GEOPHYSICS, VOL. 64, NO. 1 (JANUARY-FEBRUARY 1999); P. 208–221, 15 FIGS.

Least-squares migration of incomplete reflection data

Tamas Nemeth*, Chengjun Wu‡, and Gerard T. Schuster**

ABSTRACT

A least-squares migration algorithm is presented that reduces the migration artifacts (i.e., recording footprint noise) arising from incomplete data. Instead of migrating data with the adjoint of the forward modeling operator, the normal equations are inverted by using a preconditioned linear conjugate gradient scheme that employs regularization. The modeling operator is constructed from an asymptotic acoustic integral equation, and its adjoint is the Kirchhoff migration operator. We tested the performance of the least-squares migration on synthetic and field data in the cases of limited recording aperture, coarse sampling, and acquisition gaps in the data. Numerical results show that the least-squares migrated sections are typically more focused than are the corresponding Kirchhoff migrated sections and their reflectivity frequency distributions are closer to those of the true model frequency distribution. Regularization helps attenuate migration artifacts and provides a sharper, better frequency distribution of estimated reflectivity. The least-squares migrated sections can be used to predict the missing data traces and interpolate and extrapolate them according to the governing modeling equations. Several field data examples are presented. A ground-penetrating radar data example demonstrates the suppression of the recording footprint noise due to a limited aperture, a large gap, and an undersampled receiver line. In addition, better fault resolution was achieved after applying least-squares migration to a poststack marine data set. And a reverse vertical seismic profiling example shows that the recording footprint noise due to a coarse receiver interval can be suppressed by least-squares migration.

INTRODUCTION

A standard seismic migration operator can be regarded as the adjoint of a seismic forward modeling operator (Claerbout, 1992) as used in the iterations of full waveform inversion (Tarantola, 1984, 1987). Although an adjoint operator is a useful approximation to the inverse of the forward modeling operator, it is not the exact inverse. The standard migration operator (e.g., Kirchhoff) can be modified to approximate the exact inverse operator. For example, true-amplitude migration/inversion schemes (Bleistein, 1984; Hanitzsch et al., 1994) were derived to compensate for the spreading losses in the modeling operator. These schemes use high-frequency asymptotics to derive an analytic inverse represented by an integral equation, assuming as infinite recording aperture. Gray (1997) gives an excellent comparison of the true-amplitude methods. Similarly, Miller et al. (1987) modified the migration weights according to the generalized Radon transform. Carrion et al. (1991) presented a wavefront-set method to estimate the effects of a limited recording aperture. In most cases seismic data cannot be inverted for the correct earth model because either the modeling equations are inconsistent with the data or an unlimited recording aperture is not feasible. In such cases the exact inverse can be approximated by a generalized inverse which can be formulated as a least-squares migration (LSM) operator. The operation of LSM is equivalent to a linearized inversion of the reflectivity for a given velocity distribution. Tarantola (1987) provided the basic formalism for such an inversion.

The following authors applied linearized inversion to a variety of seismic problems. LeBras and Clayton (1988) applied iterative inversion to backscattered acoustic waves to obtain velocity and density perturbations. Beydoun and Mendes (1989) inverted elastic waves to obtain velocity perturbations. Lambare et al. (1992) applied an iterative acoustic asymptotic inversion to a North Sea data set and accurately reconstructed velocity perturbations. Cole and Karrenbach (1992) used this approach to reduce migration artifacts due to the aperture

Manuscript received by the Editor July 22, 1996; revised manuscript received June 1, 1998.
*Formerly Department of Geology and Geophysics, University of Utah, Salt Lake City, Utah 84112; currently Chevron Petroleum Technology Co., 1300 Beach Boulevard, La Habra, California 90631. E-mail: nemt@chevron.com.
‡Formerly Department of Geology and Geophysics, University of Utah, Salt Lake City, Utah 84112; currently 12471 Jack Bell Drive, Richmond, British Columbia V6V 2R9, Canada. E-mail: chengjun@mines.utah.edu.
**Department of Geology and Geophysics, University of Utah, Salt Lake City, Utah 84112. E-mail: schuster@mines.utah.edu.

limitation of the observed diffractions. Schuster (1993) generalized the least-squares formulation by adding constraint terms to the objective function and applied it to crosswell synthetic data. Sevink and Herman (1994) applied a Krylov-subspace preconditioner to the main diagonal of the related Hessian matrix to accelerate convergence. Ronen et al. (1995) used dealiasing dip moveout (DMO) (a partial prestack migration) to invert for the stacked zero-offset data in cases of moderate data acquisition irregularities and predict the missing data. Mittet et al. (1997) used an iterative elastic migration to image complex geology from offset vertical seismic profiling (VSP) data. Wu and McMechan (1996) implemented a 2-D full-waveform inversion of double-couple earthquake sources where the velocity and density distributions were assumed to be known a priori. They report increased model resolution and posterior covariance after inversion. Duquet (1996) applied a linearized inversion technique based on paraxial approximation to the wave equation. He showed that the method can remove artifacts contaminating the individual depth images by integrating a priori information in the inverse problem.

Favoring least-squares migration over Kirchhoff migration or asymptotic inversion is justified when the data are incomplete. This can be explained by using Huygens' principle: a wavefront can be represented by a composite of secondary wavefronts that originate from point sources along an earlier wavefront. Kirchhoff migrated images are obtained by back-projecting the observed data into the medium, where every point of the medium is considered to be a possible diffractor. A necessary condition to image every diffraction point is that the observed data along the aperture should be complete. According to the slice-projection theorem (Kak and Slaney, 1988), an object can be completely imaged if it is illuminated from all angles. For example, to completely reconstruct a point-wise diffractor with measurements along the surface, continuous measurements along an extremely wide recording aperture are necessary. Furthermore, for discretely sampled data, the spatial frequency of the objects in the view directions should satisfy the spatial Nyquist sampling criterion. If the above conditions are not satisfied, the data are incomplete and the object cannot be exactly reconstructed.

There are several cases when the completeness condition is not fulfilled in geophysics: truncated recording apertures, coarse source-receiver distributions, and gaps in the recording lines. A truncated recording aperture corresponds to multiplying the complete data by a boxcar function or convolving the spectrum of the data with the corresponding sinc function (Kak and Slaney, 1988). Data with large gaps in the recording aperture correspond to multiplying complete data by a set of nonoverlapping boxcar functions. Coarse source and/or geophone sampling introduces aliasing artifacts due to the wraparound effect in the model spectrum. All of these migration artifacts contribute to the "recording footprint noise" seen in migrated sections.

The purpose of this paper is to evaluate the performance of LSM for migrating incomplete data. The first part presents the projector matrix associated with Kirchhoff migration and establishes the relationship between incomplete data and the matrix equations. The next section compares the performance of preconditioned Kirchhoff migration and LSM for several cases of incomplete data. Following that, we describe three applications of LSM to field data. Suggestions on improving the LSM method are given in the Discussion section.

BASIC EQUATIONS

Assume that the linear forward modeling operator $\underset{\sim}{\mathbf{L}}$ satisfies

$$\mathbf{p} = \underset{\sim}{\mathbf{L}}\mathbf{m}, \tag{1}$$

where $\mathbf{p}$ is a vector of modeled data and $\mathbf{m}$ is the earth reflectivity model vector. The observed data $\mathbf{p}_o$ are described by $\mathbf{p}_o = \underset{\sim}{\mathbf{L}}_o\mathbf{m}_o$ where $\mathbf{m}_o$ is the true earth reflectivity model vector and $\underset{\sim}{\mathbf{L}}_o$ is the forward modeling operator for the actual earth model. Unless stated otherwise, we assume that $\underset{\sim}{\mathbf{L}} = \underset{\sim}{\mathbf{L}}_o$.

Kirchhoff migration uses the transpose of the forward modeling operator in equation 1:

$$\mathbf{m}_k = \underset{\sim}{\mathbf{L}}^T\mathbf{p}, \tag{2}$$

where $\mathbf{m}_k$ is the Kirchhoff migrated section. Substituting equation (1) into equation (2) yields

$$\mathbf{m}_k = \underset{\sim}{\mathbf{L}}^T\underset{\sim}{\mathbf{L}}\mathbf{m}. \tag{3}$$

The matrix $\underset{\sim}{\mathbf{L}}^T\underset{\sim}{\mathbf{L}}$ is a Hessian and defines $\mathbf{m}_k$ as an $\underset{\sim}{\mathbf{L}}^T\underset{\sim}{\mathbf{L}}$-filtered version of $\mathbf{m}$. The Kirchhoff migration operator will correctly reconstruct the actual earth model vector if $\underset{\sim}{\mathbf{L}}^T\underset{\sim}{\mathbf{L}}$ is an identity matrix $\underset{\sim}{\mathbf{I}}$. In most cases $\underset{\sim}{\mathbf{L}}^T\underset{\sim}{\mathbf{L}}$ is not an identity matrix and has the following characteristics: the elements in the main diagonal are nonuniform and differ from unity, and elements off the main diagonal are nonzero.

To illustrate these characteristics, the explicit form of the integral operator corresponding to $\underset{\sim}{\mathbf{L}}^T\underset{\sim}{\mathbf{L}}$ is (see Appendix A)

$$K(x, x') = \int ds h_s(s) A_{sx} A_{sx'} \times \int dr h_r(r) A_{xr} A_{x'r} R(\tau_{sx} + \tau_{xr} - \tau_{sx'} - \tau_{x'r}), \tag{4}$$

where x and x' denote the locations associated with $\mathbf{m}_k$ and $\mathbf{m}$, respectively; τ_{sx} is the wave propagation time from the source at s to the reflector at x; τ_{xr} is the wave propagation time from the reflector to the receiver at r; A_{sx} and A_{xr} are the amplitude terms (estimated from the transport equation) that account for geometrical spreading from, respectively, the source to the reflector and the reflector to the receiver; $W = W(t)$ represents the time history of the source wavelet; R is the temporal cross-correlation of $\ddot{W}$ delayed by the time $(\tau_{sx} + \tau_{xr} - \tau_{sx'} - \tau_{x'r})$, and $h_s(s)$ and $h_r(r)$ are the source and receiver sampling functions, respectively.

The following observations can be made from expression (4):

(1) The main diagonal has the form

$$K(x_0, x_0) = R(0) \int ds h_s(s) A^2_{sx_0} \int dr h_r(r) A^2_{x_0r}. \tag{5}$$

This expression shows that the value of the main diagonal elements (for a given source wavelet) depends on the spreading losses ($A^2_{sx_0}$, $A^2_{x_0r}$) and on the discontinuous sampling of the wavefield as governed by $h_s(s)$ and $h_r(r)$. Kirchhoff migration schemes are usually augmented to compensate for the spreading losses; this results in the analytic inversion formulas of Bleistein (1984) and others. The effects of discontinuous sampling can be mitigated by normalizing $K(x_0, x_0)$ by the numerical hit count in each CDP bin. This compensation scheme can be interpreted as applying $diag(\underset{\sim}{\mathbf{L}}^T\underset{\sim}{\mathbf{L}})^{-1}$ as a preconditioning matrix.

(2) Each column of $\underset{\sim}{\mathbf{L}}^T\underset{\sim}{\mathbf{L}}$ is the response to a fixed element x' in $\mathbf{m}$ applied to the migrated section $\mathbf{m}_k$, after modeling and migration. For one trace, the migration response will be along migration ellipses described by x that satisfy $\tau_{sx} + \tau_{xr} = \tau_{sx'} + \tau_{x'r}$. For complete data, the migration response is the sum of responses for individual traces (with their corresponding weighting coefficients) and migration yields a focused image around $x = x'$ for each diffractor point at x'. In such cases $\underset{\sim}{\mathbf{L}}^T\underset{\sim}{\mathbf{L}}$ is diagonally dominated and $\underset{\sim}{\mathbf{L}}^T \approx \underset{\sim}{\mathbf{L}}^{-1}$. For incomplete data, the migration response is not focused only around $x = x'$ because the missing data lead to incomplete cancellation of artifacts and thus significant off-diagonal elements exist in $\underset{\sim}{\mathbf{L}}^T\underset{\sim}{\mathbf{L}}$. In this case, $\underset{\sim}{\mathbf{L}}^T \not\approx \underset{\sim}{\mathbf{L}}^{-1}$, and some kind of matrix inversion method is necessary.

Thus, if a source-receiver distribution results in incomplete data, Kirchhoff migration will produce migration artifacts in the migrated section. To minimize these artifacts, we shall solve for $\mathbf{m}$ by minimizing the following objective function:

$$P(\mathbf{m}) = \|\underset{\sim}{\mathbf{L}}\mathbf{m} - \mathbf{p}_o\|^2 + \varepsilon^2\|\underset{\sim}{\mathbf{C}}\mathbf{m} - \underset{\sim}{\mathbf{C}}\mathbf{m}_{apr}\|^2. \qquad (6)$$

The first term on the right-hand side of equation (6) is the data misfit function and the second term is a regularization term. $\underset{\sim}{\mathbf{C}}$ denotes a linear operator acting on $\mathbf{m}$, specified for each application using a priori information. A priori information about the model might be incorporated with the $\mathbf{m}_{apr}$ term, and ε^2 is the regularization weight. Expression (6) assumes that the C-filtered model residual and the data residual probability distributions are Gaussian.

The model that minimizes equation (6) is (Tarantola, 1987)

$$\mathbf{m} = (\underset{\sim}{\mathbf{L}}^T\underset{\sim}{\mathbf{L}} + \varepsilon^2\underset{\sim}{\mathbf{C}}^T\underset{\sim}{\mathbf{C}})^{-1}\left(\underset{\sim}{\mathbf{L}}^T\mathbf{p}_o + \varepsilon^2\underset{\sim}{\mathbf{C}}^T\underset{\sim}{\mathbf{C}}\mathbf{m}_{apr}\right), \qquad (7)$$

and solves the normal equation

$$(\underset{\sim}{\mathbf{L}}^T\underset{\sim}{\mathbf{L}} + \varepsilon^2\underset{\sim}{\mathbf{C}}^T\underset{\sim}{\mathbf{C}})\mathbf{m} = \underset{\sim}{\mathbf{L}}^T\mathbf{p}_o + \varepsilon^2\underset{\sim}{\mathbf{C}}^T\underset{\sim}{\mathbf{C}}\mathbf{m}_{apr}. \qquad (8)$$

In practice $\mathbf{m}$ is solved using an iterative scheme. In this paper, a preconditioned conjugate gradient scheme (see Appendix B) is used to solve equation (8) for the model $\mathbf{m}$. This procedure is denoted as least-squares migration.

APPLICATIONS TO SYNTHETIC DATA

In this section, the least-squares migration is applied to synthetic high-frequency seismic reflection data. The goal of these tests is to demonstrate the merits and limitations of LSM. We will study the performance of LSM in the cases of limited data aperture, coarsely sampled data, and data with gaps in the receiver line.

Limited aperture

The objective here is to demonstrate that LSM can compensate for migration artifacts due to limited recording apertures. Figure 1a shows a 12-point diffractor model where each diffractor has a reflectivity value of 1 and the background propagation velocity is 2000 m/s. Figure 1b shows the common-shot gather of seismograms generated by a source at 0 m horizontal position and recorded by geophones with 1-m station increments. The Kirchhoff and least-squares migrated sections are shown in Figure 1c and 1d, respectively; the latter section has fewer artifacts, as demonstrated in Figure 1e. Iterations were continued until the energy of the misfit seismograms reached the threshold value of 0.001 of the input data energy (also denoted as normalized objective function value). A total of 51 iterations were needed to reach this threshold level, and no regularization was applied.

The frequency distributions of the reflectivities in Figures 1a, 1c, and 1d are depicted in the histograms of Figure 1f, where the x-axis corresponds to the reflectivity values and the y-axis corresponds to the frequencies (occurrences) of the given reflectivity in the migrated sections. Least-squares migration has produced a sharper frequency distribution of reflectivities than has Kirchhoff migration and the LSM frequency distribution is a better approximation to the true frequency distribution denoted by the vertical arrows.

Coarse wavefield sampling

Spatially undersampled data introduce recording footprint noise into the migrated sections. The goal here is to show that LSM can partially attenuate the recording footprint noise and fill in the data gap with traces by using the least-squares migrated section.

Figure 2a shows an undersampled version of Figure 1a that is obtained by zeroing out every second trace. The preconditioned Kirchhoff migrated section is shown in Figure 2b. Both aliasing and limited aperture artifacts are observable in this migrated section. Figure 2c shows the least-squares migrated section after 37 iterations (with regularization applied) where the footprint noise has been considerably attenuated. The predicted (both the known and missing) data are depicted in Figure 2d and show a highly accurate prediction.

Gaps in the data

Gaps in recording lines are often treated as if zeros were recorded where the traces are missing. The goal here is to demonstrate that the inclusion of gaps in the data leads to artifacts in the least-squares migrated sections. A decimated version of the seismograms in Figure 1b, with two gaps in the data, one from −4 to 7 m and the other from 14 to 21 m, is shown in Figure 3a. The gaps are designed so that the data are both aliased and devoid of the small angle diffractivity.

Figures 3c and 3e show the final least-squares migrated section and the corresponding objective function when zero amplitude traces are assumed to be recorded in the gapped intervals, respectively. Many iterations are necessary to reduce the objective function value to 0.001, yet the migrated section is defocused. Figures 3d and 3f show, respectively, the final least-squares migrated section and the corresponding objective function when the zeroed traces are excluded from the data (i.e., no data are assumed to be recorded in the gapped intervals, so that they do not enter the matrix system of equations). Fewer iterations are necessary to reach the same objective function value and the migrated section is more focused than is that of Kirchhoff migrated section in Figure 3b. No regularization was applied to these data.

The LSM migrated section, however, has some aliasing artifacts due to the aliased nature of the data. The diffractor images on both sides of the model are nearly correct, but the diffractor images in the middle column are only partly focused at

their true location because of the exclusion of the small-angle diffraction data.

Effect of regularization

Aliasing artifacts can be attenuated by regularizing the equations that lead to the migrated section in Figure 3d. Regularization reduces the size of the model space by imposing constraints on the models. The dotted, dash-dotted, and dashed lines in Figure 4d show the frequency distributions of reflectivities for Figures 3b, 3c, and 3d, respectively. These frequency distributions indicate that a better reflectivity model estimate should produce a frequency distribution with a narrower peak. This requirement is satisfied for our L2 regularization scheme by using damping with the diagonal matrix $\mathbf{C}$ containing ones and zeros. The zero values are assigned to the columns containing the diffractors and their neighbors; ones to all other columns. The migrated section in Figure 4b shows that the aliasing artifacts were successfully damped. More importantly, most of the energy of the middle-column diffractors is focused at their true locations (Figure 4c). The resulting frequency distribution is a better approximation to the true distribution (Figure 4d). Note, however, that in complex cases the separation of aliased and non-aliased energy is not trivial; in those cases, either more sophisticated $\mathbf{C}$ matrices are necessary or the chosen objective function should better approximate the assumed true frequency distribution.

Incorrect migration velocity

The goal is to demonstrate the sensitivity of LSM to changes in the migration velocity. Figure 5 shows the results of using incorrect velocities by migrating the data with a velocity that is 10% lower than the actual velocity. As depicted in Figure 5, the LSM images for different objective function values diverge from the original model while predicting the observed data (compare Figure 5 with Figure 1).

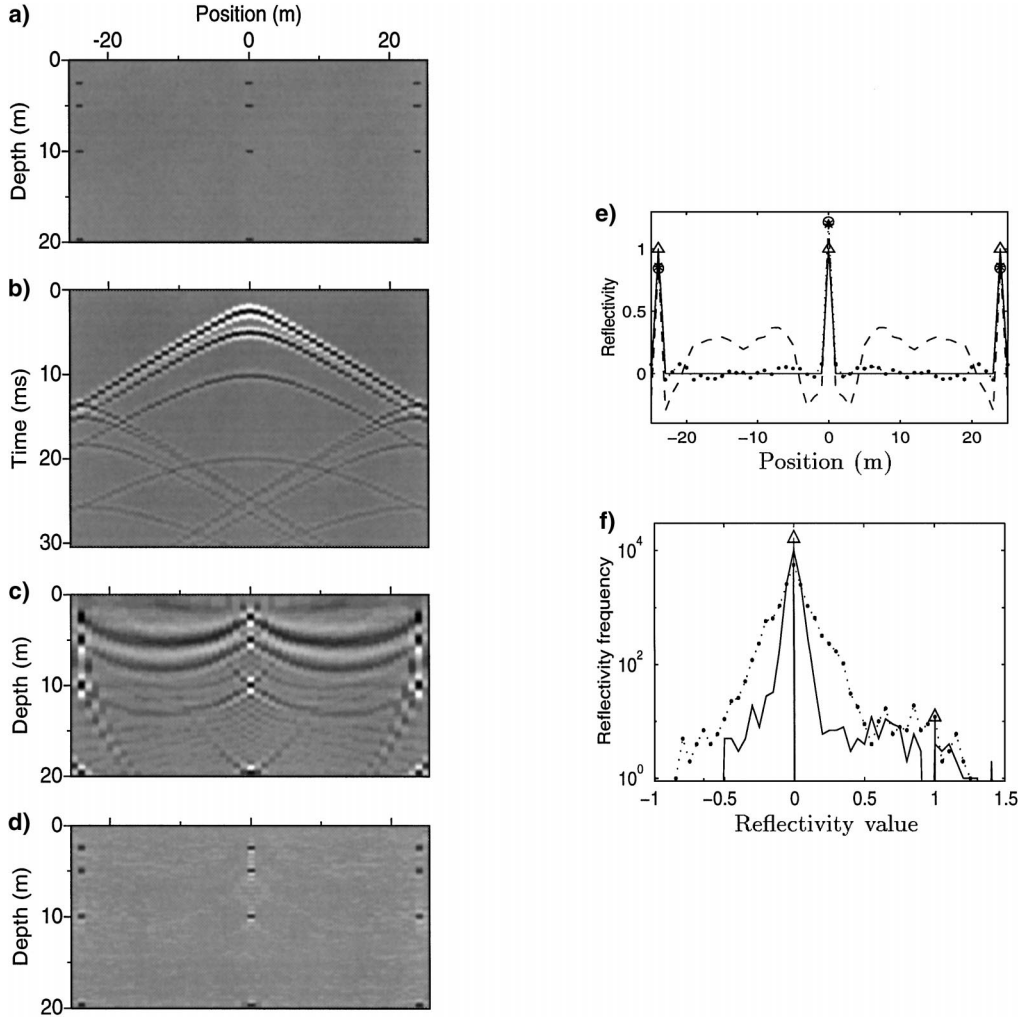

FIG. 1. Effects of limited aperture on the 12-point diffractor model. The velocity of the migration and the forward modeling correspond to the true parameters of the medium. (a) the 12-point diffractor model, (b) synthetic seismograms (common-shot gather with the source at 0 m and central source frequency of 1000 Hz), (c) Kirchhoff migrated section, (d) least-squares migrated section, (e) reflectivities at 5-m depth [the solid line shows the true reflectivities from (a), the stars and the circles show the estimated reflectivities from (c) and (d), respectively], (f) frequency distributions of the reflectivities [the two arrows at 0 and 1 indicate the true frequency distribution the; dotted and the solid lines show the distributions from (c) and (d), respectively].

The migration artifacts seen in these migrated sections follow two trends. First, the strong migration-ellipse artifacts in Figure 5a become more attenuated with increasing iteration number. Then, in order to predict the data, new model artifacts are introduced, as seen in Figures 5c and 5d. These trends can also be detected in the frequency distributions of these migrated sections (Figure 5f). The phenomenon of slight improvement in the model estimates followed by their deterioration is not unique to this example and is an indicator of incorrect migration parameters. Another indicator is that the convergence of the objective functions in cases of incorrect migration parameters (Figures 3d and 5e) is significantly slower than in cases with proper parameters (Figures 3f and 4a).

Fault model

Attenuating migration artifacts not only focuses the images of isolated diffractors, it also reveals otherwise hidden features.

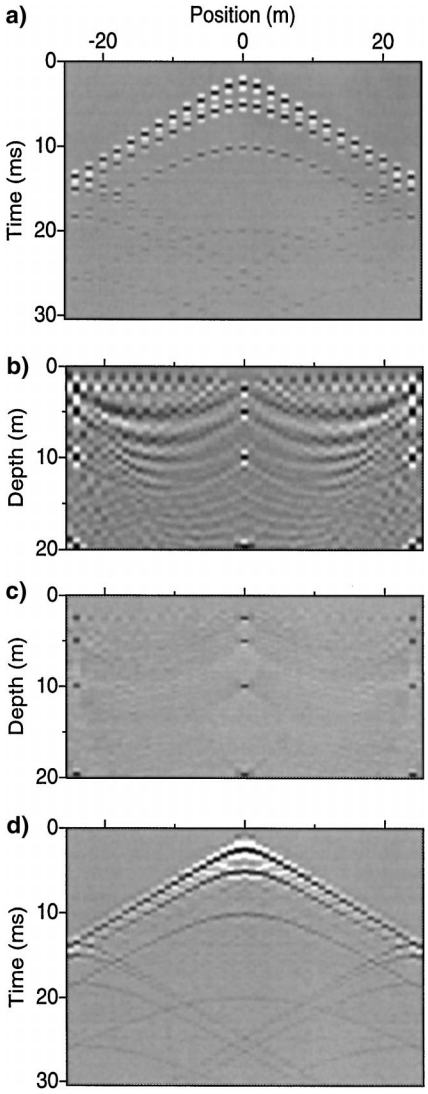

FIG. 2. Effects of coarse wavefield sampling on the 12-point diffractor model. (a) Synthetic seismograms where every other trace is missing, (b) Kirchhoff migrated section, (c) least-squares migrated section, (d) predicted unaliased seismograms using the least-squares migrated section.

The fault model in Figure 6a illustrates this claim. A common-shot gather (Figure 6b) was created where the source is at the 0-m horizontal position and the receivers have a 1-m spacing. Figure 6c shows the Kirchhoff migrated section, and Figure 6d shows the least-squares migrated section. Note the improvement in the depositional sequence at around 15-m depth. More comparisons between Kirchhoff migration and least-squares migration are shown in Nemeth (1996).

SEG/EAGE Overthrust data

The Society of Exploration Geophysicists (SEG) and European Association of Geoscientists and Engineers (EAGE) designed the acoustic Overthrust model shown in Figure 7a, and provided us with the corresponding poststack traces computed by a finite-difference solution to the acoustic wave equation. The data along a 2-D line are shown in Figure 7b. This example shows a successful application of LSM and highlights some of the problems with LSM at the same time. The problems usually arise when the LSM modeling operator does not take into account some wavemodes present in the data or it is not an accurate enough approximation of the actual wave propagation. In particular, some off-plane reflections are present in the data and the amplitude changes for some diffractors do not follow simple amplitude characteristics due to the complex geology.

First, we muted the observed data at 2 s and later to image only reflections originating above 4.5 km. In this way, the strong free-surface multiples were eliminated. Then, we applied Kirchhoff migration to the Overthrust data. The resulting migrated section is shown in Figure 8a. After that we applied LSM to the Overthrust data. The resulting migrated section is shown in Figure 8b. There, the data residual is 1%, and it took 31 iterations to reach this level. The problems pointed out above can be observed in the zone near a point at 6.5-km lateral position and 1.8-km depth (image resembling a partially collapsed diffractor, but its origin is an off-plane reflection) and near a point at 7-km lateral position and 3-km depth (partially collapsed diffractor). These artifacts are a clear reminder of possibilities and the limitations with using LSM (and other migration methods).

GROUND-PENETRATING RADAR DATA

In this section, the LSM method is applied to ground-penetrating radar (GPR) field data. GPR field data were collected at Point of the Mountain, near the southern edge of the Salt Lake Valley, Utah. Point of the Mountain is a large sand bar formed by the southward net transport of sand and gravel along the western shore of Pleistocene lake Bonneville (Machette, 1988). The survey site is located on the southern edge of Point of the Mountain (Figure 9) and is on an outcrop exposure. The purpose of the GPR experiment is to map the depositional sequences within the sand bar.

For the field measurements, 50-MHz antennas were used. Several common-offset radargrams were acquired along profiles A, B, and C. Along profile A, radargrams with offsets 2, 4, 6, 8, and 10 m were acquired. For profile B, radargrams with offsets 2, 4, 6, and 8 m were collected, whereas for profile C, a 2-m common-offset gather was collected. The interval between the adjacent traces is 0.5 m. The initial data processing included elimination of the high-frequency static shifts due to

the occasional source and receiver mispositioning, and low-pass filtering and removal of the strong first-arrival signals. No amplitude enhancement was applied to the traces. The velocity distribution for LSM was chosen as 0.15 m/ns (half of the speed of light in vacuum) for dry sands (Davis and Annan, 1989) in the upper 15 m, and 0.13 m/ns for the second layer corresponding to wetter sands. The attenuation rate for modeling and migration was chosen to be 0.01 dB/m and corresponds to typical attenuation values for dry sands (Davis and Annan, 1989).

The 8-m common-offset gather of profile A is depicted in Figure 10a. The gather has an acquisition gap in the 26–38 m interval. The preconditioned Kirchhoff migrated section is shown in Figure 10b. It contains two types of artifacts. First are the migration ellipses directly below the interval of the acquisition gap. The second type of artifact is the set of curved migration ellipses near the edges of the migrated section, resulting from the limited aperture of the data. As demonstrated in Figure 10c, these artifacts are attenuated after 20 iterations of the LSM method. Unlike the Kirchhoff migrated section, the width of the gap in the least-squares migrated section corresponds to the width of the gap in the common-offset gather, indicating a better lateral resolution for LSM than for Kirchhoff migration. The data residual decreased to 0.5% from the initial 36%. Figures 10d and 10e show the predicted common-offset gathers for Kirchhoff and least-squares migration, respectively. Consistent with the previous synthetic tests the least-squares migrated section provided an accurate prediction of the data.

Figure 11a shows the 8-m common-offset gather along profile A, except the traces are 1:4 subsampled. Most of the reflections are now aliased. Kirchhoff migration of these aliased data produces artificial reflectors (Figure 11b) with geologically unreasonable dips. To attenuate these reflectors, LSM is applied to the Figure 11a data where the regularization matrix is $\mathbf{C} = a(\partial/\partial x) + b(\partial/\partial z)$, $a = \cos(5^\circ)$, $b = \sin(5^\circ)$, and the argument of the trigonometric functions is the dip of the upper layers in Figure 11b. The least-squares migrated section after 20 iterations is depicted in Figure 11c. The data residual decreased to 0.3% from the initial 34%, and the recording footprint noise due to subsampling is significantly attenuated. Additionally, Figure 11d shows the predicted data from the Figure 11c model. Comparison of Figure 11d with Figure 10 a suggests a good correlation between the measured and predicted data.

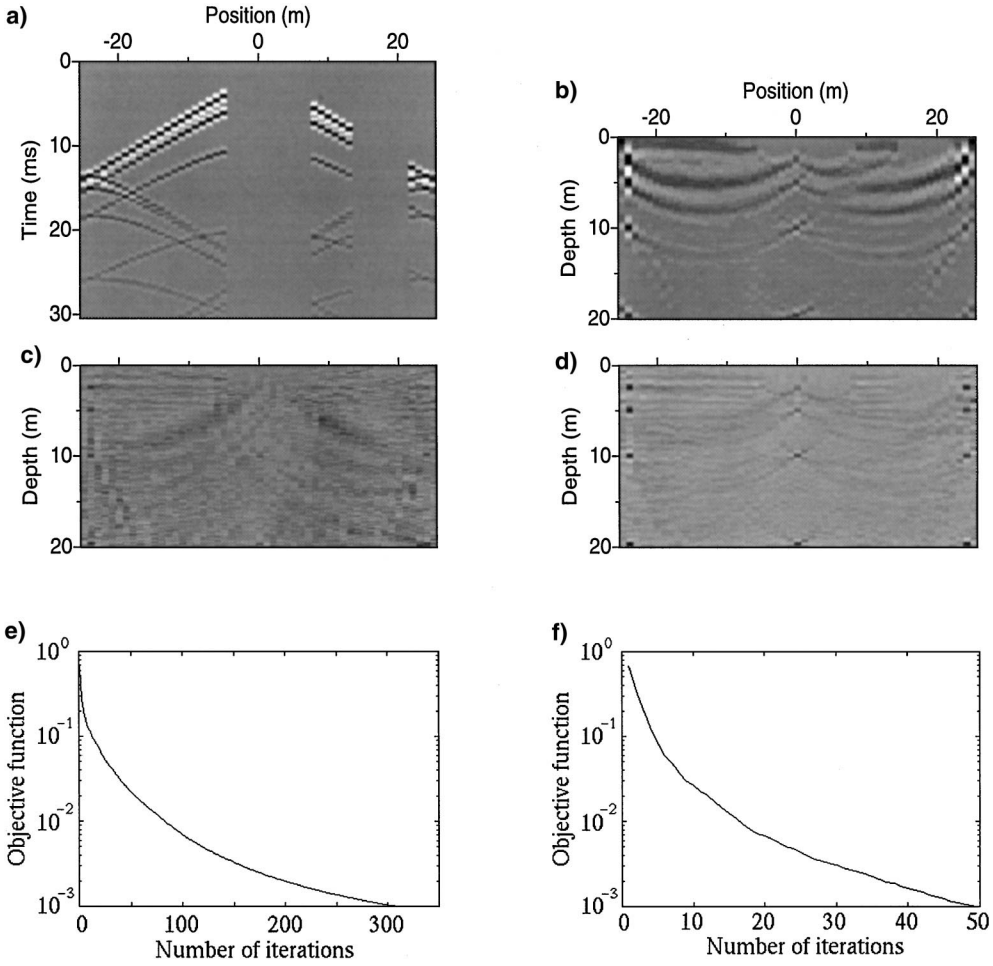

FIG. 3. Demonstration of migration artifacts associated with irregular gaps in the data for the 12-point diffractor model. (a) Synthetic seismograms, b) Kirchhoff migrated section, (c) Least-squares migrated section corresponding to the final iteration in (e), (d) Least-squares migrated section corresponding to the final iteration in (f), (e) normalized objective function obtained by treating the gaps in (a) as if zeros were recorded, (f) normalized objective function obtained by excluding the gaps from the data.

In the stacked migrated sections for each line (Figure 12), two distinctive depositional sequences can be distinguished. The first is about 10-m thick (from 2–3 m to about 13 m in the figures). It is characterized by layers moderately dipping (less than 5°) to the southwest. The second sequence is characterized by steeply dipping 25°–30°) layers. This interpretation is confirmed by the sedimentary features observed in a nearby open gravel quarry.

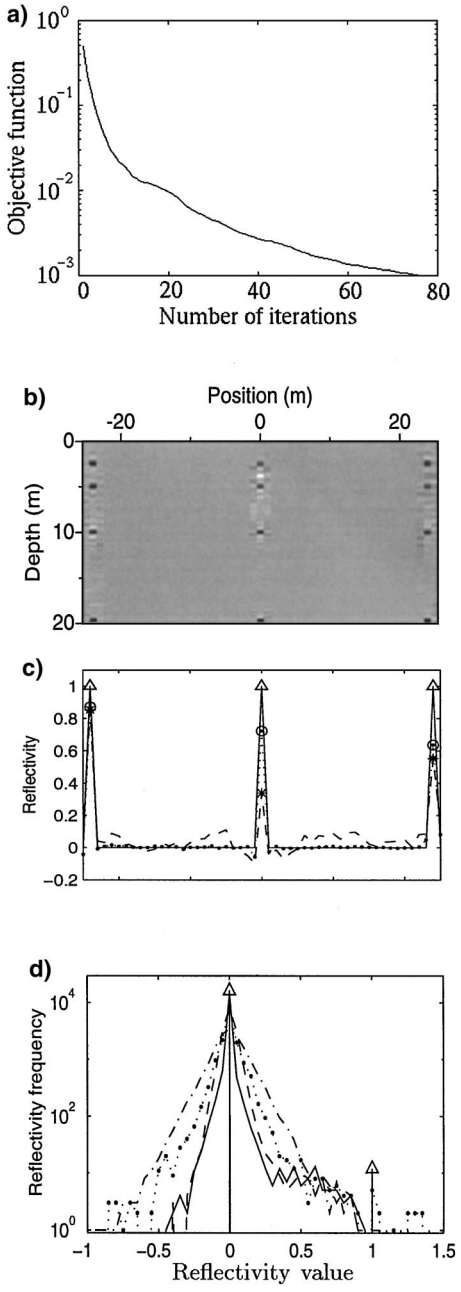

FIG. 4. The same as in Figure 3 except regularization is used to damp the aliased events. (a) Normalized objective function, (b) Least-squares migrated section, (c) Reflectivities at 5 m depth (the solid line shows the true reflectivities, the stars and the circles show the estimated reflectivities from Figure 3d and Figure 4b, respectively; note that regularization helps predict a more accurate reflectivity for the diffractor at 0 m horizontal position), (d) frequency distributions of the reflectivities in Figures 3 and 4 (the two arrows at 0 and 1 indicate the true frequency distribution, the dotted line corresponds to Figure 3b, the dash-dotted line to Figure 3c, the dashed line to Figure 3d, and the solid line to Figure 4b).

REVERSE VSP DATA

In this section, the LSM method is applied to reverse VSP (RVSP) field data for the purpose of imaging the subsurface reflectivity structure. The RVSP data were acquired near Friendswood, Texas, and the description of the experiment is given in Chen et al. (1990). The data consisted of 23 common-receiver gathers, with offsets ranging from 8 to 200 m. Each common-receiver gather contained 98 downhole sources evenly distributed over the depth range of 10 to 340 m. The interval between the adjacent downhole sources is 3 m and between the adjacent surface receivers is 8 m. The data preprocessing included the following steps: (1) bandpass filtering (90–600 Hz) to remove both the high- and low-frequency noise, (2) first-arrival traveltime picking, (3) direct-wave removal, (5) trace amplitude normalization, (6) f-k filtering to separate the up- and downgoing waves, (7) traveltime tomography to estimate the background velocity distribution. Figure 13b shows a seismogram recorded at a 135-m offset from the source well after the preprocessing.

The data were initially migrated using Kirchhoff migration. The resulting migrated section is shown in Figure 13c. The Kirchhoff migrated section contains familiar migration ellipses. These artifacts cannot be compensated by stacking all the prestack migrated sections because the wavefield sampling along the receiver axis is coarse, resulting in recording footprint noise in the stacked migrated section.

To remove these artifacts, least-squares migration is applied to the data. Here we employ regularization by using a linear filter operator $\mathbf{C} = \partial/\partial x$ and $\mathbf{C}^T = -\partial/\partial x$ in equation (6). The regularization parameter ε^2 is exponentially decreased from 2 to 0 as the iterations proceed. In this way, the regularization term dominates the early iterations, and the misfit term is dominant at later iterations. Figure 13d depicts the resulting least-squares migrated section. The migration ellipses are severely attenuated. The normed data residual decreased to 17% after 20 iterations from the initial 79%. The predicted seismograms using the Kirchhoff and least-squares migrated sections are shown in Figures 14a and 14b, respectively.

Then the LSM method was applied to the 23 common-receiver gathers, and prestack migrated sections were obtained. Before stacking the prestack migrated sections, a VSP residual moveout correction (Nemeth, 1994) was performed to compensate for distortion induced by the incorrect background velocity distribution. The corrected stacked migrated section is shown in Figure 14d. Figure 14c shows a good correlation between the migrated data trace and the well-log synthetics near the source well.

MARINE CDP POSTSTACK DATA

The LSM method was applied to CDP poststack data collected in the Gulf of Mexico (courtesy of Mobil). Figure 15a shows the poststack data along a 2-D line. This example demonstrates the benefits of increased spatial resolution with LSM images, even though the data are not spatially undersampled.

These data were migrated with conventional Kirchhoff migration to give the migrated image shown in Figure 15b. The LSM images after 30 iterations are shown in Figure 15c. As in the synthetic data examples, the LSM images have reflectivity strengths that are better balanced. The CDP interval for these data is 25 m, so these data were not aliased and, consequently, the acquisition footprint due to receiver sampling is small.

DISCUSSION AND CONCLUSIONS

A least-squares migration algorithm is presented that suppresses migration artifacts (i.e., recording footprint noise) associated with incomplete data, where the data incompleteness arises from a finite recording aperture, coarse source/receiver sampling, and irregular recording gaps. The artifacts associated with incomplete data can sometimes be attenuated by stacking many migrated sections. There are many cases, however, when the fold of the data does not allow for effective stacking. Such cases include 3-D acquisition gaps, coarsely sampled data (i.e., a sparse-offset VSP survey), the edges of the migrated sections (surface common-midpoint 2-D, 3-D sections) or bad-data sections in the observed data. Least-squares migration can improve the lateral resolution of the subsurface images without requiring costly acquisition of denser data sets.

Kirchhoff migration and the measured incomplete data are often augmented to reduce the artifacts due to incomplete data, using some a priori information. For Kirchhoff migration these procedures involve tapering of the traces near the edges of the recording aperture and limiting the incidence angles for imaging; for the measured data, interpolation and extrapolation can be used to fill in the missing data gaps. If the a priori information is sufficient, these procedures greatly improve the migrated section. In the common case of insufficient or inaccurate a priori information, however, these procedures might be ineffective and introduce additional artifacts in the migrated sections. In contrast, least-squares migration helps reduce artifacts in a natural way by generating the model that predicts the observed data in a least-squares sense.

The present version of LSM uses a preconditioning matrix to speed up convergence. The off-diagonal elements of the preconditioning matrix are zero, and the diagonal elements are calculated analytically as described in Appendix B. This version of LSM requires a few tens of iterations to compute the migrated sections shown in this paper. With an identity matrix $\mathbf{I}$ preconditioner, it would have taken a few hundred iterations to arrive at the same results! A few tens of iterations, however, are computationally still too expensive for many data sets. Better preconditioning matrices might reduce the number of

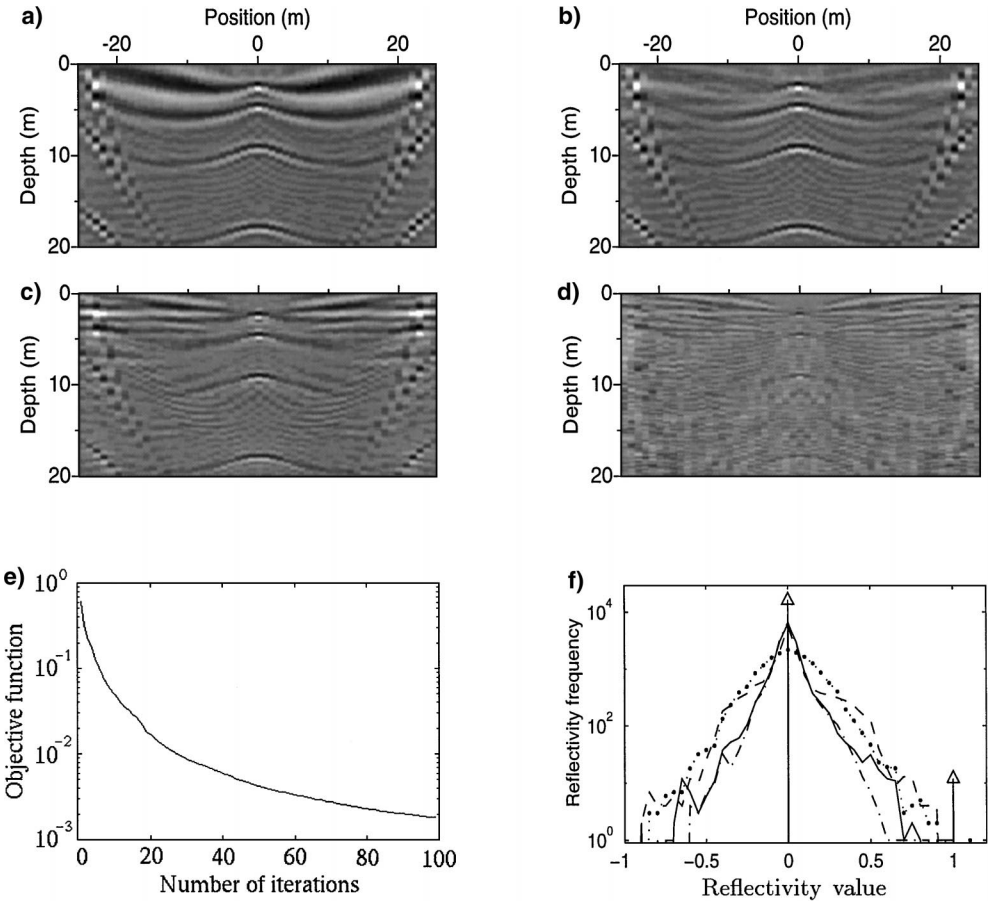

FIG. 5. The same as in Figure 1 except the migration velocity is taken to be 0.9 times the true velocity. (a) Kirchhoff migrated section, (b) migrated section at objective function value of 0.25, (c) migrated section at 0.05, (d) migrated section at 0.0018, (e) normalized objective function, (f) frequency distributions of the reflectivities in the above migrated sections [the two arrows at 0 and 1 indicate the true frequency distribution, the dashed line corresponds to (a), the solid line to (b), the dash-dotted line to (c), and the dotted line to (d)].

iterations to an acceptable few. These preconditioning matrices should be based on the sparse off-diagonal structure of the Hessian matrix.

It is an interpretive decision to choose the number of iterations or the acceptable data residual level. Our experience indicates that for good quality seismic data a residual level of 1–3% is acceptable. At this level, the amplitudes in the residual data are already evenly distributed and reach the noise level. We followed these two stopping criteria except for simple numerical models with no added noise.

Careful data management also decreases the relative cost of LSM in subsequent iterations. In general case LSM costs $R = 1 + 2\,iter$ Kirchhoff migrations, where $iter$ is the final iteration in LSM. If we can split migration into (precomputed) reusable and nonreusable parts, the computational ratio becomes $R = 1 + 2\,iter[(nr + io)/(nr + ru)]$, where nr is the computational time for the nonreusable part, ru is the computational time for the reusable part and io is the access time for the reusable part. In case of LSM, ru includes the traveltime

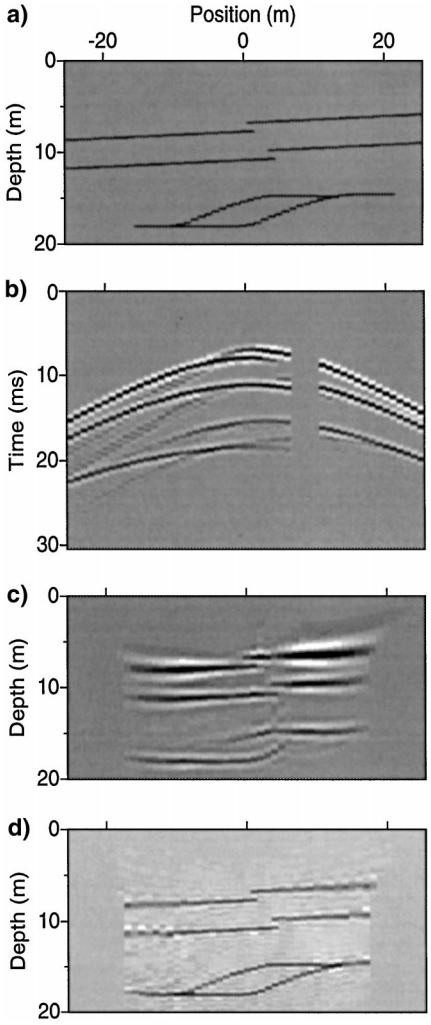

FIG. 6. The fault model used to demonstrate the capability of LSM to image hidden or weaker diffractors. (a) The fault model, (b) synthetic seismograms, (c) Kirchoff migrated section, (d) least-squares migrated section.

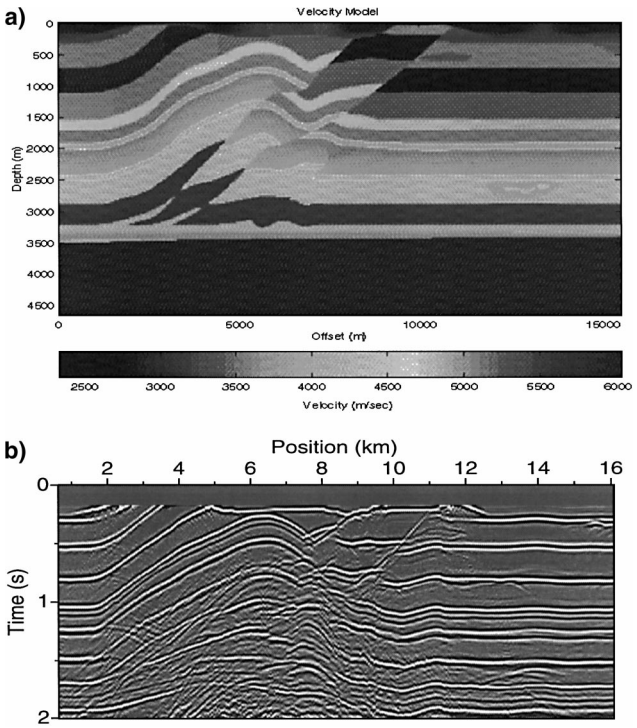

FIG. 7. Test with the Overthrust velocity model. (a) Overthrust velocity model associated with the Phase A data set (SEG/EAGE 3-D Modeling Series No. 1), (b) zero-offset Phase A data set associated with the SEG Overthrust model.

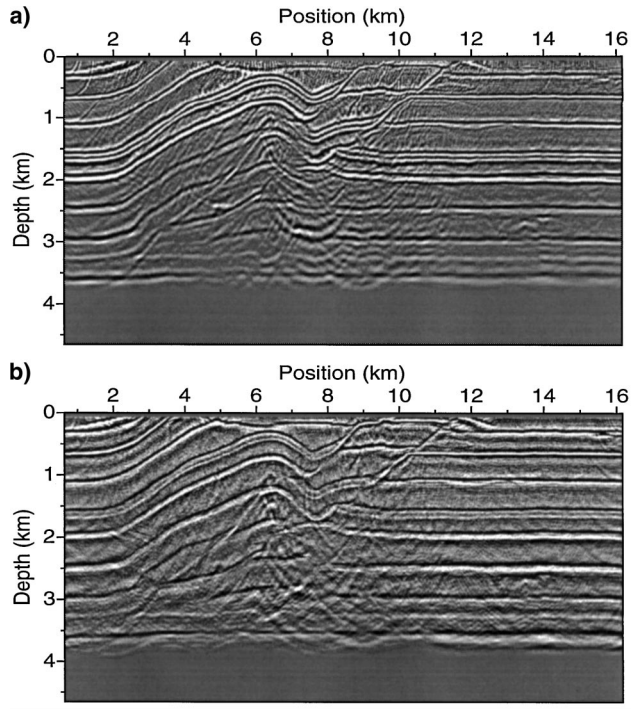

FIG. 8. Poststack migration images of the SEG Overthrust model. (a) Image obtained by using Kirchhoff migration [automatic gain control (AGC) has been applied], (b) image obtained by least-squares migration.

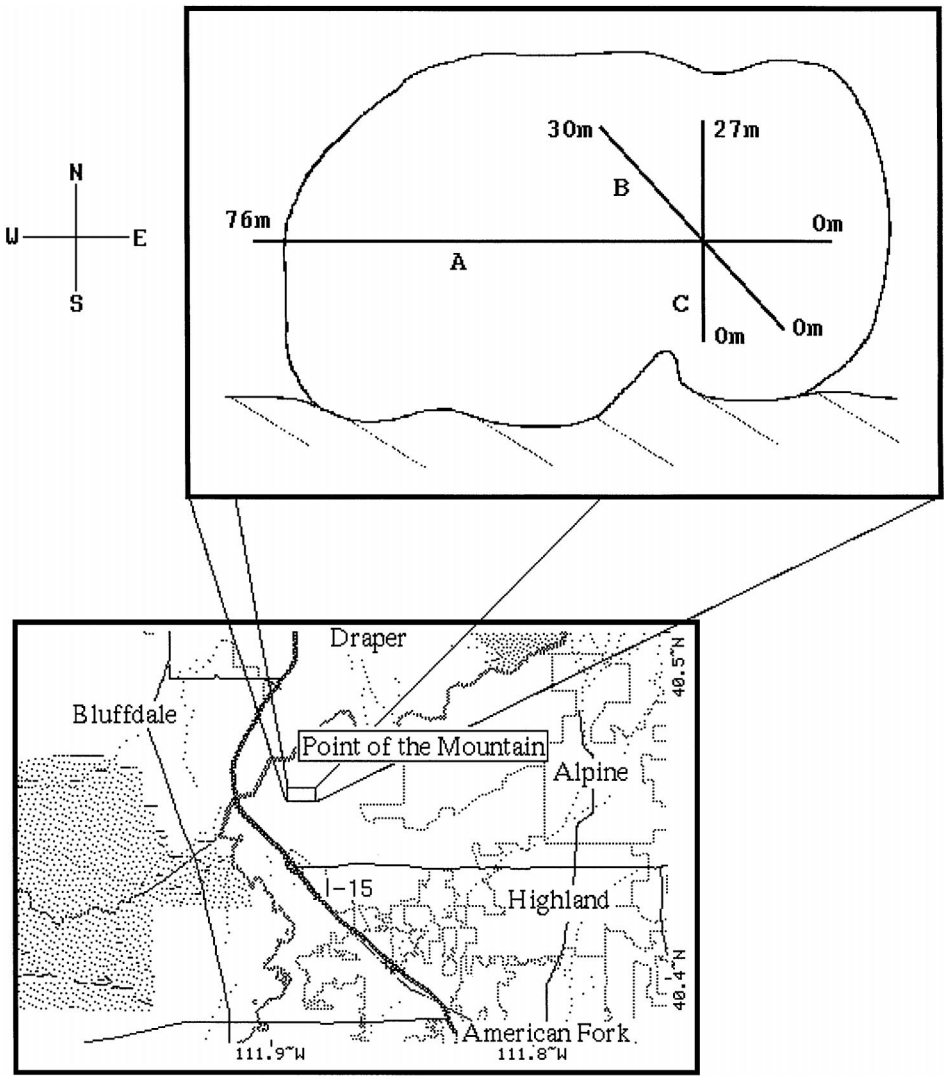

FIG. 9. Sketch of the field experiment site at Point of the Mountain, Utah. Point of the Mountain is a sand bar formed along the shoreline of Pleistocene lake Bonneville. The experiment site is an outcrop exposure with sand gravel and a thin veneer of cemented gravel on the surface. South of the experiment site, Point of the Mountain abruptly drops off due to the depositional bar form. In other directions, the surface is covered with soil and vegetation. The outline of the site is marked with a solid line.

table formation and provides significant saving for large models where $io \ll ru$.

Further improvements to LSM can be made by (1) finding the means to determine the proper regularization matrix $\underline{\mathbf{C}}$, (2) using non-Gaussian assumptions about the model and data residual distributions, and (3) making LSM more robust in the presence of noise.

ACKNOWLEDGMENTS

We are grateful for the financial support from the members of the 1994–1996 University of Utah Tomography and Modeling/Migration (UTAM) Consortium. We thank Exxon Production Research for providing the Friendswood RVSP data. This research was partially supported by NSF grant EAR-9315787. The comments and suggestions of Associate Editor George McMechan, Ken Larner, and Paul Stoffa, and an anonymous reviewer are gratefully acknowledged.

REFERENCES

Beydoun, W., and Mendes, M., 1989, Elastic ray-Born L^2 migration/inversion: Geophys. J. Internat., **97**, 151–160.
Bleistein, N., 1984, Mathematical methods for wave phenomena: Academic Press.
Carrion, P., Sato, H., and Buono, A., 1991, Wavefronts sets analysis of limited aperture migration sections: Geophysics, **56**, 778–784.
Chen, S. T., Zimmermann, L. J., and Tugnait, J. K., 1990, Subsurface imaging using reversed vertical seismic profiling and crosshole tomographic methods: Geophysics, **55**, 1478–1487.
Claerbout, J. F., 1992, Earth soundings analysis: Processing versus inversion: Blackwell Scientific Publications, Inc.
Cole, S., and Karrenbach, M., 1992, Least squares Kirchhoff migration: SEP-75, 101–110; available at http://sepwww.stanford.edu/public/docs/sep75/toc_html/index.html.
Davis, J., and Annan, A., 1989, Ground-penetrating radar for high resolution mapping of soil and rock stratigraphy: Geophys. Prosp., **37**, 531–551.
Duquet, B., 1996, Improving seismic imaging in complex geologic structures: Ph.D. diss. (in French), Univ. of Paris XIII.
Gray, S., 1997, True-amplitude seismic migration: A comparison of three approaches: Geophysics, **62**, 629–638.

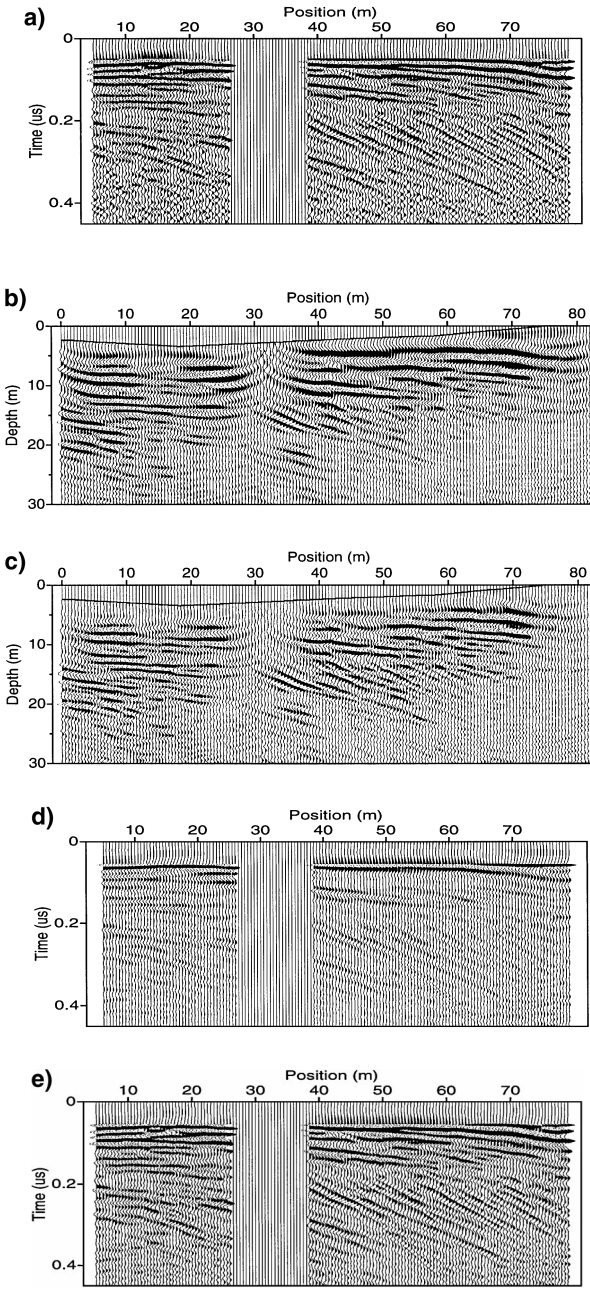

FIG. 10. Common-offset sections (offset equals 8 m) for ground penetrating radar data collected along profile A at Point of the Mountain, Utah. The upper solid line in the migrated sections delineates the topographic surface. (a) Observed radargrams, (b) Kirchhoff migrated section, (c)least-squares migrated section, (d) radargrams predicted with Kirchhoff migrated section, (e) radargrams predicted with least-squares migrated section. The same AGC gain is used to display (a), (d), and (e).

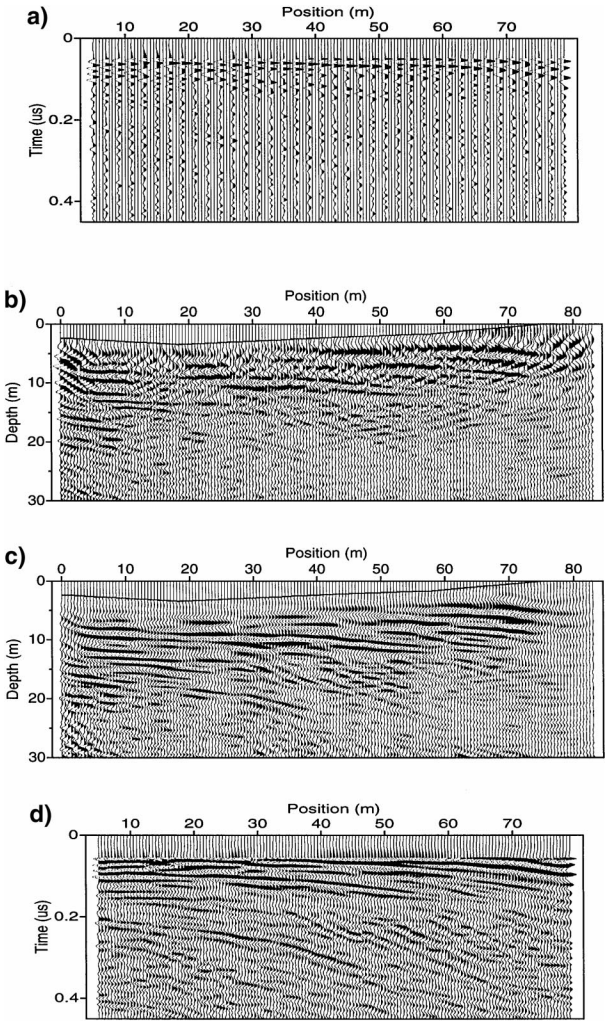

FIG. 11. Test with the 1:4 spatially subsampled 8-m common-offset section of profile A. (a) Observed radargrams, (b) Kirchhoff migrated section, (c) least-squares migrated section, (d) predicted radargrams using the least-squares migrated section.

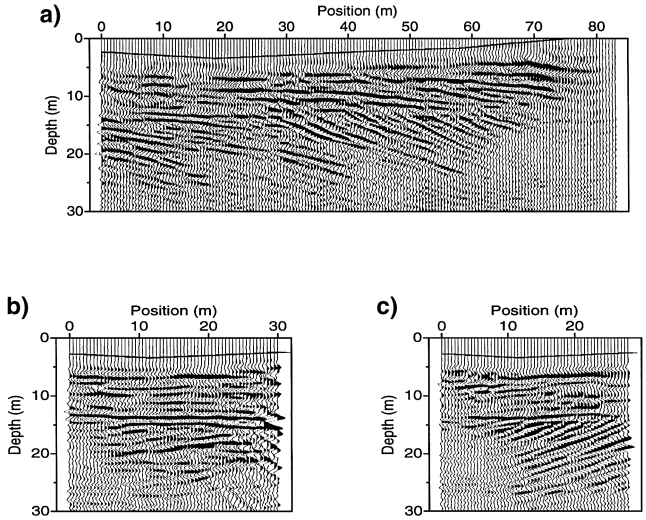

FIG. 12. Migrated sections obtained by LSM of the GPR data for (a) profile A obtained from stacking migrated common-offset gathers at offsets of 2, 4, 6, 8, and 10 m; (b) profile B obtained from stacking migrated common-offset gathers at offsets of 2, 4, 6, 8 m; (c) profile C obtained from common-offset gather at 2 m offset. The start and end positions of the migrated sections correspond to the ones on the sketch in Figure 9. Two depositional sequences can be distinguished in the migrated sections.

Hanitzsch, C., Schleicher, J., and Hubral, P., 1994, True-amplitude migration of 2D synthetic data: Geophys. Prosp., **42**, 445–462.
Kak, A., and Slaney, M., 1988, Principles of computerized tomographic imaging: Inst. Electr. Election. Eng. Press.
Lambare, G., Virieux, J., Madariaga, R., and Jin, S., 1992, Iterative asymptotic inversion in the acoustic approximation: Geophysics, **57**, 1138–1154.
LeBras, R., and Clayton, R., 1988, An iterative inversion of backscattered acoustic waves: Geophysics, **53**, 501–508.
Machette, M., 1988, In the footsteps of G. K. Gilbert—Lake Bonneville and neotectonics of the eastern Basin and Range Province: Geol. Soc. Am. Field Trip Guidebook 12: Utah Geol. Min. Surv. Miscellaneous Publication 88-1.
Miller, D., Oristaglio, M., and Beylkin, G., 1987, A new slant on seismic imaging: Migration and integral geometry: Geophysics, **52**, 943–964.

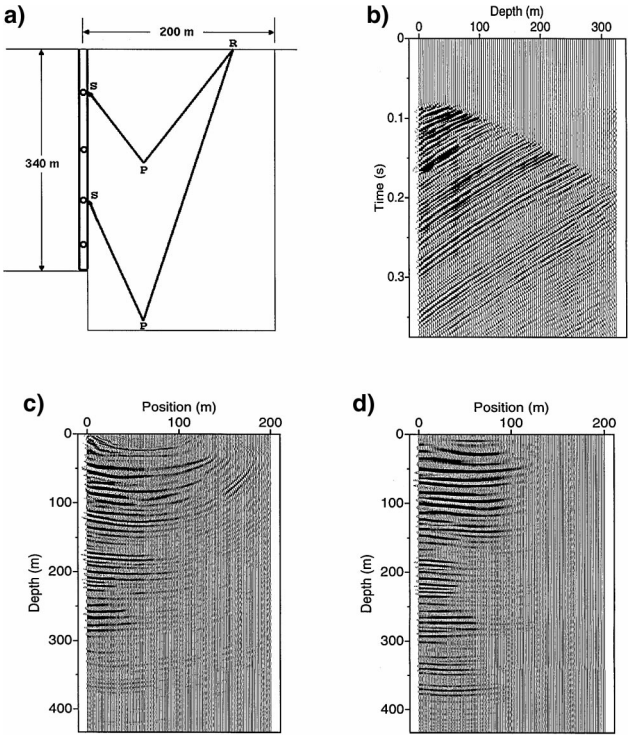

FIG. 13. The Friendswood RVSP experiment. (a) Sketch of the acquisition geometry (S = source, P = reflector, R = receiver, the reflections on common-receiver gathers are along the raypaths $\overline{RPS}$), (b) processed common-receiver gather (the receiver is at 135-m horizontal position), (c) Kirchhoff migrated section, (d) least-squares migrated section. Note that the migration artifacts are attenuated in (d).

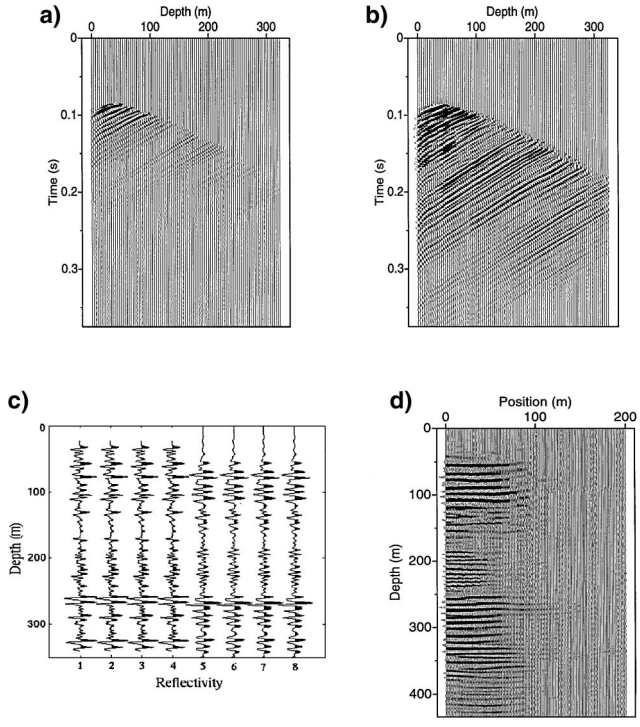

FIG. 14. The Friendswood RVSP experiment. (a) Common-receiver gather predicted with Kirchhoff migrated section, (b) common-receiver gather predicted with least-squares migrated section, (c) well-log synthetics (traces 1–4) and migrated data trace (5–8) at the source well, (d) the stacked migrated section after residual moveout correction.

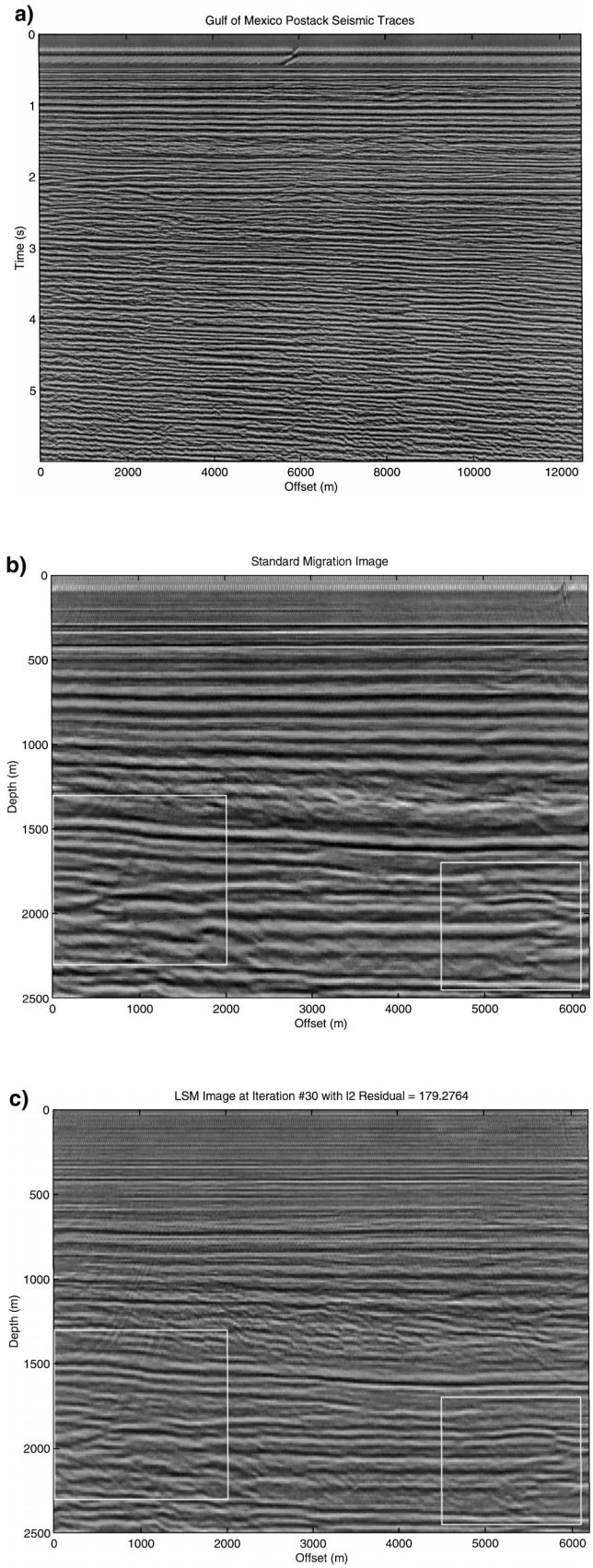

FIG. 15. The Gulf of Mexico data experiment. (a) Poststack traces; (b) Poststack Kirchhoff migrated section; (c) Poststack LSM migrated section. It appears that the LSM migrated section is more highly resolved than the Kirchhoff migration section.

Mittet, R., Hokstad, K., Helgesen, J., and Canadas, G., 1997, Imaging offset VSP data with an elastic iterative migration scheme: Geophys. Prosp. **45**, 247–268.
Nemeth, T., 1994, Migration and coherency stacking of the Friendswood reverse VSP data: 64th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 97–100.
——— 1996, Imaging and filtering by least-squares migration: Ph.D. diss., Univ. of Utah.
Ronen, S., Nichols, D., Bale, R., and Ferber, R., 1995, Dealiasing DMO: good-pass, bad-pass and unconstrained: 65th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 743–746.
Schuster, G. T., 1993, Least-squares crosswell migration: 63th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 110–113.
Sevink, A., and Herman, G., 1994, Fast iterative solution of sparsely sampled seismic inverse problems: Inverse Problems, **10**, 937–948.
Tarantola, A., 1984, Inversion of seismic reflection data in the acoustic approximation: Geophysics, **49**, 1259–1266.
——— 1987, Inverse problem theory. Methods for data fitting and model parameter estimation: Elsevier Science Publ. Co.
Wu, Y., and McMechan, G. A., 1996, Elastic full-waveform inversion for earthquake source parameters: Geophys. J. Internat., **127**, 61–74.

APPENDIX A

THE INTEGRAL EQUATION FORMULATION OF THE HESSIAN MATRIX $\underset{\sim}{\mathbf{L}}^T\underset{\sim}{\mathbf{L}}$ FOR KIRCHHOFF MIGRATION

Equation (1) can be written explicitly as

$$p(r,t\mid s,0)=\int m(x)W(t)*G(r,t\mid x,0)\,*\ddot{G}(x,t\mid s,0)\,dx, \tag{A-1}$$

where $G(x,t\mid s,0)$ and $G(r,t\mid x,0)$ are Green's functions for the sources at s and the scattering point at x and for same at the scattering point and the receivers at r, respectively, and $W(t)$ is the source wavelet (Bleistein, 1984). The zeroth-order asymptotic Green's function is

$$G(x,t\mid x',0)=A_{xx'}\,\delta(t-\tau_{xx'}), \tag{A-2}$$

where $A_{xx'}$ is the solution of the transport equation and $\delta(t-\tau_{xx'})$ is the traveltime between x and x'.

Substituting this Green's function into equation (A-1) yields, for those sources and receivers that are sampled by the $h_s(s)$ source sampling function and the $h_r(r)$ receiver sampling function,

$$p(r,t\mid s,0)=\int m(x)A_{sx}A_{xr}\,\ddot{W}(t-\tau_{sx}-\tau_{xr})\,dx. \tag{A-3}$$

Similarly, we can get the Kirchhoff migration equation

$$m_k(x)=\int dsh_s(s)\int drh_r(r)\times\left(\int p(r,t\mid s,0)A_{sx}A_{xr}\,\ddot{W}(t-\tau_{sx}-\tau_{xr})\,dt\right). \tag{A-4}$$

Substituting equation (A-3) into equation (A-4), and interchanging the integrals, the Kirchhoff migration equation can also be written as

$$m_k(x)=\int dx'm(x')\int dsh_s(s)A_{sx}A_{sx'}\int drh_r(r)A_{xr}A_{x'r}\times\left(\int\ddot{W}(t-\tau_{sx}-\tau_{xr})\ddot{W}(t-\tau_{sx'}-\tau_{x'r})\,dt\right). \tag{A-5}$$

The kernel of the outermost integral, $K(x,x')$, can be written as

$$K(x,x')=\int dsh_s(s)\int drh_r(r)A_{sx}A_{sx'}A_{xr}A_{x'r}\times R(\tau_{sx}+\tau_{xr}-\tau_{sx'}-\tau_{x'r}), \tag{A-6}$$

where $R(\tau_{sx}+\tau_{xr}-\tau_{sx'}-\tau_{x'r})$ denotes the temporal crosscorrelation of $\ddot{\mathbf{W}}$ with its $(\tau_{sx}+\tau_{xr}-\tau_{sx'}-\tau_{x'r})$ shifted version as seen in the innermost integral in equation (A-5). Then the Kirchhoff migration equation is written as

$$m_k(x)=\int K(x,x')m(x')\,dx', \tag{A-7}$$

and its discretized version can be represented as

$$\mathbf{m}_k=\underset{\sim}{\mathbf{K}}\mathbf{m}, \tag{A-8}$$

where matrix $\underset{\sim}{\mathbf{K}}$ is also the Hessian matrix $\underset{\sim}{\mathbf{L}}^T\underset{\sim}{\mathbf{L}}$. Kernel $K(x,x')$ can be regarded as a projection applied to the function $m(x)$ to give $m_k(x)$. For a fixed diffractor point x_f, $K(x,x_f)$ shows the response of the system to a point diffractor source. If only one trace is migrated, the response is the migration-ellipse impulse response. If a sufficient number of traces is used, the response will approximately reconstruct the point reflectivity at x_f from the migration ellipses.

APPENDIX B

THE PRECONDITIONED AND REGULARIZED CONJUGATE GRADIENT SCHEME

In this section, the preconditioned and regularized conjugate gradient scheme is derived. Equation (6) also can be written as

$$\begin{pmatrix}\underset{\sim}{\mathbf{L}}\\ \varepsilon\underset{\sim}{\mathbf{C}}\end{pmatrix}^T\begin{pmatrix}\underset{\sim}{\mathbf{L}}\\ \varepsilon\underset{\sim}{\mathbf{C}}\end{pmatrix}\mathbf{m}=\begin{pmatrix}\underset{\sim}{\mathbf{L}}\\ \varepsilon\underset{\sim}{\mathbf{C}}\end{pmatrix}^T\begin{pmatrix}\mathbf{p}_o\\ \varepsilon\underset{\sim}{\mathbf{C}}\mathbf{m}_{apr}\end{pmatrix}. \tag{B-1}$$

To speed up convergence, preconditioning of the previous equations is necessary. Choose the preconditioning matrix $\underset{\sim}{\mathbf{W}}$

$$\underset{\sim}{\mathbf{W}}^T\begin{pmatrix}\underset{\sim}{\mathbf{L}}\\ \varepsilon\underset{\sim}{\mathbf{C}}\end{pmatrix}^T\begin{pmatrix}\underset{\sim}{\mathbf{L}}\\ \varepsilon\underset{\sim}{\mathbf{C}}\end{pmatrix}\underset{\sim}{\mathbf{W}}\underset{\sim}{\mathbf{W}}^{-1}\mathbf{m}=\underset{\sim}{\mathbf{W}}^T\begin{pmatrix}\underset{\sim}{\mathbf{L}}\\ \varepsilon\underset{\sim}{\mathbf{C}}\end{pmatrix}^T\begin{pmatrix}\mathbf{P}_o\\ \varepsilon\underset{\sim}{\mathbf{C}}\mathbf{m}_{apr}\end{pmatrix}. \tag{B-2}$$

Based on this equation, a good $\underset{\sim}{\mathbf{W}}\underset{\sim}{\mathbf{W}}^T$ approximates the inverse of the Hessian matrix as seen by left-multiplying equation (B-2) with $\underset{\sim}{\mathbf{W}}$. An approximate choice for $\underset{\sim}{\mathbf{W}}$ might be a diagonal matrix with $1/\sqrt{diag(\underset{\sim}{\mathbf{L}}^T\underset{\sim}{\mathbf{L}}) + \varepsilon^2 diag(\underset{\sim}{\mathbf{C}}^T\underset{\sim}{\mathbf{C}})}$ where the first term can be calculated analytically as in asymptotic inversion of Bleistein (1984) and the computation of the second term is also feasible. Introducing new variables $\hat{\underset{\sim}{\mathbf{L}}} = \underset{\sim}{\mathbf{L}}\underset{\sim}{\mathbf{W}}$, $\hat{\underset{\sim}{\mathbf{C}}} = \underset{\sim}{\mathbf{C}}\underset{\sim}{\mathbf{W}}$ and $\hat{\mathbf{m}} = \underset{\sim}{\mathbf{W}}^{-1}\mathbf{m}$, equation (B-2) can be rewritten as

$$\left(\hat{\underset{\sim}{\mathbf{L}}}^T\hat{\underset{\sim}{\mathbf{L}}} + \varepsilon^2\hat{\underset{\sim}{\mathbf{C}}}^T\hat{\underset{\sim}{\mathbf{C}}}\right)\hat{\mathbf{m}} = \hat{\underset{\sim}{\mathbf{L}}}^T\mathbf{p}_o + \varepsilon^2\hat{\underset{\sim}{\mathbf{C}}}^T\underset{\sim}{\mathbf{C}}\mathbf{m}_{apr}. \qquad \text{(B-3)}$$

The conjugate gradient scheme based on equation (B-3) can be written as

$$\hat{\mathbf{m}}_0 = 0; \quad \mathbf{d}_o = \begin{pmatrix} \mathbf{p}_o \\ \underset{\sim}{\mathbf{C}}\mathbf{m}_{apr} \end{pmatrix};$$

$$\mathbf{g}_0 = \begin{pmatrix} \hat{\underset{\sim}{\mathbf{L}}} \\ \varepsilon^2\hat{\underset{\sim}{\mathbf{C}}} \end{pmatrix}^T \begin{pmatrix} \mathbf{p}_o \\ \underset{\sim}{\mathbf{C}}\mathbf{m}_{apr} \end{pmatrix}; \quad \mathbf{h}_0 = \mathbf{g}_0;$$

for $n = 0, 1, 2, \ldots$

$$\mathbf{f}_n = \begin{pmatrix} \hat{\underset{\sim}{\mathbf{L}}} \\ \hat{\underset{\sim}{\mathbf{C}}} \end{pmatrix} \mathbf{h}_n$$

$$\alpha_n = \frac{\|\mathbf{g}_n\|^2}{\|\hat{\underset{\sim}{\mathbf{L}}}\mathbf{h}_n\|^2 + \varepsilon^2\|\hat{\underset{\sim}{\mathbf{C}}}\mathbf{h}_n\|^2}$$

$$\hat{\mathbf{m}}_{n+1} = \hat{\mathbf{m}}_n + \alpha_n\mathbf{h}_n; \quad \mathbf{d}_{n+1} = \mathbf{d}_n - \alpha_n\mathbf{f}_n$$

$$\mathbf{g}_{n+1} = \begin{pmatrix} \hat{\underset{\sim}{\mathbf{L}}} \\ \varepsilon^2\hat{\underset{\sim}{\mathbf{C}}} \end{pmatrix}^T \mathbf{d}_{n+1}$$

if $\mathbf{g}_{n+1} < \eta$ or $\mathbf{d}_{n+1}/\mathbf{d}_o < \mu$ then quit

$$\beta_n = \frac{\|\mathbf{g}_{n+1}\|^2}{\|\mathbf{g}_n\|^2}$$

$$\mathbf{h}_{n+1} = \mathbf{g}_{n+1} + \beta_n\mathbf{h}_n$$

end;

$$\mathbf{m} = \underset{\sim}{\mathbf{W}}\hat{\mathbf{m}}$$

where $\mathbf{g}_n$ is the nth regularized gradient vector, $\mathbf{h}_n$ is the nth conjugate direction vector, $\mathbf{d}_n$ is the nth regularized data residual vector (the zeroth iteration is the data itself), $\mathbf{f}_n$ is the nth forward modeled conjugate direction vector, α_n and β_n are weighting coefficients, and η and μ are predetermined accuracy limits. This scheme implicitly calculates the matrix-vector operations. The matrices are never formed explicitly; instead, they are replaced by a forward and an adjoint modeling subroutine that returns the resulting vector.