GEOPHYSICS, VOL. 65, NO. 2 (MARCH-APRIL 2000); P. 574–583, 11 FIGS., 5 TABLES.

Separation of signal and coherent noise by migration filtering

Tamas Nemeth*, Hongchuan Sun‡, and Gerard T. Schuster‡

ABSTRACT

A key issue in wavefield separation is to find a domain where the signal and coherent noise are well separated from one another. A new wavefield separation algorithm, called migration filtering, separates data arrivals according to their path of propagation and their actual moveout characteristics. This is accomplished by using forward modeling operators to compute the signal and the coherent noise arrivals. A linearized least-squares inversion scheme yields model estimates for both components; the predicted signal component is constructed by forward modeling the signal model estimate. Synthetic and field data examples demonstrate that migration filtering improves separation of P-wave reflections and surface waves, P-wave reflections and tube waves, P-wave diffractions, and S-wave diffractions. The main benefits of the migration filtering method compared to conventional filtering methods are better wavefield separation capability, the capability of mixing any two conventional transforms for wavefield separation under a general inversion framework, and the capability of mitigating the signal and coherent noise crosstalk by using regularization. The limitations of the method may include more than an order of magnitude increase in computation costs compared to conventional transforms and the difficulty of selecting the proper modeling operators for some wave modes.

INTRODUCTION

All seismic data sets contain both signal and noise, and the principal task of data processing is to extract the signal component from the noisy data. In many cases noise is coherent, consisting of several unwanted wave modes. Examples of coherent noise are multiple reflections, air waves, surface waves, tube waves, and converted waves. To remove coherent noise, we sometimes transform the data into a domain where the signal and coherent noise are well separated from one another.

Many transforms offer convenient ways to filter out the unwanted wave modes. The most common transforms are the Fourier and Radon transforms, where the filtering can eliminate specified frequency bands or dip ranges in the data, respectively. Other useful operators and separation methodologies are presented in many papers. For example, Harlan et al. (1984) use a local linear Radon transform and hyperbolic stack to separate reflections from diffractions. Because the Radon transform allows for limited dips and apertures in the summation window, they are able to cleanly separate the reflections from the other wave modes. Biondi (1992) and Kostov and Biondi (1991) use localized beam stacks rather than slant stacks to improve the resolution of the seismic image. Thorson and Claerbout (1985) separate primary and multiple reflections by applying a normal moveout to the seismic data and mute the multiples in the velocity–time domain. Lumley et al. (1995) extend this method with the use of regularization and a time-variant modeling operator. Abma (1995) uses a least-squares inversion method to predict both the signal and the noise in CDP data. He designs signal- and noise-prediction filters and applies these filters to the data in an inversion process to recover either the noise or signal. In his case the prediction operators are restricted to linear moveout operations. Blonk et al. (1995) introduce an elastodynamic inverse scattering method to estimate and remove cross-line scattered surface waves that show hyperbolic moveout and resemble body wave reflections. Duquet and Marfurt (1997) filter coherent noise during the prestack depth migration process by applying a depth-variant velocity filter. The filter design requires a least-squares migration followed by residual moveout velocity estimation. Hu and White (1998) use adaptive beamforming for multiple suppression. They apply minimum variance unbiased beam-forming to extract coherent signals and obtain superior results compared to the conventional Radon transforms. Spitz (1998) also uses prediction filters to model and subtract coherent noise in a pattern recognition approach. He suggests the use of higher dimensional spaces to decrease

Manuscript received by the Editor September 29, 1998; revised manuscript received April 29, 1999.
*Formerly Univ. of Utah, Dept. of Geology and Geophysics, Salt Lake City, Utah 84112; currently Chevron Petroleum Tech. Co., 935 Gravier Street, New Orleans, Louisiana 70112. E-mail: nemt@chevron.com.
‡Univ. of Utah, Dept. of Geology and Geophysics, Salt Lake City, Utah 84112. E-mail: hsun@mines.utah.edu; schuster@mines.utah.edu.

the probability that the signal and coherent noise patterns coincide.

These methods assume that a properly chosen operator transforms the signal and coherent noise components into nonoverlapping domains. However, such an operator may not exist for the given signal and coherent noise components. For example, the apexes of primary and multiple reflections can be described equally well with both primary and multiple reflection modeling operators. In the case of aliased surface waves or tube waves, the apparent dip of the noise can be the same as the actual dip of reflected waves. In such cases some of the coherent noise energy and the signal energy are intermingled in the transform domain (Marfurt et al., 1996), i.e., there is crosstalk between the signal and the coherent noise. Claerbout (1990, 1992) studies the crosstalk phenomenon and finds that properly chosen linear and non-linear weighting coefficients might be necessary to mitigate the signal and coherent noise separation.

In this paper, we describe a new separation algorithm called migration filtering, which separates data arrivals according to their path of propagation. The algorithm accounts for the temporal moveout and the amplitudes of the actual data arrivals by using the respective forward modeling operators, so it is not restricted to uniformly weighted hyperbolic or linear moveouts. In the algorithm the data are modeled by the two modeling operators in the least-squares sense, and model estimates for the signal and coherent noise components are obtained. Then, the estimate of the signal model component is used to predict the signal component in the observed data. Using modeling operators allows one to describe each component with the minimum number of parameters; therefore, it increases the resolution and separability of the two components.

BASIC EQUATIONS

In this section the basic equations of the migration filtering algorithm are derived. The observed data $\mathbf{d}'$ are the sum of signal ($\mathbf{d}'_s$) and coherent noise ($\mathbf{d}'_n$) components

$$\mathbf{d}' = \mathbf{d}'_s + \mathbf{d}'_n. \tag{1}$$

The observed data vector $\mathbf{d}'$ on the left-hand side of equation (1) is a known quantity, and the signal and coherent noise components ($\mathbf{d}'_s$ and $\mathbf{d}'_n$) on the right-hand side are unknowns. Consequently, any filtering scheme based on equation (1) is underdetermined. Similarly, many transformations of equation (1) yield an underdetermined system in the transform domain. A typical feature of underdetermined schemes is that many nonzero transform-domain model parameters can represent a data parameter because of the inadequate constraints intrinsic to underdetermined problems. As a result, the model resolution in the transform-domain may be suboptimal. Therefore, it is essential to find a transform where the signal and the coherent noise transform-domain model estimates are well separated from each other.

In conventional filtering methods the observed data are usually transformed into a transform domain and the signal is extracted by using a passband filter of the chosen transformation. The coherent noise is given by the reject band of the transformation, and usually the pass and reject bands do not overlap. But a convenient transformation with the necessary resolution for data separation may not exist. To increase resolution and separability, we use separate operators for each desired wave mode, such that the operators provide maximum resolution for the given wave mode. The transformation is given as a combination of these operators. Here, the operators are chosen to be the forward modeling operators for the specified wave modes. These operators can account for the actual moveout and amplitudes of the data arrivals; thus they need fewer parameters to represent the data in the transform domain, thereby enhancing resolution, as illustrated in Figure 1.

In Figure 1 two events, the signal **a** and the coherent noise **b** are modeled using uniform amplitudes and hyperbolic moveout (signal), and nonuniform amplitudes and nonhyperbolic moveout (coherent noise). The two events are denoted by gray areas, with the lower end of the areas indicating the moveout and the width of the areas indicating the amplitudes. Figure 1a depicts fitting these events with uniform amplitude hyperbolic moveout curves, denoted by a solid line for the moveout and arrows for the amplitudes along the moveout curve. Figure 1b shows the corresponding transform domain images for the transformed signal **A** (black dot) and coherent noise **B** (black dot and cross-hatched area). Note that **B** is not focused since many fitting curves are necessary to model the event and some of the coherent noise energy is leaking into the signal passband (white area). Figure 1c depicts fitting these events with curves obtained by forward modeling, and Figure 1d shows the

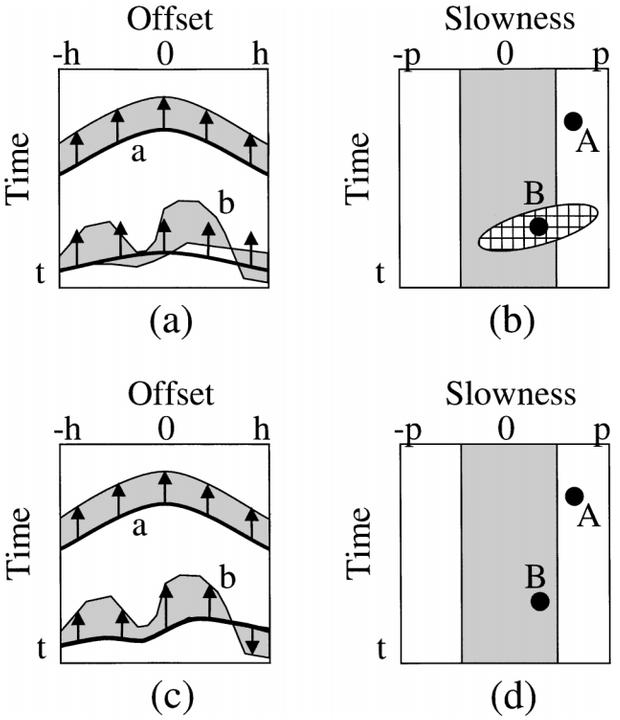

FIG. 1. Illustration of the enhanced resolution capabilities of migration filtering. (a) Signal **a** and coherent noise **b** (gray areas) are fitted with uniform amplitude (arrows) hyperbolic fitting curves (solid lines). (b) Transform domain images of signal (**A**) and coherent noise (**B**). The gray area denotes the coherent noise reject band. The white area is the signal passband. (c) Signal and coherent noise are fitted with curves obtained by forward modeling. (d) Transform domain images of signal and coherent noise from (c).

corresponding transform domain images. Here, the coherent noise image is focused and there is no leakage into the passband.

In detail, assume that the forward modeling operator $\mathbf{L}$ is represented by a combination of signal and coherent noise operators ($\mathbf{L}_s\ \mathbf{L}_n$), where

$$\mathbf{d}' = \mathbf{L}\,\mathbf{m}' = (\mathbf{L}_s \quad \mathbf{L}_n)\begin{pmatrix}\mathbf{m}'_s\\ \mathbf{m}'_n\end{pmatrix} = \mathbf{L}_s\mathbf{m}'_s + \mathbf{L}_n\mathbf{m}'_n, \qquad (2)$$

where $\mathbf{m}'$ is the model vector consisting of two parts: the model $\mathbf{m}'_s$ describing the signal $\mathbf{d}'_s$ and the model $\mathbf{m}'_n$ describing the coherent noise $\mathbf{d}'_n$. The forward modeling operators $\mathbf{L}_s$, $\mathbf{L}_n$ and models $\mathbf{m}'_s$, $\mathbf{m}'_n$ are chosen such that equation (2) is overdetermined and the model parameters are resolvable. To solve for $\mathbf{m}'$, the following parametric functional is formed:

$$P(\mathbf{m}) = \|\mathbf{L}\mathbf{m} - \mathbf{d}'\|^2. \qquad (3)$$

The model estimate $\mathbf{m}$ that minimizes equation (3) is

$$\mathbf{m} = (\mathbf{L}^T\mathbf{L})^{-1}\mathbf{L}^T\mathbf{d}'. \qquad (4)$$

To test the validity of using forward modeling operator $\mathbf{L}$, a more detailed analysis on the resolution and the invertibility of the least-squares migration operator is given in Appendices A and B.

The two key steps in our model-based filtering scheme are (a) to estimate $\mathbf{m}_s$ from equation (4) and (b) to find the signal estimate by computing $\mathbf{d}_s = \mathbf{L}_s\,\mathbf{m}_s$. The practical implementation of equation (4) is achieved by using a conjugate gradient scheme described in Nemeth (1996).

AN APPLICATION TO SYNTHETIC DATA

In this section some numerical tests are conducted to demonstrate the separation capabilities of the migration filtering method. The method is tested on synthetic data computed for the multiple P- and S-wave scatterer models depicted in Figure 2. Some point scatterers generate diffractions with both compressional- ($v_p = 3000$ m/s) and shear-wave velocities ($v_s = 1750$ m/s), and some generate diffractions with only compressional- or shear-wave velocities. The modeling operators and their values are given in Tables 1 and 2. The objective is to separate the compressional velocity scatterers (here, signal) from the shear velocity scatterers (here, coherent noise) in the common-shot gather. Both the correct and incorrect ($\pm 7\%$) velocity distributions are used for migration filtering. The incorrect velocity distribution is used to evaluate the robustness of migration filtering in the presence of velocity errors.

The compressional and the shear velocity scatterer models are shown in Figures 2a and 2b, and generate diffractions shown in Figures 2c and 2d, respectively. The observed common-shot gather is depicted in Figure 2e, where the shot location is at 0 m offset. Only this single gather was used for the subsequent

Table 1. Notations for Tables 2–5.

$d_s(.)$ = predicted signal
$m_s(.)$ = signal model estimate
$d_n(.)$ = predicted coherent noise
$m_n(.)$ = coherent noise model estimate
$W(t)$ = time history of the source
τ_{PR} = traveltime between image point and receiver
τ_{SP} = traveltime between source and image point
τ_{SR} = traveltime between source and receiver
$1/A_{PR}$ = amplitude decay between image point and receiver
$1/A_{SP}$ = amplitude decay between source and image point
$1/A_{SR}$ = amplitude decay between source and receiver
x_r, z_r, or $\vec{x}_r$ = spatial coordinates of receiver
x_s, z_s, or $\vec{x}_s$ = spatial coordinates of source
$\delta()$ = Dirac delta function

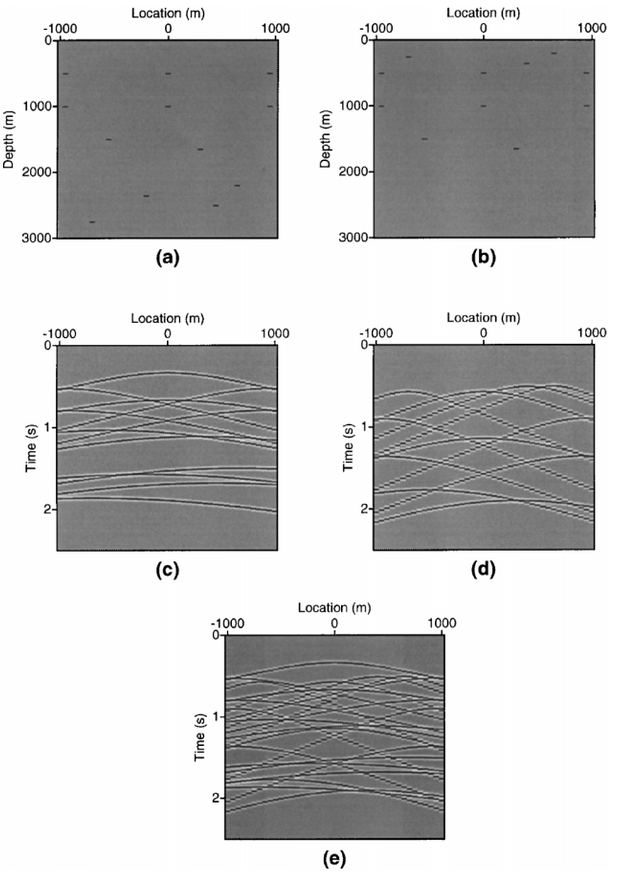

FIG. 2. Test with multiple scatterers. (a) Point scatterers for compressional waves. (b) Point scatterers for shear waves. (c) Single shot gather obtained from (a). (d) Single shot gather obtained from (b). (e) Observed shot gather obtained by stacking (c) and (d). Conventional filtering methods have difficulty in completely separating the compressional and shear components from a single gather.

Table 2. Signal and coherent noise modeling operators for separating *P*- and *S*-diffractions in the synthetic example.

	Explicit modeling formulas	Velocity restrictions
$\mathbf{L}_s$	$d_s(x_r, z_r, x_s, z_s, t) = \iint m_s(x,z)\ \frac{\delta(t-\tau_{PR}-\tau_{SP})}{A_{SP}A_{PR}} * W(t)\,dx\,dz$	$v = 3000\ \frac{m}{s}$
$\mathbf{L}_n$	$d_n(x_r, z_r, x_s, z_s, t) = \iint m_n(x,z)\ \frac{\delta(t-\tau_{PR}-\tau_{SP})}{A_{SP}A_{PR}} * W(t)\,dx\,dz$	$v = 1750\ \frac{m}{s}$

tests. The receiver interval is 50 m, and the maximum offset is 1000 m. Figure 3 depicts the results of migration filtering where the correct $\mathbf{L}_s$ and $\mathbf{L}_n$ operators are used. Figures 3a and 3b show the estimated scatterers, Figures 3c and 3d depict the estimated compressional and shear components, and Figure 3e shows the estimated data. Note that the estimated compressional and shear components approximately resemble their original components. The artifacts seen in the estimated gathers are mostly from the tangency problem, that is, when the apparent moveout curves of the compressional and shear scattered energy are locally collinear.

Figure 4 depicts the results of migration filtering, where the incorrect operators $\mathbf{L}_s$ and $\mathbf{L}_n$ are applied. Figures 4a and 4b show the estimated image of the scatterers, Figures 4c and 4d depict the estimated compressional and shear components, and Figure 4e shows the estimated data. Note that the incorrect velocity distribution introduces additional artifacts in the estimated data components. This is because the tangency problem is now different, since different velocities give rise to different apparent moveout curves. In our experience, the migration filtering method appears to be effective if the migration velocity errors are not more than 10% of the correct velocity. Outside this range, the apparent moveout of the predicted arrivals might differ too much from the actual arrivals.

These tests demonstrate that the accurate specification of operators $\mathbf{L}_s$ and $\mathbf{L}_n$ allows a satisfactory separation of signal and coherent noise even from a single gather. The information about the spatially variable moveouts and amplitudes for each component were captured the most compact way by the modeling operators. Therefore, it allowed the most optimal resolution of the two components. The tests also show the limitation of the method in the present form illustrated in the tangency problem. This limitation arises when the spaces of operators $\mathbf{L}_s$ and $\mathbf{L}_n$ overlap; it is studied in Appendix B.

TUBE-WAVE ATTENUATION IN CROSSWELL FIELD DATA

The migration filtering method is applied to a crosswell common-shot gather to attenuate tube waves in the data. The common-shot gather contains both reflected waves and tube waves, and it is f-k filtered in the range from 0.011 to 0.03 ms/ft (Figure 5a). The tube waves are aliased and their apparent dip coincides with the true dip of the reflections. For the reflected waves the modeling operator is the acoustic modeling operator

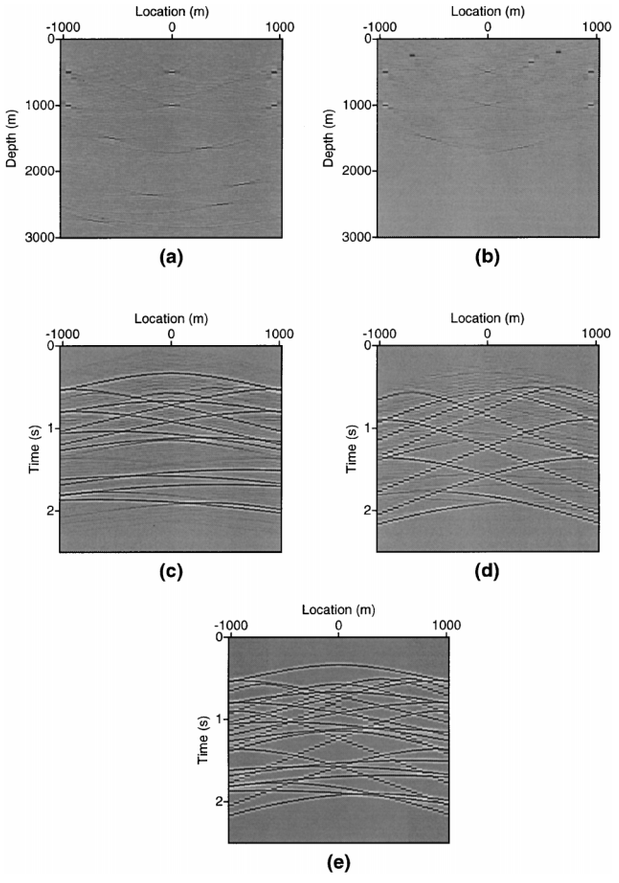

FIG. 3. Test with multiple scatterers. (a) Migrated section estimate of the signal model. (b) Migrated section estimate of the coherent noise model. (c) Single shot gather predicted from the estimated signal model. (d) Single shot gather predicted from the estimated coherent noise model. (e) Predicted shot gather using both the estimated signal and coherent noise models. The migration velocity is the correct one.

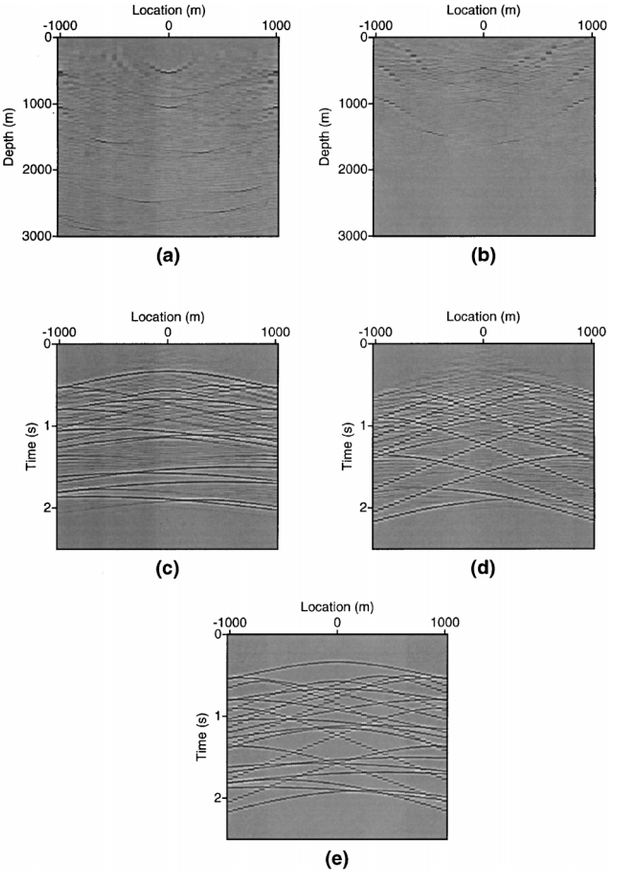

FIG. 4. The same as in Figure 3 except the compressional velocity distribution is 7% higher and the shear velocity distribution is 7% lower than the correct one. Acceptable separation has been achieved in the presence of realistic errors in migration velocity.

that employs a high-frequency approximation (see Table 3). The velocity distribution used for the forward modeling is obtained by traveltime tomography. The tube waves are modeled by using a modeling operator with linear moveout characteristics, as depicted in Table 3.

To assess the difficulty of removing these tube waves, we first examine how well the modeling operators can predict the data. The common-shot gather is f-k filtered, and the remaining filtered data (reflections and tube waves) should be predictable by the reflection modeling operator because the tube waves have the same apparent moveout as the real reflections. The tube-wave modeling operator is more restricted. It can predict most of the tube waves (except some local amplitude variations), but it cannot predict most of the reflections because they are not characterized by periodic arrivals with a linear moveout and a tube-wave velocity. Based on these considerations, it is expected that there will be some residual coherent noise energy left in the predicted signal component.

Figure 5b shows the reconstructed common-shot gather, and Figure 5c shows the reconstructed reflections after 15 iterations. The normalized objective function is 18% of the original objective function, where most of the misfit energy originates from the lower-left corner of the data in Figure 5a. This part of the data is not accessible because of restrictions on the dimension of the computational model. Some of the tube waves are removed, but considerable tube-wave energy still remains in the data after migration filtering. Apparently the reflection modeling operator is capable of predicting some of the tube waves because they are aliased in the same apparent dip direction as the reflections. Figure 5d shows the reconstructed tube waves after migration filtering.

Figures 6a–6d show the f-x spectrum of the observed data, the reconstructed data, the estimated reflections, and the estimated tube waves. The energy of the tube waves is partially removed from the estimated reflections, and the reflection

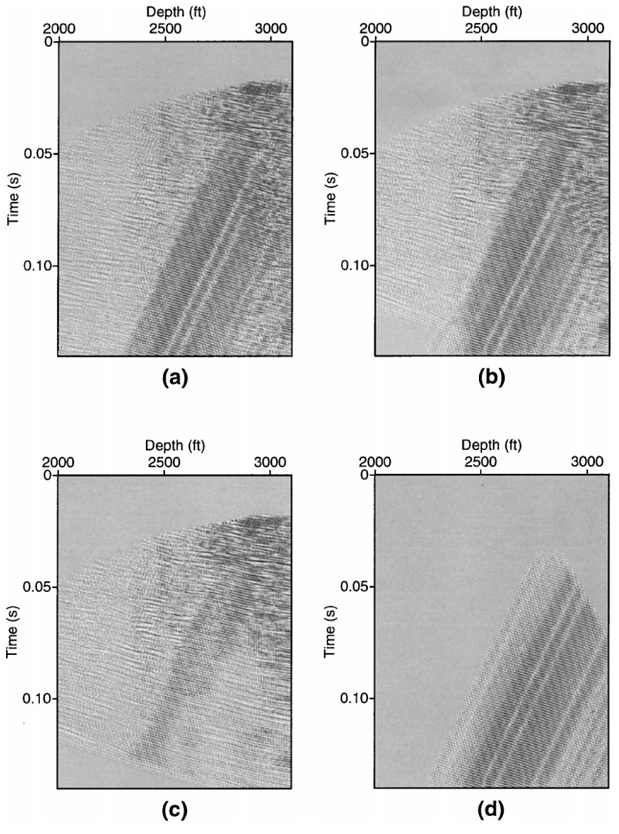

FIG. 5. Separation of crosswell reflections and tube waves. (a) A filtered crosswell common-shot gather. Note the aliased tube waves characterized by the periodic linear events that slant from left to right. The apparent dip of the aliased tube waves corresponds to the dip of the reflected waves. (b) The reconstructed common-shot gather; (c) the reconstructed reflections; (d) the reconstructed tube waves.

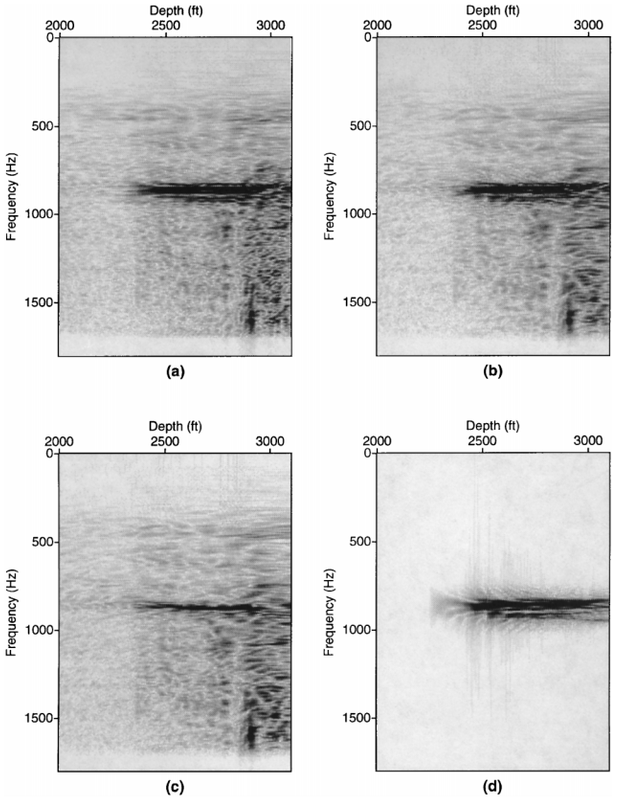

FIG. 6. Separation of crosswell reflections and tube waves. (a) The f-x spectrum of the crosswell common-shot gather; (b) the f-x spectrum of the reconstructed common-shot gather; (c) the f-x spectrum of the reconstructed reflections; (d) the f-x spectrum of the reconstructed tube waves.

Table 3. Signal and coherent noise modeling operators for separating reflections and tube waves in crosswell data.

	Explicit modeling formulas	Velocity restrictions
$\mathbf{L}_s$	$d_s(x_r, z_r, x_s, z_s, t) = \iint m_s(\vec{x}) \, \frac{\delta(t-\tau_{PR}-\tau_{SP})}{A_{SP}A_{PR}} * W(t)\, dx\, dz$	$v = v_{\text{migration}}$
$\mathbf{L}_n$	$d_n(z_r, z_s, t) = \int m_n(z)\, \delta\,(t - \tau_{SZ} - \tau_{ZS}) * W(t)\, dz$	$v = 4570\ \frac{ft}{s}$

patterns are now revealed. However, the estimated reflections still contain considerable tube-wave energy.

ATTENUATION OF SURFACE-RELATED WAVES IN 3-D FIELD DATA

Migration filtering is now used to attenuate surface-related waves from a 3-D common-shotpoint (CSP) gather collected in West Texas. The traces contain two dominant wave components: reflections and surface-related waves. The surface-related waves are partly comprised of Rayleigh waves, direct waves, and shallow multiple reflections. The low-frequency Rayleigh waves are first removed by a low-cut (0–20 Hz) filter to give the results shown in Figure 7. Although the surface-related waves do not pose a serious threat for imaging 2-D or 3-D in-line data, they become increasingly challenging for 3-D off-line data because these waves have an x–t moveout character that is hyperbolic on the far-offset off-lines.

The signal modeling operator (for the reflections) is selected to be the high-frequency acoustic modeling operator. The velocity for this operator is the migration velocity, and the migrated reflectivity events are obtained by the adjoint modeling operator (Kirchhoff migration). The coherent noise modeling operator for the surface-related waves is chosen to be a point scatterer modeling operator with a range of velocities from 7000 to 11 000 ft/s (2134–3354 m/s). The scatterers are restricted to be no deeper than 100 ft (26.3 m) from the surface. A summary of the modeling operators is given in Table 4.

A difficulty in separating wave modes is that the diffraction-based reflection modeling operator can predict most of the arrivals for each observed shot gather. Therefore we need to constrain the range of this operator. For this purpose we use geological a priori information by limiting the dip of the imaged reflections within a range of $\pm 15^\circ$ from the horizontal because the expected maximum dip of the subsurface layers does not exceed this range. In this way the constrained modeling operator cannot predict most of the surface-related waves. The surface-wave modeling operator predicts events with off-line hyperbolic moveouts in the given velocity range. Most of the surface-related waves satisfy this criterion, except for arrivals strongly affected by local surface velocity variations or amplitude fluctuations. The surface-wave modeling and reflection modeling operators are not likely to predict these fluctuations. As a result, we expect these features will remain in the residual data after inversion.

Figure 8a depicts an off-line shot gather taken from traces 642–720 in the 3-D data set in Figure 7. Figures 8b–8d depict the estimated signal, coherent noise, and residual data components. The misfit function value is 9% of the original data.

The estimated reflections and surface-related waves differ from each other in several features. The frequency range of the two components is different. The frequency ranges of the

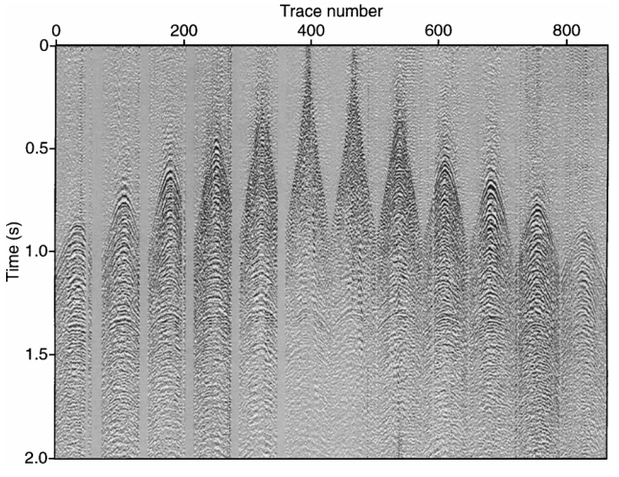

FIG. 7. The 3-D shot gather after 20-Hz low-cut filtering. The maximum in-line offset for the gathers is 7800 ft (2378 m); the maximum cross-line offset is 8600 ft. (2622 m)

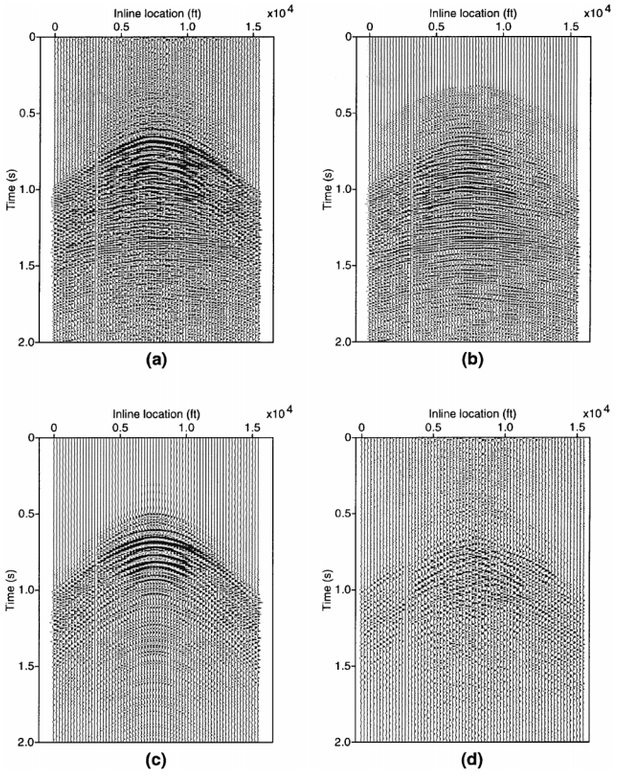

FIG. 8. Separation of reflected and surface-related waves in 3-D data. (a) An off-line common-shot gather extracted from the 3-D common-shot gather; (b) the reconstructed reflections; (c) the reconstructed surface-related waves; (d) the residual data between the observed and the reconstructed wave modes.

Table 4. Signal and coherent noise modeling operators for separating reflections and surface-related waves in 3-D data.

	Explicit modeling formulas	Velocity restrictions
$\mathbf{L}_s$	$d_s(\vec{x}_r, \vec{x}_s, t) = \iiint m_s(\vec{x}) \, \frac{\delta(t-\tau_{PR}-\tau_{SP})}{A_{SP}A_{PR}} * W(t) \, d\vec{x}$	$v = v_{\text{migration}}$
$\mathbf{L}_n$	$d_n(x_r, x_s, 0, t) = \iint m_n(x_s, y, v) \, \frac{\delta(t-\tau_{SR})}{A_{SR}} * W(t) \, dy \, dv$	$7000 \, \frac{ft}{s} < v < 11\,000 \, \frac{ft}{s}$

reflections and louder surface-related waves are 30–82 Hz and 2–40 Hz, respectively. The frequencies for arrivals near the apexes (where the x–t moveouts are similar) are different for the two components, with a small overlap in the 30–40-Hz range. Near the flanks of the hyperbolic moveout curves, however, both components have similar frequencies but their dip is different. The frequency range of the residual data is 20–65 Hz, but the main events in the center of the gather have the same frequency as the surface-related waves. The dip of the residual data also corresponds to the dip typical for the surface-related waves. Considering all these observations, we conclude that the estimated reflections and surface-related waves have different characteristics. Thus, migration filtering is successful in separating the surface-related waves and the reflection arrivals. The characteristics for the residual data are similar to those of the surface-related waves and dissimilar to those of the reflections.

SURFACE-WAVE ATTENUATION IN 2-D FIELD DATA

The migration filtering method is now applied to 2-D seismic field data to attenuate surface waves. The data were collected by Arco at their test site several miles west of Sulphur Springs, Texas. Figure 9a shows a portion of the data, which will be used as the input data for the migration filtering algorithm. Note the strong surface waves, which we will attempt to eliminate by migration filtering and f-k filtering. The source is located at the surface with an offset of 0 ft, there are 60 geophones located at the surface with offsets ranging from 1800 to 3570 ft (549–1088 m), the geophone interval is 30 ft (9.14 m), and the time sampling interval is 4 ms. No group arrays were used for these data.

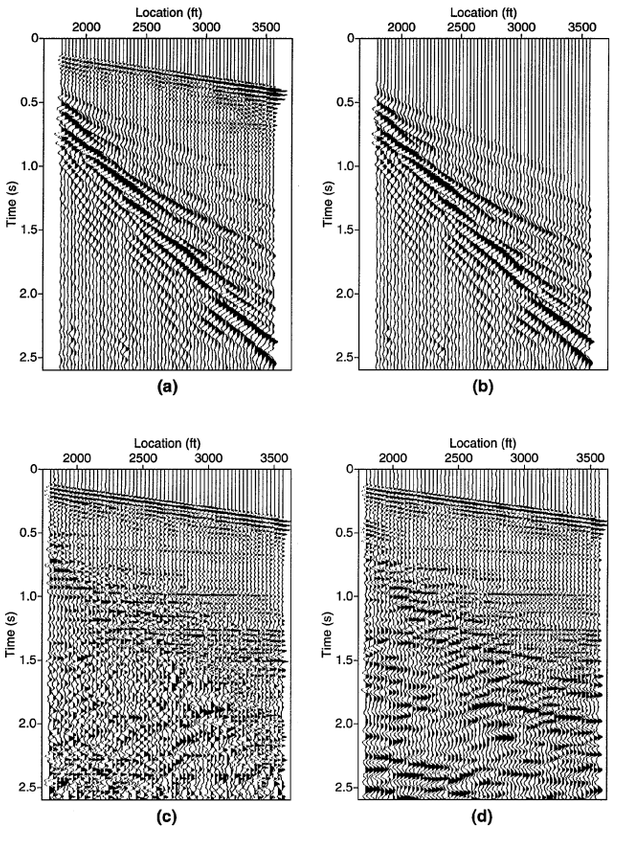

FIG. 9. Separation of reflected and surface waves in 2-D data. (a) A common-shot gather after 1–55-Hz bandpass filtering and muting; (b) the reconstructed surface waves; (c) the reconstructed reflections [1 s automatic gain control (AGC) is used for display]; (d) the predicted reflection by using f-k fan filtering (with 1 s AGC).

Figure 10a shows the f-k Fourier spectrum of the original traces. In the spectrum, the energy is focused into two frequency zones: 4 to 21 Hz, which is predominantly the surface-wave energy, and 22 to 40 Hz, which contains the reflection energy. We expect that migration filtering will effectively remove surface waves without attenuating the reflection energy.

The migration velocity is taken as a constant value of 8000 ft/s (2439 m/s). This use of a constant migration velocity also has important implications for computational efficiency because Kirchhoff's method of migration can be replaced by the much more efficient Stolt-type f-k migration. By inspection of the common-shot gathers, the surface-wave migration velocity is estimated to range from 850 to 2750 ft/s (259–838 m/s). The source wavelet is extracted directly from the seismograms. The central frequency of the Ricker wavelet is taken to be 10 Hz for the surface-wave migration and 30 Hz for the reflection migration operator. A summary of the modeling operators is given in Table 5.

Figure 9b shows the predicted surface waves after 50 iterations, which agree very well with the apparent surface waves in the original seismograms shown in Figure 9a. Figure 10b shows

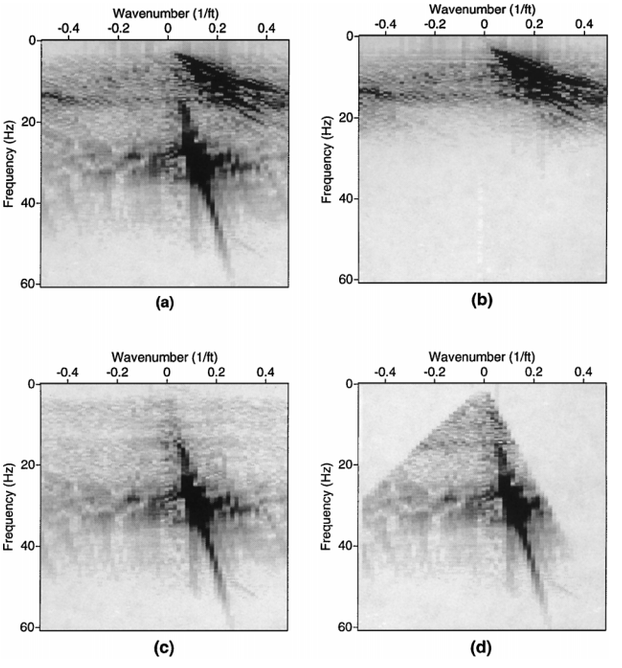

FIG. 10. Separation of reflected and surface waves in 2-D data. (a) The f-k spectrum of a common-shot gather shown in Figure 9; (b) the f-k spectrum of the reconstructed surface-related waves; (c) the f-k spectrum of the reconstructed reflections; (d) the f-k spectrum of the fan-filtered observed data.

Table 5. Signal and coherent noise modeling operators for separating reflections and surface waves in 2-D data.

	Explicit modeling formulas	Velocity restrictions
$\mathbf{L}_s$	$d_s(x_r, z_r, x_s, z_s, t) = \iint m_s(x, z)\, \frac{\delta(t-\tau_{PR}-\tau_{SP})}{A_{SP}A_{PR}} * W(t)\, dx\, dz$	$v = v_{\text{migration}}$
$\mathbf{L}_n$	$d_n(x_r, x_s, t) = \iint m_n(x_s, y, v)\, \frac{\delta(t-\tau_{SR})}{A_{SR}} * W(t)\, dy\, dv$	$850\ \frac{ft}{s} < v < 2750\ \frac{ft}{s}$

the f-k Fourier spectrum of the predicted surface waves, which also agrees well with the original surface-wave energy shown in Figure 10a.

Figure 9c shows the migration-filtered common-shot gathers after 50 iterations. Compared with the original data in Figure 9a, this migration-filtered section contains less surface-wave energy, indicating an effective removal of surface waves. Figure 10c shows the f-k Fourier spectrum of the migration-filtered common-shot gather. Compared with the original data spectrum in Figure 10a, most of the surface-wave energy has been filtered out, but there is no significant attenuation of reflected wave energy. In comparison, Figure 9d shows f-k filtered common-shot gathers, and Figure 10d depicts their f-k spectrum. Migration filtering succeeds to correctly predict both components' f-k spectrum without the surgical cut required by f-k filtering.

DISCUSSION AND CONCLUSIONS

The key element in separating the signal and coherent noise is to find operators that separate these components in some transform domain. In this paper we described an algorithm that separates data arrivals according to their actual path of propagation. The modeling operators are not restricted to linear or hyperbolic moveout and uniform amplitudes. The benefits of using the modeling operators are (1) better wavefield separation capability as a result of increased resolution (also fewer artifacts) in the respective transform-domain model estimates, (2) capability of mixing any two conventional transforms for wavefield separation under a general inversion framework, and (3) capability of mitigating the crosstalk problem by adding model- and data-dependent regularization terms to the algorithm (Nemeth, 1996). Results with synthetic and field data suggest that migration filtering provides better separation than the separation obtained by using some other conventional transform because the moveout and amplitude information of the data is captured in the most compact way.

The modeling operators might not be able to separate the two components completely. For example, separation of tangent data components will be incomplete. The synthetic and field data examples show different separation results depending on the predictability of one component with the other component's operator. The 2-D surface wave and the 3-D surface-related wave separations resulted in an almost complete separation of coherent energy from signal, while some residual coherent energy was left in the signal for the tube-wave case and the synthetic example. To further improve these last two cases, some regularization may be necessary. Model-based regularization can mitigate the crosstalk by using constraints on the model domain parameters, and data component-based regularization provides additional information by using data constraints (Nemeth, 1996). Further research on the regularization issues is also necessary.

The effectiveness of the migration filtering method relies on the accurate specification of the two modeling operators, $\mathbf{L}_s$ and $\mathbf{L}_n$. In practice, only approximate estimates of these modeling operators are available, and they may have (slightly) differing ranges than those of the original operators. For example, the velocity distributions used for modeling might be suboptimal, leading to some artifacts in separation. However, the numerical examples suggest that migration filtering can tolerate moderate velocity errors ($\pm 7\%$). Random noise and additional unspecified wave modes also affect the separation capabilities of the migration filtering method. If these wave modes lie beyond the spaces spanned by the modeling operators, the filtering algorithm will leave them in the residual vector.

The present form of migration filtering is cast into a least-squares inversion formalism where numerous (in practice, a few tens of) iterations may be necessary to achieve acceptable convergence. This is usually the consequence of using inefficient modeling operators $\mathbf{L}_s$ and $\mathbf{L}_n$ for the inversion. To speed up convergence, both $\mathbf{L}_s$ and $\mathbf{L}_n$ can be preconditioned to balance the convergence rates of $\mathbf{L}_s$ and $\mathbf{L}_n$. Of course, a general computational speed-up can be achieved by providing computationally faster modeling operators.

Possible extensions of the migration filtering method include (1) separation of primaries and multiples, (2) separation of reflections and surface waves, (3) separation of S-waves and converted P-S-waves, (4) separation of reflections and aliased tube waves, (5) separation of on-line and off-line reflections in a 3-D survey, and (6) separation of multicomponent data (Wang and Nemeth, 1997).

ACKNOWLEDGMENTS

We are grateful for the financial support from the members of the 1996–1997 University of Utah Tomography and Modeling/Migration (UTAM) Consortium. The comments and suggestions of associate editor Mihai Popovici and two anonymous reviewers are gratefully acknowledged.

REFERENCES

Abma, R., 1995, Least-squares separation of signals and noise using multidimensional filters: Ph.D. Dissertation, Stanford Univ.

Biondi, B., 1992, Velocity estimation by beam stack: Geophysics, **57**, 1034–1047.

Blonk, B., Herman, G., and Drijkoningen, G., 1995, An elastodynamic inverse scattering method for removing scattered surface waves from field data: Geophysics, **60**, 1897–1905.

Claerbout, J. F., 1990, Eliminating crosstalk: The Leading Edge, **9**, 38–40.

———1992, Earth soundings analysis: Processing versus inversion: Blackwell Scientific Publications, Inc.

Duquet, B., and Marfurt, K., 1997, Filtering coherent noise during prestack depth migration: 67th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1262–1265.

Harlan, W., Claerbout, J., and Rocca, F., 1984, Signal/noise separation and velocity estimation: Geophysics, **49**, 1869–1880.

Hu, T., and White, R., 1998, Robust multiple suppression using adaptive beamforming: Geophys. Prosp., **46**, 227–248.

Kostov, C., and Biondi, B., 1991, Improved resolution of slant stacks using beam stack, *in* Gardner, G., and Lu, L., eds., Slant-stack processing: Soc. Expl. Geophys., 358–360.
Lumley, D., Nichols, D., and Rekdal, T., 1995, Amplitude-preserved multiple suppression: 65th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1460–1463.
Marfurt, K., Schneider, R., and Mueller, M., 1996, Pitfalls in using conventional and discrete Radon-transforms on poorly sampled data: Geophysics, **61**, 1467–1482.
Nemeth, T., 1996, Imaging and filtering by least-squares migration: Ph.D. Dissertation, Univ. of Utah.
Spitz, S., 1998, Pattern recognition and subtraction of coherent noise: 60th Ann. Conf., Eur. Assoc. Geosci. Eng., Extended Abstracts, paper 1-09.
Tarantola, A., 1987, Inverse problem theory: Methods for data fitting and model parameter estimation: Elsevier Science Publ. Co., Inc.
Thorson, J., and Claerbout, J. F., 1985, Velocity-stack and slant-stack stochastic inversion: Geophysics, **51**, 2727–2741.
Wang, Y., and Nemeth, T., 1997, Multi-component separation of PP and SS by a least-squares migration method: Synthetic and field data tests: 67th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts, 1222–1225.

APPENDIX A

MATRIX EXPRESSIONS FOR THE LEAST-SQUARES MIGRATION OPERATOR

In this section the explicit matrix expressions are derived for the Hessian matrix associated with equation (4), the least-squares migration operator, and the data resolution matrix.

The Hessian matrix is given as

$$\mathbf{H} = \begin{pmatrix} \mathbf{H}_{11} & \mathbf{H}_{12} \\ \mathbf{H}_{21} & \mathbf{H}_{22} \end{pmatrix} = \mathbf{L}^T\mathbf{L} = \begin{pmatrix} \mathbf{L}_s^T\mathbf{L}_s & \mathbf{L}_s^T\mathbf{L}_n \\ \mathbf{L}_n^T\mathbf{L}_s & \mathbf{L}_n^T\mathbf{L}_n \end{pmatrix}. \quad \text{(A-1)}$$

Its inverse, $\mathbf{D} = \mathbf{H}^{-1}$, is given as

$$\begin{aligned} \mathbf{D}_{11} &= \left(\mathbf{H}_{11} - \mathbf{H}_{12}\,\mathbf{H}_{22}^{-1}\,\mathbf{H}_{21}\right)^{-1}, \\ \mathbf{D}_{12} &= -\left(\mathbf{H}_{11} - \mathbf{H}_{12}\,\mathbf{H}_{22}^{-1}\,\mathbf{H}_{21}\right)^{-1}\mathbf{H}_{12}\,\mathbf{H}_{22}^{-1}, \\ \mathbf{D}_{21} &= -\left(\mathbf{H}_{22} - \mathbf{H}_{21}\,\mathbf{H}_{11}^{-1}\,\mathbf{H}_{12}\right)^{-1}\mathbf{H}_{21}\,\mathbf{H}_{11}^{-1}, \\ \mathbf{D}_{22} &= \left(\mathbf{H}_{22} - \mathbf{H}_{21}\,\mathbf{H}_{11}^{-1}\,\mathbf{H}_{12}\right)^{-1}. \end{aligned} \quad \text{(A-2)}$$

Similar expressions can be found in Tarantola (1987, problems 4.2, 4.3, 4.4). Substituting the terms of the Hessian matrix into the above expressions yields the inverse of the Hessian:

$$\mathbf{H}^{-1} = \begin{pmatrix} \left[\mathbf{L}_s^T(\mathbf{I}-\mathbf{N}_n)\mathbf{L}_s\right]^{-1} & -\left[\mathbf{L}_s^T(\mathbf{I}-\mathbf{N}_n)\mathbf{L}_s\right]^{-1}\mathbf{L}_s^T\mathbf{L}_n\left(\mathbf{L}_n^T\mathbf{L}_n\right)^{-1} \\ -\left[\mathbf{L}_n^T(\mathbf{I}-\mathbf{N}_s)\mathbf{L}_n\right]^{-1}\mathbf{L}_n^T\mathbf{L}_s\left(\mathbf{L}_s^T\mathbf{L}_s\right)^{-1} & \left[\mathbf{L}_n^T(\mathbf{I}-\mathbf{N}_s)\mathbf{L}_n\right]^{-1} \end{pmatrix}, \quad \text{(A-3)}$$

where

$$\begin{aligned} \mathbf{N}_s &= \mathbf{L}_s\left(\mathbf{L}_s^T\mathbf{L}_s\right)^{-1}\mathbf{L}_s^T, \\ \mathbf{N}_n &= \mathbf{L}_n\left(\mathbf{L}_n^T\mathbf{L}_n\right)^{-1}\mathbf{L}_n^T \end{aligned} \quad \text{(A-4)}$$

are the data resolution matrices of the signal and the coherent noise, respectively. Matrices $(\mathbf{I}-\mathbf{N}_s)$ and $(\mathbf{I}-\mathbf{N}_n)$ are called signal and coherent noise rejector matrices (with $\mathbf{N}_s$ and $\mathbf{N}_n$ being projector matrices), respectively, because they filter out the signal and the coherent noise components from the data. For example, if the signal filter matrix is applied to the data described by equation (2), the signal component is filtered out and part of the coherent noise component lying outside space defined by $\mathbf{I}-\mathbf{N}_s$ is passed ($(\mathbf{I}-\mathbf{N}_s)\,\mathbf{d}' = (\mathbf{I}-\mathbf{N}_s)\mathbf{L}_n\mathbf{m}_n'$).

The least-squares migration operator is given by

$$\mathbf{H}^{-1}\mathbf{L}^T = \begin{pmatrix} \left[\mathbf{L}_s^T(\mathbf{I}-\mathbf{N}_n)\mathbf{L}_s\right]^{-1}\mathbf{L}_s^T(\mathbf{I}-\mathbf{N}_n) \\ \left[\mathbf{L}_n^T(\mathbf{I}-\mathbf{N}_s)\mathbf{L}_n\right]^{-1}\mathbf{L}_n^T(\mathbf{I}-\mathbf{N}_s) \end{pmatrix}. \quad \text{(A-5)}$$

Equation (A-5) shows how the least-squares migration operator estimates the model from the data. First, the data $\mathbf{d}'$ are filtered by both the coherent noise rejector matrix $(\mathbf{I}-\mathbf{N}_n)$ and the signal rejector matrix $(\mathbf{I}-\mathbf{N}_s)$ to eliminate the coherent noise and signal components, respectively. These rejector matrices not only eliminate the above-mentioned components but may also reject part of the passed components lying within the spaces defined by the corresponding rejector matrices. Then, standard adjoint modeling operators are applied to these remaining components to give the model estimates, e.g., for the signal $\mathbf{m}_s \approx \mathbf{L}_s^T(\mathbf{I}-\mathbf{N}_n)\,\mathbf{d}'$. Finally, inverse Hessian-like operators, e.g., $[\mathbf{L}_s^T(\mathbf{I}-\mathbf{N}_n)\mathbf{L}_s]^{-1}$, are applied to these model estimates to provide a least-squares estimate of the model. If the inverse Hessian-like operators, as above, are invertible, the two components can be separated perfectly. In the opposite case some approximate inverses are used and parts of the signal and coherent noise lying in the same space spanned by both operators are only partially separated. To mitigate this situation, some regularization might be needed. For a more detailed analysis on the invertibility issue, see Appendix B.

Finally, the data resolution matrix is given as

$$\begin{aligned} \mathbf{L}\,\mathbf{H}^{-1}\mathbf{L}^T = {} & \mathbf{L}_s\left[\mathbf{L}_s^T(\mathbf{I}-\mathbf{N}_n)\mathbf{L}_s\right]^{-1}\mathbf{L}_s^T(\mathbf{I}-\mathbf{N}_n) \\ & + \mathbf{L}_n\left[\mathbf{L}_n^T(\mathbf{I}-\mathbf{N}_s)\mathbf{L}_n\right]^{-1}\mathbf{L}_n^T(\mathbf{I}-\mathbf{N}_s), \end{aligned} \quad \text{(A-6)}$$

where the first term on the right-hand side is the data resolution matrix associated with the signal and the second term is the data resolution matrix associated with coherent noise.

APPENDIX B

INVERTIBILITY OF THE HESSIAN OF THE LEAST-SQUARES MIGRATION OPERATOR

In this section the invertibility properties of the Hessian of the least-squares migration operator in equation (A-5) are analyzed. While equation (A-5) demonstrates that both the signal and the coherent noise components can be reconstructed theoretically from the data, it also shows the limitations of the reconstruction. Here, we will analyze only the upper row of expression (A-5) for its ability to reconstruct the signal estimate $\mathbf{d}_s$. A similar analysis can also be made for the case of coherent noise reconstruction with the lower row of expression (A-5).

We chose modeling operators such that $(\mathbf{L}_s^T\mathbf{L}_s)^{-1}$ and $(\mathbf{L}_n^T\mathbf{L}_n)^{-1}$ are invertible. Then the signal reconstruction is limited by the crosstalk of the coherent noise rejector matrix $(\mathbf{I}-\mathbf{N}_n)$. That is, matrix

$$\mathbf{Q}=\left[\mathbf{L}_s^T(\mathbf{I}-\mathbf{N}_n)\mathbf{L}_s\right] \tag{B-1}$$

in equation (A-5) may not be invertible. There are several cases where $\mathbf{Q}$ is not invertible because of the structure of the data resolution matrix $\mathbf{N}_n$.

$\mathbf{N}_n = \mathbf{I}$, or the range of $\mathbf{N}_n$ covers the full data range

In this case

$$\mathbf{N}_n=\mathbf{L}_n\left(\mathbf{L}_n^T\mathbf{L}_n\right)^{-1}\mathbf{L}_n^T=\mathbf{I}, \tag{B-2}$$

with $\mathbf{I}$ being the unity matrix so that the matrix $\mathbf{Q}=\mathbf{0}$ and its inverse do not exist. Here, matrix $\mathbf{0}$ denotes a matrix with zero values. In this case operator $\mathbf{N}_n$ predicts any possible data in the least-squares sense. Examples of invertible operators are the continuous Fourier transform and the continuous Radon transform over closed surfaces (related to the Fourier transform via the slice-projection theorem). Equation (B-2) also implies that operator $\mathbf{L}_n$ is invertible.

A necessary condition for the complete separation of signal and coherent noise is that the $\mathbf{L}_s$ and $\mathbf{L}_n$ operators must span different regions in the data space. For example, if $\mathbf{L}_s$ predicts all data in the 30–50-Hz range and $\mathbf{L}_n$ predicts all data in the 70–90-Hz range, then the signal and the coherent noise can be separated exactly. If one of the operators spans the full range, as in this case, complete separation is not possible without modifications. For illustration, see Figure B-1a.

$\mathbf{L}_n = \mathbf{L}_s$, or coincident operator ranges

In this case

$$\mathbf{N}_n=\mathbf{L}_s\left(\mathbf{L}_s^T\mathbf{L}_s\right)^{-1}\mathbf{L}_s^T \qquad \text{or} \qquad \mathbf{L}_n=\mathbf{L}_s, \tag{B-3}$$

so that the matrix $\mathbf{Q}=\mathbf{0}$ and its inverse do not exist. This case shows that if the two operators ($\mathbf{L}_s$ and $\mathbf{L}_n$) are too close to one another, they have the same range and the signal and coherent noise cannot be separated. For example, if both $\mathbf{L}_s$ and $\mathbf{L}_n$ predict the same data in the 30–50-Hz range, then the signal and coherent noise cannot be separated by equation (A-5). For illustration, see Figure B-1b.

Partially coincident operator ranges

This is the case when the ranges of $\mathbf{L}_s$ and $\mathbf{L}_n$ partially overlap, or $\mathbf{Q}\neq\mathbf{0}$ and its inverse still do not exist. In such cases some parts of the signal and the coherent noise are fully separable and some parts are not separable. For example, if $\mathbf{L}_s$ predicts data in the 30–50-Hz range and $\mathbf{L}_n$ predicts data in the 40–60-Hz range, then the 30–40-Hz range of data will be predicted as signal, the 50–60-Hz range of data will be predicted as coherent noise, and the 40–50-Hz range of data will be predicted by some combination of signal and coherent noise. For illustration, see Figure B-1c.

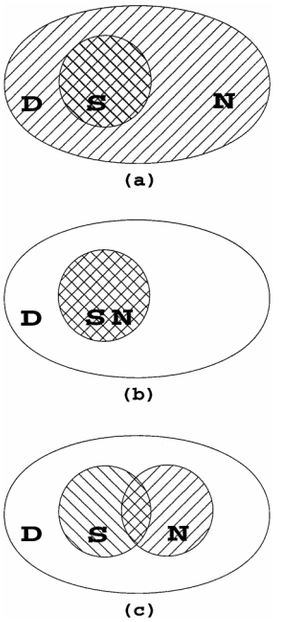

FIG. B-1. Overlapping operator ranges. **D** denotes the full data range, **S** denotes the range of the signal modeling operator, and **N** denotes the range of the coherent-noise modeling operator. (a) The range of the coherent-noise modeling operator covers the full data range; (b) the ranges of signal and coherent-noise modeling operators fully overlap; and (c) the ranges of signal and coherent-noise modeling operators partially overlap.