan, and Y. Yue, 2018, Time-domain least squares migration using the Gaussian beam summation method: Geophysical Journal International, **214**, 548–572, doi: 10.1093/gji/ggy142.

Yilmaz, Ö., 1979, Seismic data processing: SEG.

Yilmaz, Ö., and J. F. Claerbout, 1980, Prestack partial migration: Geophysics, **45**, 1753–1779, doi: 10.1190/1.1441064.

Yue, Y., Y. Liu, Y. Li, and Y. Shi, 2021, Least-squares Gaussian beam migration in viscoacoustic media: Geophysics, **86**, no. 1, S17–S28, doi: 10.1190/geo2020-0129.1.

Yue, Y., P. Sava, Z. Qian, J. Yang, and Z. Zou, 2019, Least-squares Gaussian beam migration in elastic media: Geophysics, **84**, no. 4, S329–S340, doi: 10.1190/geo2018-0391.1.

Zhang, Y., S. Gray, and J. Young, 2000, Exact and approximate weights for Kirchhoff migration: 70th Annual International Meeting, SEG, Expanded Abstracts, 1036–1039, doi: 10.1190/1.1815561.

Zhao, C., S. Wang, G. Luo, J. Wen, and J. Zhang, 2015, A highly scalable parallel algorithm for 3D prestack Kirchhoff time migration: Journal of Computer Research and Development, **52**, 869–878.

Biographies and photographs of the authors are not available.