The possibilities of least-squares migration of internally scattered seismic energy

Ali Aldawood[1], Ibrahim Hoteit[1], Mohammad Zuberi[1], George Turkiyyah[2], and Tariq AlKhalifah[1]

ABSTRACT

Approximate images of the earth's subsurface structures are usually obtained by migrating surface seismic data. Least-squares migration, under the single-scattering assumption, is used as an iterative linearized inversion scheme to suppress migration artifacts, deconvolve the source signature, mitigate the acquisition fingerprint, and enhance the spatial resolution of migrated images. The problem with least-squares migration of primaries, however, is that it may not be able to enhance events that are mainly illuminated by internal multiples, such as vertical and nearly vertical faults or salt flanks. To alleviate this problem, we adopted a linearized inversion framework to migrate internally scattered energy. We apply the least-squares migration of first-order internal multiples to image subsurface vertical fault planes. Tests on synthetic data demonstrated the ability of the proposed method to resolve vertical fault planes, which are poorly illuminated by the least-squares migration of primaries only. The proposed scheme is robust in the presence of white Gaussian observational noise and in the case of imaging the fault planes using inaccurate migration velocities. Our results suggested that the proposed least-squares imaging, under the double-scattering assumption, still retrieved the vertical fault planes when imaging the scattered data despite a slight defocusing of these events due to the presence of noise or velocity errors.

INTRODUCTION

Seismic imaging using Kirchhoff-type integrals, based on the high-frequency asymptotic solution to the wave equation, is a powerful tool to map the data recorded on the surface of the earth into representative subsurface structures (Claerbout, 1992). However, a major drawback of standard Kirchhoff migration is that it mainly images reflectors that are primarily illuminated by single-scattering energy. In other words, it focuses the seismic energy that travels from a source point and bounces only once off a subsurface scattering point before it is received at a surface receiver. The reason is that the recording surface in seismic acquisition surveys does not typically enclose the medium containing the scattering points. Therefore, in the case of a complex velocity structure, some seismic energy may bounce once off a subsurface scattering point without ending up at the recording surface (Zuberi and Alkhalifah, 2014). Figure 1a shows schematically a reflected ray that does not reach a recording station after a single reflection, pointing away from the recording surface.

Imaging internal multiples has the potential to retrieve seismic events that are poorly illuminated by single-scattering energy, such as vertical and nearly vertical fault planes or salt flanks (Aldawood et al., 2014b). Internally scattered seismic energy may have a path that could end up on the recording surface, allowing the illumination of such reflectors. Figure 1b shows schematically a doubly reflected ray with a raypath that ends up on the recording surface. Malcolm et al. (2009) propose a methodology to image internal multiples, in which they image prismatic waves and other higher order internal multiples. Behura et al. (2012) propose an iterative imaging procedure to image internal scattering along with single-scattering energy. Their methodology is based on virtual source creation using inverse scattering (Wapenaar et al., 2012). Zuberi and Alkhalifah (2013) propose a three-step interferometric imaging procedure to image first-order internal scattering energy. Zuberi and Alkhalifah (2014) generalize their interferometric technique to include higher order internal multiples. Their generalized interferometric multiple imaging method simultaneously creates and images

internal multiples given that they are kinematically allowed to be in the recorded data.

Standard Kirchhoff-based, or reverse time migration (RTM)-based imaging is performed by applying the corresponding adjoint of the forward-modeling operators to the single-scattered surface seismic data and correlating it with the forward-modeled source wavefield to predict the location of the single-scattering energy. Standard migration suffers from migration artifacts, which results in a blurred depiction of the true subsurface reflectivity distributions. This effect is primarily caused by the limited acquisition aperture, coarse source-receiver sampling, band-limited nature of the source wavelet, and low subsurface illumination (Nemeth et al., 1999; Tang, 2009). To remedy this problem, least-squares migration (LSM) can be used to enhance the quality of migrated images and suppress the migration artifacts when imaging primaries and to properly deconvolve the source wavelet and acquisition fingerprint (LeBras and Clayton, 1988; Nemeth et al., 1999). LSM provides an approximate solution for the subsurface reflectivity as demonstrated by the linearized waveform inversion (Snieder et al., 1988; Landa et al., 1989; Crase et al., 1990; Roth and Tarantola, 1992). Dai et al. (2012) adapt LSM to image blended data in a multisource RTM framework. Wang and Sacchi (2007), Herrmann and Li (2012), and Aldawood et al. (2014a) improved the least-squares imaging by imposing sparseness constraints on the least-squares solution to enhance the spatial resolution of the seismic events.

In this paper, we extend the theory of LSM to include the imaging of multiply scattered seismic energy. First, we describe how we linearize the problem of imaging internal multiples. Then, we apply the proposed linearized inversion to delineate synthetic vertical fault planes that are generally poorly illuminated by primaries. Finally, we demonstrate the robustness of the proposed linearized inversion framework to image these fault planes in the presence of white Gaussian observational noise and in the case of imaging the subsurface using erroneous migration velocities.
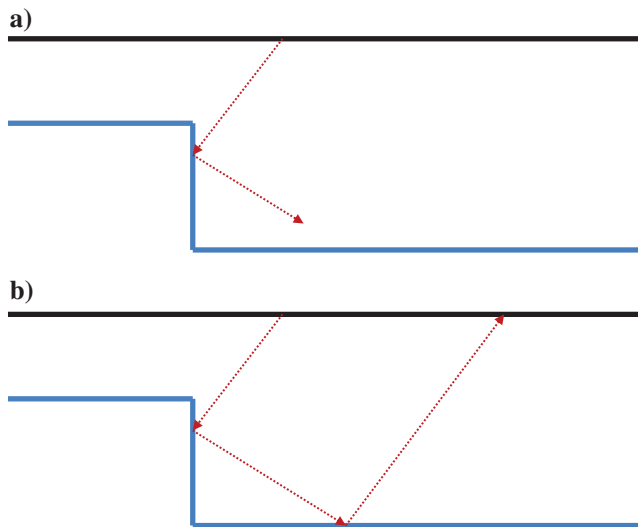
a)

b)

Figure 1. Ray diagrams: (a) A singly scattered ray is reflected off the vertical fault plane away from the recording surface, and (b) a doubly reflected ray ends up at the recording surface illuminating the vertical fault plane.

SINGLE-SCATTERING MODELING AND INVERSION

The Kirchhoff forward-modeling operator $\mathbf{L}$ is a linear operator that maps the subsurface reflectivity distribution to single-scattered seismic data. Mathematically, generating synthetic seismic data corresponding to a reflectivity model of the earth $m(x)$ is obtained using the following forward-modeling operator in the frequency domain:

$$d(s, g, \omega) = \int_V m(x) W(\omega) e^{i\omega(\tau_{sx} + \tau_{xg})} \mathrm{d}x, \qquad (1)$$

where $d(s, g, \omega)$ is the modeled single-scattered seismic data recorded at a surface receiver g due to a surface source located at s for a particular frequency ω. A point in the subsurface volume V is given by x, and $W(\omega)$ denotes the source wavelet. In addition, τ_{sx} is the traveltime from the source point s to the subsurface scattering point x and τ_{xg} is the traveltime from the subsurface scattering point x to the receiver point g. In matrix form, the modeling can be represented by the following matrix-vector multiplication:

$$\mathbf{d} = \mathbf{Lm}, \qquad (2)$$

where $\mathbf{d}$ is the recorded single-scattered data vector and $\mathbf{m}$ is the subsurface reflectivity vector. A standard migrated image is obtained by applying the adjoint of the forward-modeling operator $\mathbf{L^T}$ to the single-scattered seismic data (Claerbout, 1992). The adjoint operator linearly maps the single-scattered seismic data to subsurface structures of the earth as follows:

$$m_{\mathrm{mig}}(x) = \int_s \int_g \int_\omega d(s, g, \omega) W^*(\omega) e^{-i\omega(\tau_{sx} + \tau_{xg})} ds dg d\omega, \qquad (3)$$

where $m_{\mathrm{mig}}(x)$ is the standard migrated image and $W^*(\omega)$ is the complex conjugate of the source wavelet. Although the integration is over the model space x for modeling seismic data, it is over the data space s, g, and ω for imaging recorded data. Physically, the single-scattered data $d(s, g, \omega)$ are back propagated from the receiver point g to the trial imaging point at x and the source function is forward propagated from the source point s to the trial imaging point at x. Then, the image value at the trial imaging point x is given by the zero-lag cross correlation between the forward and backward propagated wavefields, which is obtained by the summation over frequencies in equation 3. In a matrix representation, the Kirchhoff migration is given by a matrix-vector multiplication as follows:

$$\mathbf{m_{mig}} = \mathbf{L^T d}. \qquad (4)$$

The migration operators are adjoint operators of the forward-modeling ones as they *undo* the time and phase shifts of the modeling operators but do not account for the amplitude (Claerbout and Fomel, 2008). Substituting the expression of $\mathbf{d}$ in equation 2 into equation 4 yields

$$\mathbf{m_{mig}} = \mathbf{L^T Lm}, \qquad (5)$$

where $\mathbf{L}^T\mathbf{L}$ is known as the Hessian matrix and reflects the amount of smearing in the seismic image resulting from many factors, including the limited recording aperture, the coarse source-receiver samplings, and the band-limited nature of the source wavelet. The standard migrated image is, therefore, a linear combination of the columns of the Hessian matrix weighted by the true reflectivity model (Aster et al., 2005), providing a physical interpretation of the model resolution.

LSM yields an approximate reflectivity image that minimizes the sum of the squares of the residual vector. Mathematically, the unconstrained objective function is defined as

$$J = \frac{1}{2}\|\mathbf{Lm} - \mathbf{d}\|_2^2 = \frac{1}{2}\mathbf{m}^T\mathbf{L}^T\mathbf{Lm} - \mathbf{m}^T\mathbf{L}^T\mathbf{d} + \frac{1}{2}\mathbf{d}^T\mathbf{d}. \quad (6)$$

The minimizer of this objective function is found by computing the gradient with respect to $\mathbf{m}$ and setting it to zero:

$$\nabla_{\mathbf{m}}J = \mathbf{L}^T\mathbf{Lm} - \mathbf{L}^T\mathbf{d} = 0. \quad (7)$$

A more accurate solution to the migration problem is the least-squares migrated image $\mathbf{m}_{ls}$, which is obtained by solving equation 7, the so-called normal equation:

$$\mathbf{L}^T\mathbf{Lm}_{ls} = \mathbf{L}^T\mathbf{d}. \quad (8)$$

Gradient-based algorithms are generally used to solve this equation iteratively where data residuals are imaged at every iteration to obtain model updates. By rearranging the terms in equation 7, the gradient at every iteration is given by

$$\nabla_{\mathbf{m}}J = \mathbf{L}^T(\mathbf{Lm} - \mathbf{d}) = \mathbf{L}^T\mathbf{r}, \quad (9)$$

where $\mathbf{r}$ is the data residual vector. An initial guess is usually given by the image resulting from the standard Kirchhoff migration in equation 3. At each iteration, the forward-modeling operator in equation 1 is used to synthesize single-scattered seismic data to be compared with the recorded data. Then, the data residual is back projected using the adjoint operator in equation 3 to determine the subsurface reflectivity model updates.

LSM, under the single-scattering assumption, suppresses the migration artifacts and enhances the seismic reflectors, which are mainly illuminated by primaries (Nemeth et al., 1999). However, it hardly delineates reflectors that are mainly illuminated by internally scattered energy (Aldawood et al., 2014b). Consequently, imaging internal multiples potentially helps illuminate subsurface areas that are poorly illuminated by primaries, such as vertical and nearly vertical fault planes (Zuberi and Alkhalifah, 2013).

DOUBLE-SCATTERING MODELING AND INVERSION

The following forward-modeling operator maps the reflectivity model to doubly scattered seismic data as

$$d_d(s, g, \omega) = \int_V \int_V v(x)v(x')W(\omega)e^{i\omega(\tau_{sx}+\tau_{xx'}+\tau_{x'g})}dxdx', \quad (10)$$

where $d_d(s, g, \omega)$ is the recorded doubly scattered seismic data. $v(x)$ and $v(x')$ are the subsurface reflectivity functions at subsurface points x and x', respectively. In addition, τ_{sx} is the traveltime from the source point s to the first scattering point x. $\tau_{xx'}$ is the traveltime from the first scattering point x to the second scattering point x'. Similarly, $\tau_{x'g}$ is the traveltime from the second scattering point x' to the receiver point g.

Clearly, equation 10 shows that generating first-order internal multiples is not obtained by a linear mapping of the subsurface reflectivity model because of the multiplication of two reflectivity functions inside this double-summation modeling step. To linearize this operator, we propose replacing $v(x')$ by the fixed LSM solution $p(x')$ under the single-scattering assumption, obtained by iteratively solving the normal equation 8. Equation 10 then becomes

$$d_d(s, g, \omega) = \int_V \int_V v(x)p(x')W(\omega)e^{i\omega(\tau_{sx}+\tau_{xx'}+\tau_{x'g})}dxdx'. \quad (11)$$

Physically, the scattering points in the least-squares image $p(x)$ under the single-scattering assumption act as secondary sources to predict the doubly scattered internal multiples. To predict a first-order internal multiple, the seismic energy is forward propagated from the source point s to the known scattering point x' and scaled by its reflectivity value $p(x')$. Then, the energy is emitted from this secondary source at x' and forward propagated to the second scattering point x and scaled by its reflectivity value $v(x)$. Last, the energy is further propagated from the second scattering point to the receiver point g to predict the internally scattered first-order multiple. In a similar approach, Zhang and Duan (2012) use the RTM reflectivity image as a secondary source in the subsurface to predict the interbed and surface-related multiples in their reverse time demigration (RTDM) framework.

Our proposed prediction step of first-order multiples can then be represented by the following matrix-vector multiplication:

$$\mathbf{d_d} = \mathbf{Gv}, \quad (12)$$

where $\mathbf{G}$ is the forward-modeling operator that linearly maps the subsurface reflectivity distribution, which is mainly illuminated by the doubly scattered energy $\mathbf{v}$, to first-order internal multiples $\mathbf{d_d}$. The corresponding adjoint of this forward-modeling operator can be applied to image the doubly scattered data as follows:

$$v_{mig}(x) = \int_s \int_g \int_\omega \int_V d_d(s, g, \omega)p(x')W(\omega)$$
$$\times e^{-i\omega(\tau_{sx}+\tau_{xx'}+\tau_{x'g})}dsdgd\omega dx'. \quad (13)$$

The physical interpretation of this equation is that the subsurface reflectivity image, under the double-scattering assumption, can be obtained by applying the following steps:

1) The doubly scattered seismic data are back propagated from the receiver point g to the known scattering point x' and the seismic energy is scaled by the known reflectivity value $p(x')$ at this subsurface scattering point.

2) The scaled redatumed data at the known scattering point x' is further back propagated to the trial imaging point x.
3) The source wavefield is forward propagated from the source point s to the trial imaging point x.
4) We obtain the final image at the trial subsurface scattering point x by applying the zero-lag imaging condition of the forward and backward propagated wavefields.

In matrix form, the subsurface image, under the double-scattering assumption, $\mathbf{v_{mig}}$, is obtained by applying the adjoint operator $\mathbf{G^T}$ to the first-order internal multiples $\mathbf{d_d}$:

$$\mathbf{v_{mig}} = \mathbf{G^T d_d}. \tag{14}$$

Similarly to the single-scattering case, computing the doubly scattered migrated image is determined by a linear mapping of the doubly scattered seismic data. One can also solve a corresponding normal equation iteratively to obtain a least-squares-migrated image $\mathbf{v_{ls}}$ under the double-scattering assumption as follows:

$$\mathbf{G^T G v_{ls}} = \mathbf{G^T d_d}. \tag{15}$$

LSM of the single-scattering and double-scattering energy helps suppress the migration artifacts, remove the acquisition footprint, and deconvolve the source wavelet. It further enhances the spatial resolution of the horizontal reflectors and vertical fault planes. Equations 8 and 15 are usually solved iteratively using a gradient-based optimization algorithm. Here, we use the limited-memory Broyden, Fletcher, Goldfarb, and Shanno (L-BFGS) opti-

mization algorithm to obtain the least-squares solutions. See Appendix A for a more detailed description of the algorithm.

IMAGING VERTICAL FAULT PLANES

In the following two examples, the grid spacings of the 100×50 grids are $\Delta x = \Delta z = 30$ m and the time sampling is 4 ms. We computed 100 noise-free common shot gathers along the surface using the forward-modeling operators to generate the primaries and first-order internal multiples. A Ricker wavelet with a peak frequency of 20 Hz was used to model the seismic data. The first shot is at the top-left corner, and the last shot is at the top-right corner. The remaining equally spaced shots are placed at grid points along the earth surface. Similarly, there are 100 equally spaced receivers placed at the grid points along the earth surface.

The first velocity and reflectivity models are shown in Figure 2a and 2b, respectively. In this example, the two-layer velocity is given by 2500 and 3000 m/s from top to bottom and the common-depth-point (CDP) sampling is 15 m. The dominant frequency of 20 Hz results in a dominant wavelength of 125–150 m. Thus, the image point is sufficiently sampled at about 8–10 points per wavelength. On the other hand, the second velocity and reflectivity models are shown in Figure 3a and 3b, respectively. In this case, the velocity varies between 1500 and 3900 m/s and the CDP sampling is 15 m. Using the same Ricker wavelet, the dominant wavelength ranges from 75 to 195 m. Hence, the image point is sampled sufficiently at 5–13 points per wavelength. Our goal is to apply our proposed LSM of first-order internal multiples to properly image the vertical fault planes in these two different models.

The images obtained using the standard migration of the noise-free data, under the single-scattering assumption, are shown in Figure 4a and 4b, respectively. The panels show blurry depictions of
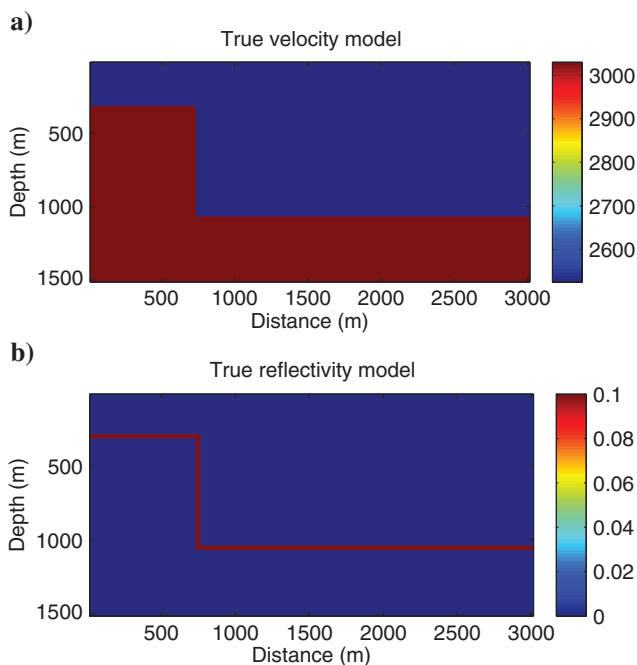

Figure 2. A vertical fault plane: (a) True velocity model. (b) True reflectivity model.
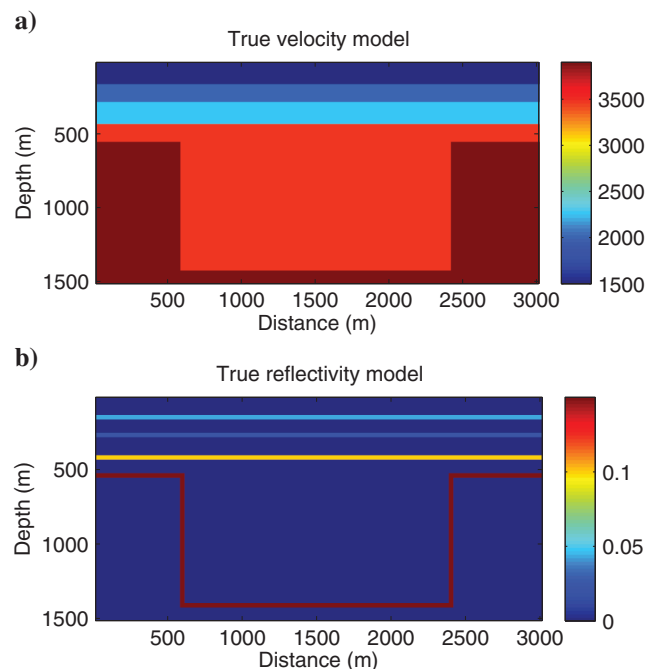

Figure 3. Buried vertical faults: (a) True velocity model. (b) True reflectivity model.

the true subsurface reflectivity distributions due to the migration artifacts, the acquisition fingerprint, and the limited bandwidth of the source wavelet. To alleviate these problems, LSM, under the single-scattering assumption, is applied to suppress the migration artifacts and sharpen the reflectors as shown in Figure 5a and 5b, respectively. After 30 quasi-Newton L-BFGS iterations, this linearized inversion mainly enhances the horizontal reflectors, which are mostly illuminated by primaries. The vertical fault planes, however, are not well delineated because they are mainly illuminated by internally scattered seismic energy.

The least-squares solutions in Figure 5a and 5b, under the single-scattering assumption, are then used in equations 11 and 13 to linearize the inversion of the first-order internal multiples using our proposed framework. These solutions are further used as an initial guess to start the proposed linearized inversion. After 11 quasi-Newton L-BFGS iterations, the LSM results, based on the double-scattering assumption, are shown in Figure 6a and 6b, respectively. Clearly, inverting the first-order internal multiples delineates the fault planes and localizes them in their true subsurface positions.

ROBUSTNESS TO OBSERVATION NOISE

A challenging question is to assess whether one can accurately delineate the vertical fault plane in the presence of white observational noise. To answer this question, we added white Gaussian random noise to the two synthetic noise-free data sets. We then used our LSM framework to image the first-order internal multiples to assess the robustness of this linearized inversion.

The least-squares images, under the double-scattering assumption, of the two noisy data sets with a signal-to-noise ratio (S/N) of −15 and −10 dB are shown in Figure 7a and Figure 7b, respectively. These solutions, obtained after 10 quasi-Newton L-BFGS iterations, still show a clear delineation of the fault planes despite the defocusing of the seismic events. However, as the data become noisy, the quality of the LSM images, under the double-scattering assumption, deteriorates compared with the noise-free images shown in Figure 5a and 5b.

ROBUSTNESS TO VELOCITY ERROR

Velocity model building to estimate accurate migration velocities is one of the major challenges in seismic imaging. Hence, another question of interest is to study whether our proposed linearized inversion of doubly scattered data is robust in the case of imaging with inaccurate migration velocities. In other words, What is the effect of velocity errors on the final image quality and on the ability of the proposed framework to resolve the vertical fault planes?

To answer these questions, we applied the linearized framework to image first-order internal multiples using slower and faster migration velocities. We first compute the least-squares solutions, under the single-scattering assumption, to image the recorded data with velocity models that are 1% faster and slower than the exact
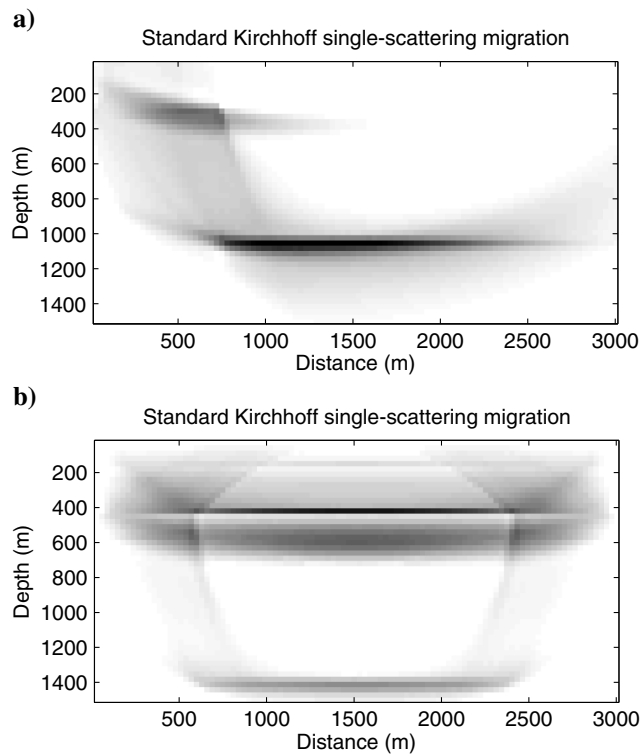

a)

Figure 4. Standard Kirchhoff migration: (a) Example 1: Kirchhoff migration of the two-layer model. (b) Example 2: Kirchhoff migration of the buried vertical faults model. The standard Kirchhoff imaging of primaries only yields a blurry representation of the subsurface structures.
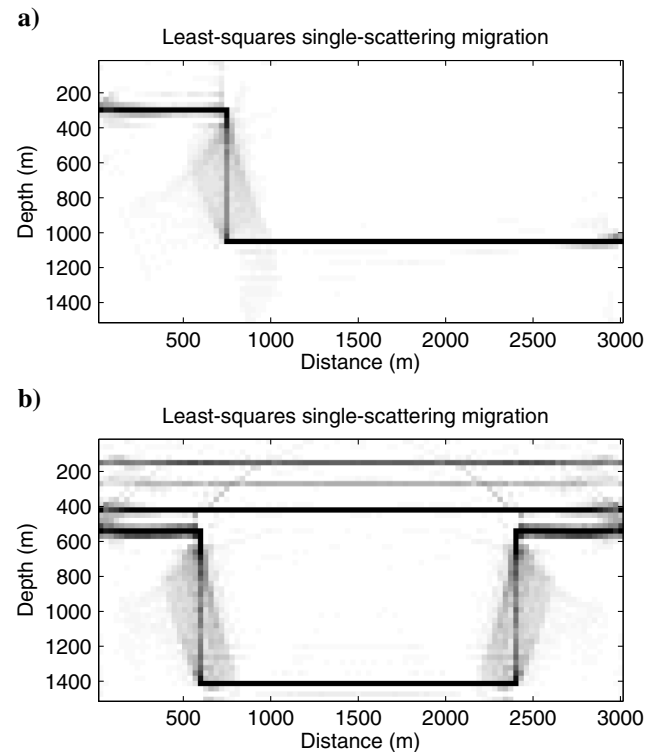
Figure 5. Imaging vertical faults: (a) Example 1: Single-scattering least-squares migrated image. (b) Example 2: Single-scattering least-squares migrated image. The LSM of primaries better emphasizes only the horizontal reflectors.

velocity model. The LSM solutions of the first and second examples, imaged using the slower and faster velocities, are plotted in Figure 8a and 8b, and in Figure 8c and 8d, respectively. One can see that these least-squares images become more defocused and mispositioned as the velocity becomes erroneous. Although the horizontal reflectors are pulled up when migrating the data using slower migration velocities, they are pushed down when the recorded data are imaged with faster migration velocities.

These images are then used as constraints and initial guesses in the linearized inversion framework under the double-scattering assumption. The least-squares solutions, obtained by imaging first-order internal multiples using the slower and faster velocities, are plotted in Figure 9a and 9b for the first example and in Figure 9c and 9d for the second example. The vertical fault planes are fairly spatially focused in case the migration velocity error is ±1% as demonstrated in these figures.

The results suggest that the proposed least-squares imaging, under the double-scattering assumption, can be sensitive to the accuracy of the migration velocity model and the accuracy of the scatterers' locations obtained by inverting the primaries. As a result, these two factors could cause more defocusing and mispositioning of the vertical fault planes compared with the horizontal reflectors, which are mainly delineated by the LSM of primaries.

DISCUSSION

Standard migration suffers from migration artifacts and low spatial resolution due to the limited acquisition aperture, coarse source-receiver sampling, limited bandwidth of the source wavelet, and low subsurface illumination as shown in Figure 4. LSM helps suppress migration artifacts, remove the acquisition fingerprint, and enhance the spatial resolution of the migrated section by iteratively finding a subsurface reflectivity model that better explains the observed data in a least-squares sense.

However, the standard and iterative LSM algorithms, under the single-scattering assumption, yield migrated sections that focus seismic events, which are mainly illuminated by primaries, such as horizontal reflectors as seen in Figures 4 and 5, respectively. We have demonstrated that modeling the doubly scattered data is mathematically not represented by a linear mapping of the subsurface reflectivity distribution. Thus, we proposed a linearization step in which we use the least-squares image, under the single-scattering assumption, as a constraint in the forward modeling and adjoint operators, which models and images first-order multiples, respectively. We then demonstrated the effectiveness of the proposed linearized inversion scheme of doubly scattered energy and its ability to localize vertical fault planes in Figure 6.

We also demonstrated the robustness of this inversion framework in the presence of white Gaussian random observational noise. Overall, the proposed least-squares imaging, under the double-scattering assumption, still retrieves the vertical fault planes when imaging the noisy data sets despite the slight defocusing of these events due to the presence of noise as illustrated in Figure 7.

We finally studied the robustness of our proposed linearized inversion when imaging data with inaccurate migration velocities. In
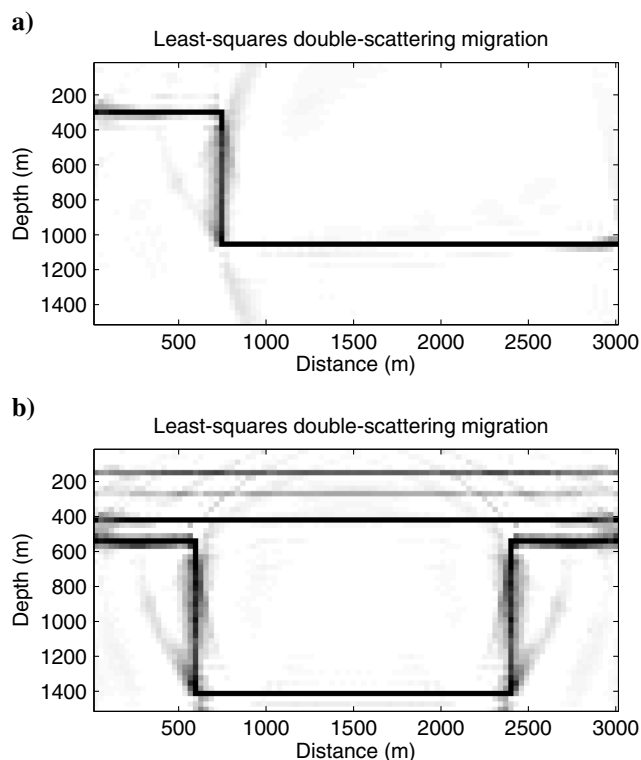

Figure 6. Imaging vertical faults: (a) Example 1: Double-scattering least-squares migrated image. (b) Example 2: Double-scattering least-squares migrated image. The LSM of first-order internal multiples better delineates the vertical fault plane.
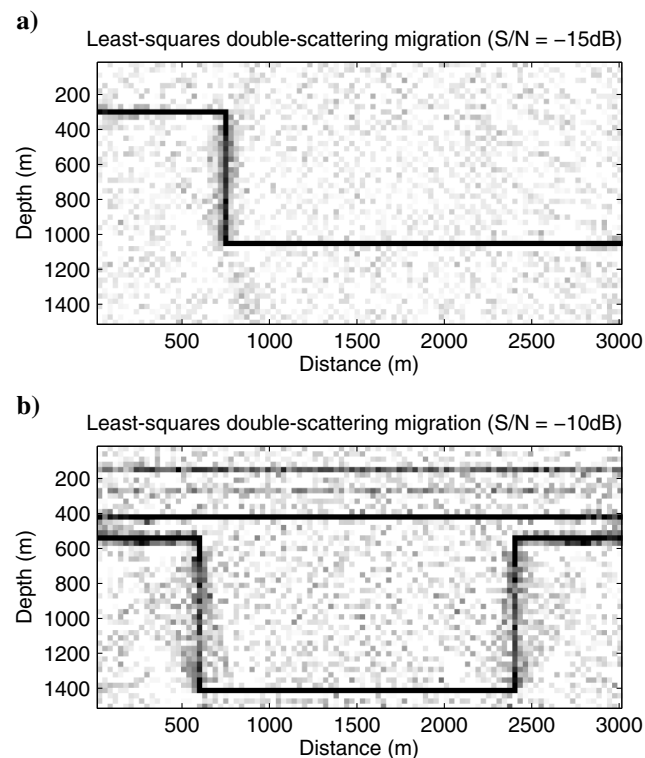

Figure 7. Double-scattering LSM of noisy data sets: (a) Example 1: double-scattering LSM with S/N = −15 dB. (b) Example 2: double-scattering LSM with S/N = −10 dB. When the data become noisy, the ability of the least-squares imaging of first-order internal multiples to delineate the vertical fault planes deteriorates.

Figure 9, we showed how the LSM of the doubly scattered data could delineate the vertical fault planes although they become mispositioned and unfocused in the presence of velocity errors. We conclude that our linearized inversion of first-order internal multiples can be sensitive to the migration velocity models and the constraining least-squares images of primaries.
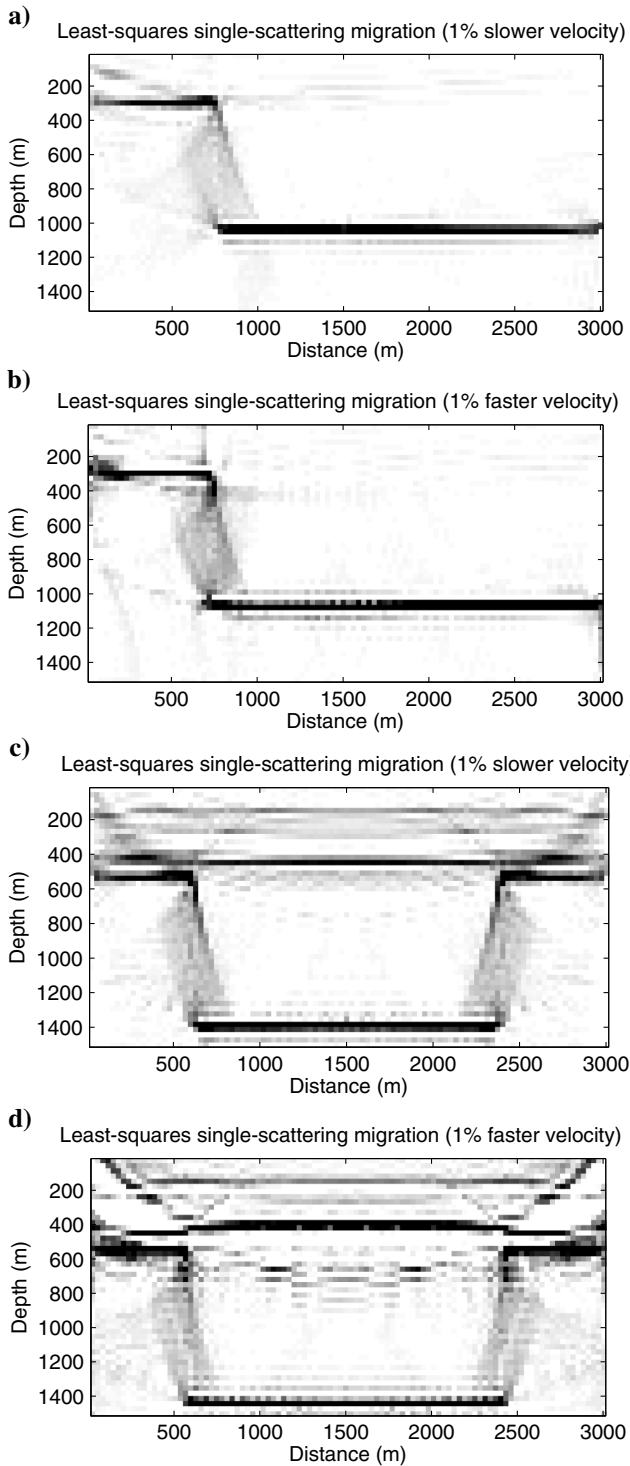

Figure 8. LSM of primaries using the slower and faster migration velocities. Example 1: LSM using (a) 1% slower velocity model and (b) 1% faster velocity model. Example 2: LSM using (c) 1% slower velocity and (d) 1% faster velocity.
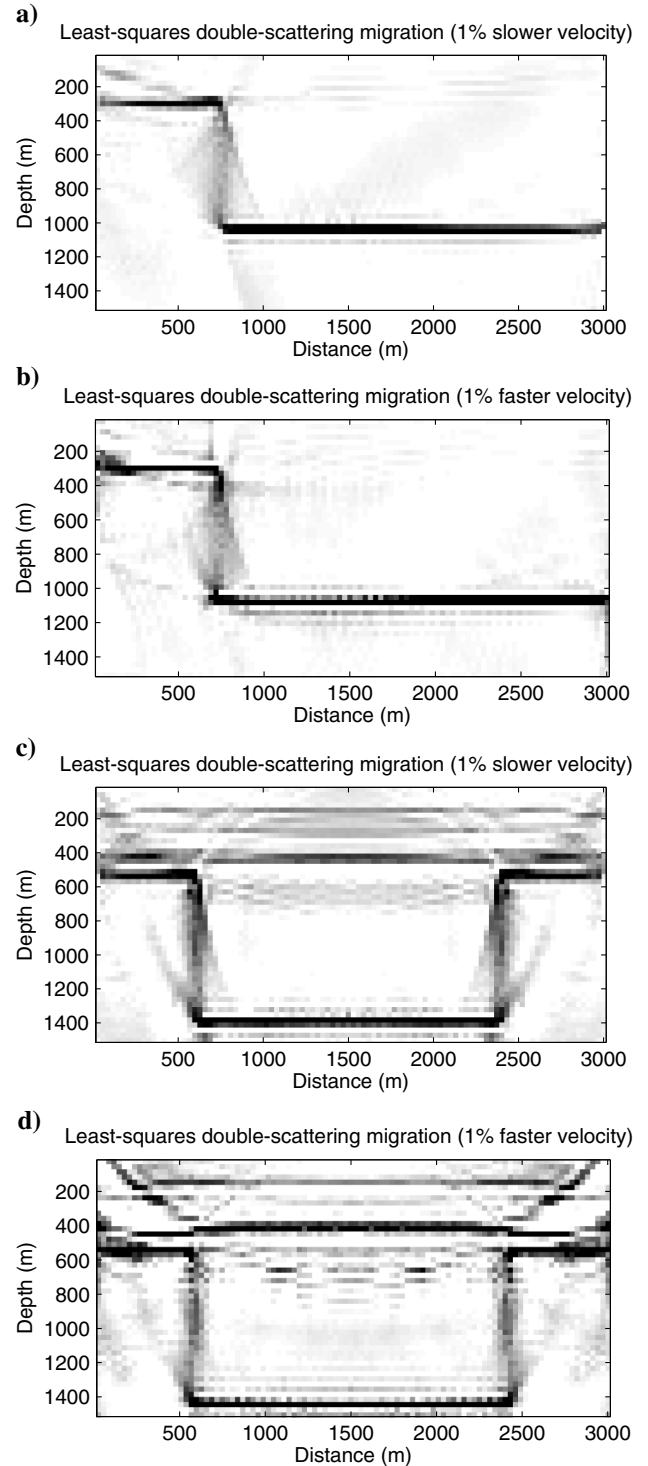
Figure 9. LSM of first-order multiples using the slower and faster migration velocities. Example 1: LSM using (a) 1% slower velocity model and (b) 1% faster velocity model. Example 2: LSM using (c) 1% slower velocity and (d) 1% faster velocity.

CONCLUSIONS

We proposed a linearized inversion framework to invert first-order internal multiples in the seismic records. The least-square image, based on the single-scattering assumption, was used as a constraint to linearize the forward modeling and adjoint operators of doubly scattered energy. Our results on synthetic data revealed the effectiveness of the proposed linearized inversion in delineating vertical fault planes, which are mainly illuminated by doubly scattered seismic energy. We further demonstrated the robustness of the LSM of first-order multiples in the presence of noise and in case of imaging using inaccurate velocities. In conclusion, we would expect the proposed LSM, under the double-scattering assumption, to work best if the data quality is high, and the single-scattered least-squares image and the migration velocity model are quite accurate.

Future work will consider using a sparsity-promoting imaging framework to obtain highly focused images with our proposed linearized inversion. It will also include the application of this least-squares imaging technique to real data examples and 3D synthetic models in an efficient RTM framework instead of the currently used Kirchhoff-based framework.

ACKNOWLEDGMENTS

APPENDIX A

GRADIENT-BASED OPTIMIZATION

The L2-norm least-squares misfit objective function has the form

$$J = \frac{1}{2}\|\mathbf{\Psi m} - \mathbf{d^{obs}}\|_2^2, \qquad (A-1)$$

where $\mathbf{\Psi}$ is a linear operator that maps the observed data vector $\mathbf{d^{obs}}$ to the vector of model parameters $\mathbf{m}$. This objective function is minimized and the model parameters are updated after each iteration by moving along the negative gradient direction (i.e., the direction of maximum descent) to obtain the next iterate using a steepest descent algorithm. The gradient is obtained by back projecting the data residual vector onto the model space using the adjoint operator. However, the steepest descent algorithm suffers from slow convergence because it only uses the gradient direction at every iteration (Aster et al., 2005). A better search direction is the quasi-Newton direction, which uses the gradient direction and the curvature information of the objective function characterized by the Hessian matrix. An approximate Hessian can be deduced by monitoring how the gradient changes with respect to the model update after each iteration.

The BFGS quasi-Newton algorithm, named after its discoverers Broyden, Fletcher, Goldfarb, and Shanno, uses an approximate Hessian $\mathbf{B}_k$ in finding the new search direction at the kth iteration given by the projection of the gradient vector onto the column space of the inverse Hessian approximation (i.e., $-\mathbf{B}_k^{-1}\nabla_\mathbf{m}J_k$) (Nocedal and Wright, 2006). The value $\mathbf{B}_k$ is a symmetric positive definite matrix that is updated after each iteration, and ∇J_k is the gradient

direction evaluated at the current iterate. In the steepest-descent algorithm, the search direction is the negative gradient direction with $\mathbf{B} = \mathbf{I}$. The BFGS formula for updating the approximate Hessian is given by

$$\mathbf{B}_{k+1} = \mathbf{B}_k - \frac{\mathbf{B}_k\mathbf{s}_k\mathbf{s}_k^T\mathbf{B}_k}{\mathbf{s}_k^T\mathbf{B}_k\mathbf{s}_k} + \frac{\mathbf{y}_k\mathbf{y}_k^T}{\mathbf{y}_k^T\mathbf{s}_k}, \qquad (A-2)$$

where $\mathbf{s}_k = \mathbf{m}_{k+1} - \mathbf{m}_k$ and $\mathbf{y}_k = \nabla_\mathbf{m}J_{k+1} - \nabla_\mathbf{m}J_k$. Notice that the difference between $\mathbf{B}_{k+1}$ and $\mathbf{B}_k$ is a rank-two matrix. The major advantage of the quasi-Newton BFGS algorithm is that it does not require the calculation of second derivatives, using the gradient changes to obtain curvature information (Nocedal and Wright, 2006). The reflectivity model is updated by moving along this quasi-Newton direction as follows:

$$\mathbf{m}_{k+1} = \mathbf{m}_k - \alpha\mathbf{B}^{-1}\nabla_\mathbf{m}J_k, \qquad (A-3)$$

where α is a scalar value that determines how much one should move along the quasi-Newton direction and it could be obtained by a numerical line search scheme (Nocedal and Wright, 2006).

The "fmincon" function, included in the MATLAB optimization toolbox, computes the minimum of a scalar objective function of several variables. It implements an L-BFGS optimization algorithm, which is based on the BFGS optimization algorithm. L-BFGS is used to solve large-scale problems, whose approximate Hessian matrices of size $n \times n$, where n is the number of knowns, cannot be computed and/or stored (Nocedal and Wright, 2006). Instead, they store just a few vectors of length n that represent the Hessian approximation implicitly. We used it here to solve the single- and the double-scattering imaging problems. The stopping criteria of the optimization is when the number of function evaluations reaches a maximum of 50 evaluations.

REFERENCES

Aldawood, A., I. Hoteit, and T. Alkhalifah, 2014a, The possibilities of compressed-sensing-based Kirchhoff prestack migration: Geophysics, **79**, no. 3, S113–S120, doi: 10.1190/geo2013-0271.1.

Aldawood, A., I. Hoteit, G. Turkiyyah, M. Zuberi, and T. Alkhalifah, 2014b, Linearized least-square imaging of internally scattered data: 76th Annual International Conference and Exhibition, EAGE, Extended Abstracts, doi: 10.3997/2214-4609.20141160.

Aster, R., B. Borchers, and C. Thurber, 2005, Parameter estimation and inverse problems (international geophysics), HAR/CDR ed.: Academic Press.

Behura, J., K. Wapenaar, and R. Snieder, 2012, Newton-Marchenko-Rose imaging: 82nd Annual International Meeting, SEG, Expanded Abstracts, doi: 10.1190/segam2012-1531.1.

Claerbout, J. F., 1992, Earth sounding analysis: Processing versus inversion: Blackwell Science.

Claerbout, J. F., and S. Fomel, 2008, Image estimation by example: Geophysical soundings image construction: Multidimensional autoregression: Stanford University.

Crase, E., A. Pica, M. Noble, J. McDonald, and A. Tarantola, 1990, Robust elastic nonlinear waveform inversion: Application to real data: Geophysics, **55**, 527–538, doi: 10.1190/1.1442864.

Dai, W., P. Fowler, and G. T. Schuster, 2012, Multi-source least-squares reverse time migration: Geophysical Prospecting, **60**, 681–695, doi: 10.1111/j.1365-2478.2012.01092.x.

Herrmann, F. J., and X. Li, 2012, Efficient least-squares imaging with sparsity promotion and compressive sensing: Geophysical Prospecting, **60**, 696–712, doi: 10.1111/j.1365-2478.2011.01041.x.

Landa, E., W. Beydoun, and A. Tarantola, 1989, Reference velocity model estimation from prestack waveforms: Coherency optimization by simulated annealing: Geophysics, **54**, 984–990, doi: 10.1190/1.1442741.

LeBras, R., and R. Clayton, 1988, An iterative inversion of backscattered acoustic waves: Geophysics, **53**, 501–508, doi: 10.1190/1.1442481.

Malcolm, A. E., B. Ursin, and M. V. De Hoop, 2009, Seismic imaging and illumination with internal multiples: Geophysical Journal International, **176**, 847–864, doi: 10.1111/j.1365-246X.2008.03992.x.

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

Nocedal, J., and S. J. Wright, 2006, Numerical optimization, 2nd ed.: Springer.

Roth, G., and A. Tarantola, 1992, Inversion of seismic waveforms using neural networks: 62nd Annual International Meeting, SEG, Expanded Abstracts, 788–791.

Snieder, R., M. Y. Xie, A. Tarantola, and A. Pica, 1988, Retrieving both the impedance variations and background velocity in reflection seismics using least-squares waveform inversion: 58th Annual International Meeting, SEG, Expanded Abstracts, 1070–1072.

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.

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

Wapenaar, K., F. Broggini, and R. Snieder, 2012, Creating a virtual source inside a medium from reflection data: Heuristic derivation and stationary-phase analysis: Geophysical Journal International, **190**, 1020–1024, doi: 10.1111/j.1365-246X.2012.05551.x.

Zhang, Y., and L. Duan, 2012, Predicting multiples using a reverse time demigration: 82nd Annual International Meeting, SEG, Expanded Abstracts, doi: 10.1190/segam2012-0520.1.

Zuberi, M., and T. Alkhalifah, 2013, Selective interferometric imaging of internal multiples: 75th Annual International Conference and Exhibition, EAGE, Extended abstracts, doi: 10.3997/2214-4609.20130197.

Zuberi, M. A. H., and T. Alkhalifah, 2014, Generalized internal multiple imaging (GIMI) using Feynman-like diagrams: Geophysical Journal International, **197**, 1582–1592, doi: 10.1093/gji/ggt527.