

High-resolution velocity spectra using eigenstructure methods

Biondo L. Biondi* and Clement Kostov*

ABSTRACT

Stacking spectra provide maximum-likelihood estimates for the stacking velocity, or for the ray parameter, of well separated reflections in additive white noise. However, the resolution of stacking spectra is limited by the aperture of the array and the frequency of the data. Despite these limitations, parametric spectral estimation methods achieve better resolution than does stacking. To improve resolution, the parametric methods introduce a parsimonious model for the spectrum of the data. In particular, when the data are modeled as the superposition of wavefronts, the properties of the eigenstructure of the data covariance matrix can be used to obtain high-resolution spectra. The traditional stacking spectra can also be expressed as a function of the data covariance matrix and directly compared to the eigenstructure spectra. The superiority of the latter in separating closely interfering reflections is then apparent from a simple geometric interpretation.

Eigenstructure methods were originally developed for use with narrow-band signals, while seismic reflections are wide-band and transient in time. Taking advantage of the full bandwidth of seismic data, we average spectra from several frequency bands. We choose each frequency band wide enough, so that we can average over time estimates of the covariance matrix. Thus, we obtain a robust estimate of the covariance matrix from short data sequences.

A field-data example shows that the high-resolution estimators are particularly attractive for use in the estimation of local spectra in which short arrays are considered. Several realistic synthetic examples of stacking-velocity spectra illustrate the improved performance of the new methods in comparison with conventional processing.

INTRODUCTION

The estimation of stacking velocities is a classic problem in exploration seismology. A related problem is the estimation of the ray parameter of a plane wave. Both problems are particular cases of a more general problem: estimating the parameters that describe the shape of the wavefront ("shape parameters") for a signal recorded at a linear array of receivers. In the near field, the wavefront is approximately spherical, and therefore the relative time delays between receivers are conveniently parameterized by stacking velocities and zero-offset traveltimes. In the far field, the wavefront is well approximated by plane waves characterized by ray parameters.

The standard approach to estimating stacking velocities is to pick the maxima of shape-parameter spectra (Taner and Koehler, 1969). The spectra are computed by the repeated application, for a sweep of shape parameters, of first a time correction that aligns the wavefront in space along the array and then a coherency measure along the spatial direction. The time correction could be normal moveout (NMO) or linear moveout (LMO); the coherency measure could be a simple stacking or the computation of a semblance function.

The stacking spectrum, and in particular the stacking-velocity spectrum, has many attractive properties. It yields the maximum-likelihood estimates of the shape parameter when the statistics of the data are Gaussian and there is only one wavefront impinging on the array. Furthermore, the estimate is robust with respect to the data's deviations from the assumed simple propagation model or from Gaussian statistics. When two or more wavefronts impinge on the array, the classical procedure still yields good estimates of their shape parameters, provided that the wavefront shapes are sufficiently different. However, when a pair of wavefronts are too closely interfering, the resulting estimates are biased. Even worse, the spectra may show only one maximum and indicate only one wavefront incident on the array.

The main disadvantage of the stacking spectrum is its poor resolution, limited by the aperture of the array and by the

Manuscript received by the Editor June 30, 1988; revised manuscript received December 27, 1988.

*Geophysics Department, Stanford University, Stanford, CA 94305-2215.

© 1989 Society of Exploration Geophysicists. All rights reserved.

frequency of the data. For narrow-band data, the estimation of shape-parameter spectra is related to the estimation of frequency spectra for time series, and the stacking spectrum is equivalent to the periodogram. Many algorithms for estimating high-resolution, shape-parameter spectra from narrow-band data have been developed, in particular for radar and sonar applications (Capon, 1969). To improve the resolution, a model for the data is introduced; this method is similar to the high-resolution methods for estimating frequency spectra (Pisarenko, 1972; Burg, 1975). When the data are modeled as a superposition of wavefronts, the covariance matrix of the data has a particular structure, and the properties of its eigenstructure can be exploited so that the resolution of the spectra is further improved (Bienvenu and Kopp, 1983; Schmidt, 1986). Eigenstructure methods can be extended to the case of broad-band signals, such as the seismic signal, by decomposing the data into narrow-band components (Wax et al., 1984; Wang and Kaveh, 1985).

In reflection seismology, high-resolution spectra are needed to resolve reflection events, for instance primaries from reflectors with conflicting dips or a primary and an intrabed or peg-leg multiple. Stacking-velocity spectra typically fail to separate interfering events at late times and high velocities. High-resolution spectra can be even more useful in the estimation of local spectra, such as local slant stacks (Sword, 1987) or beam stacks (Kostov and Biondi, 1987). The use of local spectra has two advantages over conventional methods: (1) the physical model describing the data is more accurate for shorter arrays (for instance the plane-wave approximation is appropriate only in the Fresnel zone) and (2) local spectra are more sensitive to small-scale variations in the shape parameter. On the other hand, local spectra have poor resolution, because their estimation uses shorter arrays. This limitation can be alleviated by use of high-resolution methods.

The application of the eigenstructure methods to seismic reflections is more complicated than their application to sonar or radar data, because seismic data are both wide-band and highly nonstationary, in the sense that the shape parameter of the reflections can vary rapidly with time. When processing seismic data, only a few time samples can be used for the estimation of the covariance matrix; therefore, the redundancy of information contained in all the frequency bands of the data must be exploited to obtain reliable shape-parameter spectra.

Previous applications in geophysics of the eigenstructure methods include those by Key et al. (1987) and Mars et al. (1987). Key et al. use only the eigenvalues of the covariance matrix and do not allow for interfering reflections in their data model. Mars et al. discuss the applications of eigenstructure methods to noise suppression.

In the first section, we present the eigenstructure algorithms applied to narrow-band signals. Then we compare the new methods with the classical stacking method and show a geometric explanation for the superiority of the former. The third section contains an extension of the narrow-band method to the general case of a wide-band signal. In this section, we discuss the particular features of the seismic problem and the tradeoff between resolution and robustness in the application of the method to seismic signals. The last two sections illustrate, with a field-data example and some

synthetic examples, the applications of the high-resolution methods to stacking-velocity analysis and local slant stacks.

THE NARROW-BAND METHOD

The recorded data are assumed to be complex and narrow-band, with central angular pulsation ω . The seismic signal is real and broad-band, but it can be easily transformed into a narrow-band, analytical signal by use of a band-pass filter and the Hilbert transform (Claerbout, 1976). Some advantages of working with the analytical signal, rather than with the real-valued one, are discussed by Sguazzero and Vesnaver (1987). The data recorded at the M receivers are described as the combination of W wavefronts and additive noise. The model for the data recorded by the receiver m at time t is

$$d(m, t) = \sum_{w=1}^W s_w [t - \tau_w(m; \theta_w)] + n(m, t), \quad (1)$$

where $n(m, t)$ is the additive noise, assumed to be uncorrelated with the sources $s_w(t)$. The delays of the wavefronts' arrivals at receiver m with respect to zero-offset traveltime τ_w are denoted by $\tau_w(m; \theta_w)$. When the wavefronts are plane waves, the delays are expressed as

$$\tau_w(m; \theta_w) = \theta_w \Delta x (m - 1), \quad (2)$$

where Δx is the spacing between the receivers and θ_w is the ray parameter of the w wavefront. When the wavefront shapes are hyperbolic, the delays are

$$\tau_w(m; \theta_w) = \sqrt{\bar{\tau}_w^2 + \theta_w^2 [\Delta x (m - 1)]^2} - \bar{\tau}_w, \quad (3)$$

where now θ_w is the stacking slowness. In this case, the delays are functions of two parameters: slowness and zero-offset traveltime. We consider the spectra only as functions of slowness, but the proposed method could be applied to estimate spectra that are functions of both slowness and zero-offset traveltime.

The sources $s_w(t)$ are modeled as narrow-band stochastic processes. Thus, the time shifts are complex exponentials; and the data can be expressed as

$$\begin{aligned} d(m, t) &= \sum_{w=1}^W s_w(t) e^{i\omega \tau_w(m; \theta_w)} + n(m, t) \\ &= \sum_{w=1}^W \sqrt{M} s_w(t) \frac{e^{i\omega \tau_w(m; \theta_w)}}{\sqrt{M}} + n(m, t). \end{aligned} \quad (4)$$

The recorded data are truncated in time by a window of T time samples containing the interfering reflections to be analyzed. In matrix notation the data window is expressed as

$$\mathbf{D} = \mathbf{A}(\boldsymbol{\Theta}) \mathbf{S} + \mathbf{N}, \quad (5)$$

where \mathbf{D} and \mathbf{N} are $(M \times T)$ data and noise matrices. \mathbf{S} is a $(W \times T)$ source matrix, and $\mathbf{A}(\boldsymbol{\Theta})$ is the "data steering matrix" function of the vector of parameters $\boldsymbol{\Theta}$. The data steering matrix is an $(M \times W)$ matrix formed by the W "data steering vectors" $\mathbf{a}(\theta_w)$:

$$\mathbf{a}(\theta_w) = \frac{1}{\sqrt{M}} \begin{bmatrix} e^{i\omega\tau_n(1;\theta_w)} \\ \vdots \\ e^{i\omega\tau_n(m;\theta_w)} \\ \vdots \\ e^{i\omega\tau_n(M;\theta_w)} \end{bmatrix} \quad (6)$$

Because the sources are assumed to be zero-mean stochastic processes and to be uncorrelated with the noise, the data covariance matrix is

$$\begin{aligned} \mathbf{R}_d &= E[\mathbf{D}\mathbf{D}^*] = \mathbf{A}(\Theta)E[\mathbf{S}\mathbf{S}^*]\mathbf{A}(\Theta)^* + \mathbf{R}_n \\ &= \mathbf{A}(\Theta)\mathbf{R}_s\mathbf{A}(\Theta)^* + \mathbf{R}_n, \end{aligned} \quad (7)$$

where \mathbf{R}_s and \mathbf{R}_n are the respective covariance matrices of the sources and the noise and \mathbf{D}^* is the conjugate transpose of \mathbf{D} .

For simplicity, we assume that the noise is spatially uncorrelated and has equal power at all the receivers. Equation (7) becomes

$$\mathbf{R}_d = \mathbf{A}(\Theta)\mathbf{R}_s\mathbf{A}(\Theta)^* + \sigma_n^2\mathbf{I}. \quad (8)$$

The above assumption about the statistics of the noise does not limit the generality of the method, because if \mathbf{R}_n were known, the data could either be prewhitened (Bienvenu and Kopp, 1983) or the eigenstructure of the positive definite matrix pencil (\mathbf{R}_d , \mathbf{R}_n) could be used (Schmidt, 1986).

Properties of the eigenstructure of the data covariance matrix

When the source covariance matrix \mathbf{R}_s is nonsingular (i.e., there are no pairs of fully correlated sources) and there are no linearly dependent steering vectors $\mathbf{a}_n(\theta_w)$, the matrices \mathbf{R}_s and $\mathbf{A}(\Theta)$ have rank W (here we assume that $M \geq W$); and therefore the matrix $\mathbf{A}(\Theta)\mathbf{R}_s\mathbf{A}(\Theta)^*$ also has rank W .

Let $[\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_M]$ be the eigenvalues and $[\mathbf{E}_1, \mathbf{E}_2, \dots, \mathbf{E}_M]$ be the eigenvectors of the data covariance matrix \mathbf{R}_d . Equation (8) implies that the eigenvalues of \mathbf{R}_d are of the form $\lambda = \mu^2 + \sigma_n^2$, with μ^2 the eigenvalues of $\mathbf{A}(\Theta)\mathbf{R}_s\mathbf{A}(\Theta)^*$. The rank of \mathbf{R}_s is W , so the two following properties hold true:

(1) The minimal eigenvalue is σ_n with multiplicity $M - W$. That is,

$$\lambda_{W+1} = \lambda_{W+2} = \dots = \lambda_M = \sigma_n^2. \quad (9a)$$

(2) The eigenvectors corresponding to the minimal eigenvalues are orthogonal to the data steering vectors $\mathbf{a}_n(\theta_w)$. That is,

$$\mathbf{a}(\theta_w)^*\mathbf{E}_m = 0, \quad m = W + 1, W + 2, \dots, M. \quad (9b)$$

The subspace $\mathbf{E}_{n_r} = [\mathbf{E}_{W+1}, \mathbf{E}_{W+2}, \dots, \mathbf{E}_M]$ spanned by the eigenvectors corresponding to the minimal eigenvalues is called the "noise subspace." The noise subspace is orthogonal to the "signal subspace" $\mathbf{E}_s = [\mathbf{E}_1, \mathbf{E}_2, \dots, \mathbf{E}_W]$ that coincides with the subspace spanned by the data steering vectors $\mathbf{a}(\theta_w)$.

The first property is used in the determination of the number of wavefronts impinging on the array. The second property is used in the estimation of the wavefront shapes.

Estimation of the number of wavefronts

In practice, the data covariance matrix is unknown and must be estimated from the recorded data. After subtracting the mean in the offset direction, the maximum-likelihood estimate of the covariance matrix is

$$\hat{\mathbf{R}}_d = \frac{1}{T} \mathbf{D}\mathbf{D}^*. \quad (10)$$

Property (1) in equation (9a) is not true for the eigenvalues $[\hat{\lambda}_1 \geq \hat{\lambda}_2 \geq \dots \geq \hat{\lambda}_M]$ of the estimated covariance matrix $\hat{\mathbf{R}}_d$. In general, the lowest $M - W$ eigenvalues are different from σ_n^2 . Therefore, the determination of the number of wavefronts is based on a statistical criterion.

Wax and Kailath (1985) have proposed two criteria by which the number of wavefronts can be determined. Their criteria are similar to the ones introduced by Akaike (1973) and Rissanen (1978) for model selection in system identification. These criteria minimize the difference between (1) the log-likelihood function of the maximum-likelihood estimator of the number of parameters in the model and (2) a term penalizing overparameterization of the model.

The criterion (AIC) similar to that developed by Akaike minimizes the function

$$\begin{aligned} \text{AIC}(W) &= -2 \log \left(\frac{\prod_{m=W+1}^M \hat{\lambda}_m^{1/(M-W)}}{\frac{1}{M-W} \sum_{m=W+1}^M \hat{\lambda}_m} \right)^{(M-W)T} \\ &\quad + 2W(2M - W). \end{aligned} \quad (11)$$

The Rissanen criterion, called also the minimum descriptive length (MDL) criterion, minimizes the function

$$\begin{aligned} \text{MDL}(W) &= -2 \log \left(\frac{\prod_{m=W+1}^M \hat{\lambda}_m^{1/(M-W)}}{\frac{1}{M-W} \sum_{m=W+1}^M \hat{\lambda}_m} \right)^{(M-W)T} \\ &\quad + W(2M - W) \log(T). \end{aligned} \quad (12)$$

The two criteria yield the same estimate of the number of wavefronts in most practical situations. The MDL criterion has the theoretical advantage of yielding a consistent estimate of the number of wavefronts, while the AIC tends, asymptotically, to overestimate the number of signals.

In practice, both criteria tend to overestimate the number of signals when few time samples are used in the estimate of the covariance matrix. In seismic applications, it is seldom necessary to detect more than two or three interfering wavefronts; it is therefore useful to set a low limit, two or three, for the maximum number of wavefronts W .

Estimation of the wavefront-shape parameters

Property (2) in equation (9b) ensures that the data steering vectors $\mathbf{a}(\theta_n)$ are orthogonal to the noise subspace \mathbf{E}_n and, consequently, that they belong to the signal subspace. Let $\{\mathbf{a}(\theta)\}$ be the continuum described by the steering vectors for all values of the parameter θ . Then, $\hat{\theta}$ is a solution of the estimation problem if $\mathbf{a}(\hat{\theta})$ belongs to the signal subspace \mathbf{E}_s . The solution is unique when only W vectors in $\{\mathbf{a}(\theta)\}$ belong to \mathbf{E}_s . A sufficient condition for uniqueness is that any $W + 1$ steering vectors are linearly independent. Furthermore, a set of distinct shape parameters implies steering vectors that are linearly independent when the signal is not aliased, i.e., when the spacing between the receivers is smaller than the half-wavelength of the signal.

In practice, the signal and noise subspaces are unknowns to be estimated from $\hat{\mathbf{R}}_d$. Once the number of wavefronts is determined, the maximum-likelihood estimates of the signal and noise subspaces are, respectively, $\hat{\mathbf{E}}_s = [\hat{\mathbf{E}}_1, \hat{\mathbf{E}}_2, \dots, \hat{\mathbf{E}}_W]$, and $\hat{\mathbf{E}}_n = [\hat{\mathbf{E}}_{W+1}, \hat{\mathbf{E}}_{W+2}, \dots, \hat{\mathbf{E}}_M]$, where the $\hat{\mathbf{E}}_m$ are the eigenvectors of $\hat{\mathbf{R}}_d$ (Anderson, 1963). The shape-parameter spectrum can be computed for every θ by measuring the closeness of the "test steering vector" $\mathbf{a}(\theta)$ to the estimated signal subspace $\hat{\mathbf{E}}_s$ or by measuring its orthogonality to the estimated noise subspace $\hat{\mathbf{E}}_n$. Three equivalent expressions for the shape-parameter spectrum are

$$P_{s_1}(\theta) = \sum_{m=1}^W |\mathbf{a}(\theta) * \hat{\mathbf{E}}_m|^2 \tag{13a}$$

and

$$P_{s_2}(\theta) = \frac{1}{1 - \sum_{m=1}^W |\mathbf{a}(\theta) * \hat{\mathbf{E}}_m|^2} \tag{13b}$$

when the signal subspace is used, or

$$P_{n_1}(\theta) = \frac{1}{\sum_{m=W+1}^M |\mathbf{a}(\theta) * \hat{\mathbf{E}}_m|^2} \tag{13c}$$

when the noise subspace is used. $P_{s_2}(\theta)$ and $P_{n_1}(\theta)$ yield exactly the same spectrum, while $P_{s_1}(\theta)$ has the same maxima as the others and therefore is equivalent for the estimation of the shape parameter θ .

Another useful shape-parameter spectrum is

$$P_{n_2}(\theta) = \frac{1}{\sum_{m=W+1}^M \frac{1}{\lambda_m} |\mathbf{a}(\theta) * \hat{\mathbf{E}}_m|^2} \tag{14}$$

In general, the maxima of $P_{n_2}(\theta)$ are different from the maxima of $P_{n_1}(\theta)$. $P_{n_2}(\theta)$ usually yields better estimates when the number of time samples T is small.

Correlated sources and spatial smoothing

The only assumption of the high-resolution methods presented above that would be unrealistic for seismic data is the

lack of correlation between the sources of the interfering wavefronts. When the interfering wavefronts are a primary and a multiple, their waveforms are probably highly correlated.

If two sources are fully correlated, the source covariance matrix \mathbf{R}_s is singular; and therefore, the properties in equations (9) are not true. In practice, even two highly but not fully correlated sources could be unresolvable by eigenstructure spectra.

An effective method to *uncorrelate* the sources of signal before applying the eigenstructure method is to apply "spatial smoothing" while estimating the covariance matrix from the data (Shan et al., 1985; Shan et al., 1987). Spatial smoothing increases the rank of the source covariance matrix \mathbf{R}_s .

The original array of M receivers is divided into K overlapping subarrays of $(M - K + 1)$ receivers, as shown in Figure 1. The covariance matrix $\hat{\mathbf{R}}_d^k$ can be defined for each subarray k , and the *spatially smoothed* covariance matrix is defined as

$$\hat{\mathbf{R}}_d^K = \frac{1}{K} \sum_{k=1}^K \hat{\mathbf{R}}_d^k \tag{15}$$

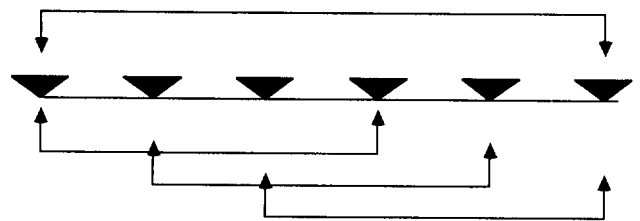
Theoretically, use of the spatial smoothing technique can be justified only for planar wavefronts. However, if the number of subarrays K is small, the spatial smoothing technique can be applied successfully to spherical wavefronts as well.

Spatial smoothing reduces the effective aperture of the array from M to $(M - K + 1)$ and, consequently, reduces slightly the resolution of the method.

COMPARISON WITH THE STACKING METHOD

For narrow-band data, the shape-parameter spectrum computed by the classical method of time corrections followed by stacking can also be expressed as a function of the covariance matrix of the data. The stacking spectrum, or power of the stack, averaged over a temporal window T samples long can be defined as

$$P_{\text{Stack}}(\theta) = \frac{1}{T} \sum_{t=1}^T \left| \frac{1}{M} \sum_{m=1}^M d[m, t + \tau(\theta, m)] \right|^2 \tag{16}$$



K subarrays, M-K+1 receivers each

FIG. 1. Subdivision of the original array of M receivers into K subarrays of $M - K + 1$ receivers for application of *spatial smoothing*.

Time shifting for narrow-band data is achieved by multiplication by a complex exponential; thus the stacking spectrum can be expressed in matrix notation as

$$\begin{aligned} P_{\text{Stack}}(\theta) &= \frac{1}{T} \|\mathbf{a}(\theta)^* \mathbf{D}\|_{L_2}^2 = \frac{1}{T} [\mathbf{a}(\theta)^* \mathbf{D}] [\mathbf{a}(\theta)^* \mathbf{D}]^* \\ &= \frac{1}{T} \mathbf{a}(\theta)^* [\mathbf{D} \mathbf{D}^*] \mathbf{a}(\theta) = \mathbf{a}(\theta)^* \hat{\mathbf{R}}_d \mathbf{a}(\theta). \end{aligned} \quad (17)$$

The stacking spectrum, as a function of the eigenvalues and eigenvectors of the data covariance matrix, is

$$P_{\text{Stack}}(\theta) = \mathbf{a}(\theta)^* \hat{\mathbf{R}}_d \mathbf{a}(\theta) = \sum_{m=1}^M \hat{\lambda}_m |\mathbf{a}(\theta)^* \hat{\mathbf{E}}_m|^2. \quad (18)$$

The stacking spectrum equation (18) differs from the eigenstructure spectrum equation (13a) in two ways: First, and most important, is the weighting by eigenvalues in the summation of equation (18). The weights in equation (18) reward those test steering vectors that are close to eigenvectors corresponding to the largest eigenvalues. On the contrary, the spectrum in equation (13a) is a projection of the test steering vectors onto the signal subspace; it does not prefer any direction in this subspace. Second, the ranges of the summations differ. In the stacking method, the summation is computed for all the eigenvectors. In the eigenstructure method, summation is limited to the eigenvectors belonging to the signal subspace. Limiting the summation reduces the effects of the noise and of the finite number of time samples used in the estimation of the data covariance matrix.

Geometric interpretation

The superior resolution achieved by the eigenstructure spectrum can be easily explained by geometric considerations, as shown in Figure 2. We consider two plane waves, generated by uncorrelated sources with equal power and recorded at noise-free receivers. The data steering vectors are $\mathbf{a}(\theta_1)$ and $\mathbf{a}(\theta_2)$. The first two eigenvectors \mathbf{E}_1 and \mathbf{E}_2 define the signal subspace in the M -dimensional vector space Ω_M spanned by the M orthonormal steering vectors $[\mathbf{a}(\theta = 0), \mathbf{a}(\theta = \pm\Delta\theta), \mathbf{a}(\theta = \pm2\Delta\theta), \dots, \mathbf{a}(\theta = +M\Delta\theta/2)]$, where $\Delta\theta = 2\pi/\omega M\Delta x$. The continuum $\{\mathbf{a}(\theta)\}$, described by the test steering vectors for all θ , is a nonlinear manifold in the vector space Ω_M . This manifold intersects the signal subspace only in $\mathbf{a}(\theta_1)$ and $\mathbf{a}(\theta_2)$.

The eigenstructure spectrum in equation (13a) is a projection of the test steering vectors onto the signal subspace; it thus has two maxima exactly at θ_1 and θ_2 , independent of the angle between the data steering vectors. By contrast, as discussed in the next paragraph, the stacking spectrum has a spurious maximum at $(\theta_1 + \theta_2)/2$ when the angle between the two data steering vectors is too small.

The cosine of the angle α between the two data steering vectors is given by their dot product, assumed here to be a real quantity (for simplification of the algebra),

$$\begin{aligned} \cos \alpha &= \mathbf{a}^*(\theta_1) \mathbf{a}(\theta_2) \\ &= \frac{\sin \{\omega \Delta x M [(\theta_2 - \theta_1)/2]\}}{M \sin \{\omega \Delta x [(\theta_2 - \theta_1)/2]\}} = \frac{\sin (2M\beta)}{M \sin (2\beta)}, \end{aligned} \quad (19)$$

with $\beta = \omega \Delta x (\theta_2 - \theta_1)/4$.

The eigenvalues of the data covariance matrix are

$$\lambda_1 = S(1 + \cos \alpha),$$

$$\lambda_2 = S(1 - \cos \alpha),$$

and

$$\lambda_m = 0 \quad \text{for } m > 2. \quad (20)$$

The difference between the first two eigenvalues increases as the angle between the data steering vectors decreases. The first two eigenvectors are

$$\mathbf{E}_1 = \frac{\mathbf{a}(\theta_1) + \mathbf{a}(\theta_2)}{\sqrt{2 + 2 \cos \alpha}}$$

and

$$\mathbf{E}_2 = \frac{\mathbf{a}(\theta_1) - \mathbf{a}(\theta_2)}{\sqrt{2 - 2 \cos \alpha}}. \quad (21)$$

The test steering vector $\mathbf{w} = \mathbf{a}[(\theta_1 + \theta_2)/2]$ is closer to the first eigenvector than is either $\mathbf{a}(\theta_1)$ or $\mathbf{a}(\theta_2)$. Therefore, the smaller the angle α between the data steering vectors, the larger the weight given by the stacking method [equation (18)] to the test steering vector \mathbf{w} in comparison with the weights of the steering vectors $\mathbf{a}(\theta_1)$ and $\mathbf{a}(\theta_2)$. For small α , the stacking spectrum would exhibit a single, spurious maximum for the shape parameter equal to $(\theta_1 + \theta_2)/2$.

When the two plane waves are correlated in time, determining their resolution—by either the stacking spectrum or the eigenstructure spectrum—is more difficult than when they are uncorrelated. Assuming that S_{12} , the off-diagonal term of the source covariance matrix \mathbf{R}_s , is real, the eigenvalues of the data covariance matrix are:

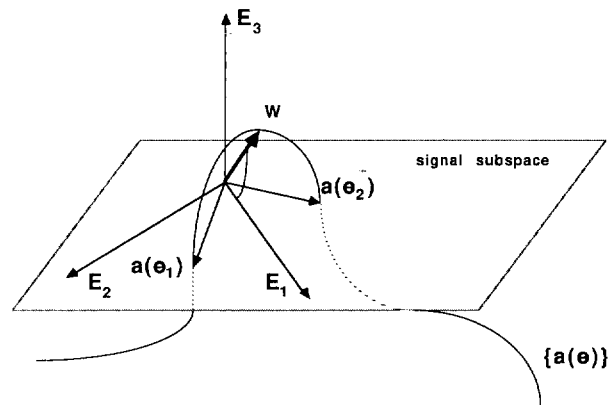


FIG. 2. Geometric interpretation of the stacking spectrum and of the eigenstructure spectrum for two incident plane waves. The first two eigenvectors \mathbf{E}_1 and \mathbf{E}_2 define the *signal subspace*. The data steering vectors $\mathbf{a}(\theta_1)$ and $\mathbf{a}(\theta_2)$ lie in the signal subspace; therefore the eigenstructure spectrum shows two maxima for $\mathbf{a}(\theta_1)$ and $\mathbf{a}(\theta_2)$. The vector \mathbf{w} belongs to the continuum $\{\mathbf{a}(\theta)\}$, but not to the signal subspace. When the angle α between the data steering vectors is too small, the stacking spectrum shows a maximum for \mathbf{w} instead of two maxima for $\mathbf{a}(\theta_1)$ and $\mathbf{a}(\theta_2)$.

$$\lambda_1 = (S + S_{12})(1 + \cos \alpha),$$

$$\lambda_2 = (S - S_{12})(1 - \cos \alpha),$$

and

$$\lambda_m = 0 \quad \text{for } m > 2, \quad (22)$$

while the eigenvectors are the same as they are for uncorrelated sources [equation (21)]. The correlation between the sources decreases the "effective angle" between the data steering vectors. When the sources are fully correlated, only one eigenvalue is different from zero. Then the stacking method and the eigenstructure spectrum yield the same spectra.

Figure 3a shows a stacking spectrum superimposed upon an eigenstructure spectrum computed using equation (13b). The two spectra are normalized to one. The data are two monochromatic plane waves with frequency 25 Hz and ray parameters $\theta_1 = 0.2$ s/km and $\theta_2 = 0.25$ s/km. The array has 40 receivers at 10 m spacing; the time sampling is 4 ms. With these parameters, the angle α between the two data steering vectors is approximately 50° . The stacking spectrum has not resolved the two plane waves and shows a maximum at $(\theta_1 + \theta_2)/2$. The eigenstructure spectrum has resolved the two plane waves and shows two maxima exactly at θ_1 and θ_2 .

Figure 3b shows the stacking spectrum and the eigenstructure spectrum obtained when the sources are fully correlated. Neither method resolves the two correlated signals. On the other hand, after spatial smoothing is applied, as

presented in the previous section, the eigenstructure method resolves the two wavefronts. Figure 3c shows the spectra obtained when spatial smoothing is applied by dividing the original array into 11 subarrays. The spatially smoothed covariance matrix is used for the computation of both spectra.

In this section, we considered the simple case of two planewaves and noiseless data. We did not consider the effect of datalength limitations on the quality of the estimate of the data covariance matrix. A more general discussion of the statistical performances of the eigenstructure method is in Kaveh and Barabell (1986) and Wang and Kaveh (1986).

THE WIDE-BAND METHOD

Seismic data are wide-band; before the narrow-band methods described above can be applied, the data must be decomposed into frequency components, either by Fourier transform or by filtering with a bank of band-pass filters.

The results obtained for different frequencies can be combined to yield a robust solution to the original wide-band estimation problem. This combination is possible because the number of wavefronts and the shape parameters remain the same for all frequencies. The frequency components can be combined at different stages of the processing: averages can be made of either the final estimates of the spectra (Wax et al., 1984) or of the estimates of the covariance matrices of the different components (Wang and Kaveh, 1985). In the latter case, the frequency components must be linearly

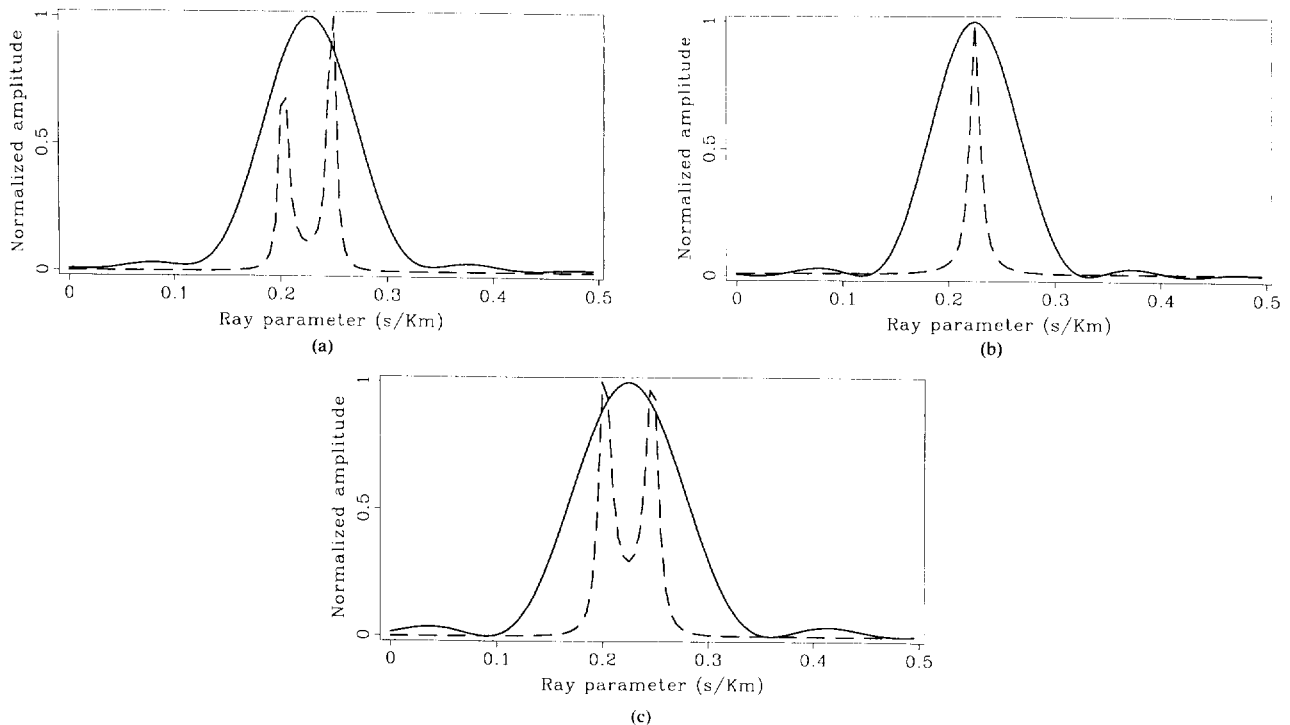


FIG. 3. Stacking spectrum (solid line) and eigenstructure spectrum (dashed line) for two monochromatic plane waves. (a) Plane waves with $\theta_1 = 0.2$ s/km and $\theta_2 = 0.25$ s/km. The angle α between the two data steering vectors is 50° . Only the eigenstructure method resolves the two wavefronts. (b) The plane waves are fully correlated. Neither method resolves the two wavefronts. (c) The plane waves are fully correlated and spatial smoothing is applied. The eigenstructure spectrum resolves the two wavefronts.

transformed so that the covariance matrices are approximately *coherent* with each other. The rationale of this method is that averaging the covariance matrices increases the statistical robustness of the estimates. This result is particularly important to the application to seismic data, in which only a few time samples can be used in the estimation of the covariance matrices. The disadvantage of combining the covariance matrices is that they are only approximately coherent.

In our computations, we combine both methods described above: we average correlation matrices for nearby frequencies, and average spectra from different frequency bands.

First the data are time corrected according to the moveout $\tau(m, \bar{\theta})$, where $\bar{\theta}$ is an estimate of the shape parameter. This estimate can be obtained, for instance, by stacking spectrum analysis. The time correction increases the coherency of the different frequency components of the wavefronts with shape parameters close to $\bar{\theta}$. Before the time correction is applied, the phase difference between two frequency components, with angular frequencies ω_1 and ω_2 and common shape parameter $\theta + d\theta$, is

$$\Delta\phi = (\omega_1 - \omega_2)\tau(m, \bar{\theta} + d\theta).$$

After the time correction, the same phase difference becomes

$$\Delta\phi = (\omega_1 - \omega_2)\tau(m, d\theta).$$

A bank of band-pass filters is used to decompose the time-corrected data into a few wide frequency bands. The data covariance matrix is estimated for each band. Estimating the covariance matrix for a wide frequency band implicitly averages the estimates of the covariance matrices of the different frequency components within the band. Therefore, the wider the frequency bands, the more robust the estimates of the covariance matrices. Robustness is gained at the expense of resolution, because the different frequency

components inside each band are not completely coherent, even after the time correction.

To determine the number of wavefronts W impinging on the array of receivers, we select the value that minimizes the sum of the MDL criteria [equation (12)] for each of the F frequency bands,

$$\sum_{f=1}^F \text{MDL}_f(W) = \sum_{f=1}^F \left[-2 \log \left(\frac{\prod_{m=W+1}^M \hat{\lambda}_m(f)^{1/(M-W)}}{\frac{1}{M-W} \sum_{m=W+1}^M \hat{\lambda}_m(f)} \right)^{(M-W)T} + W(2M-W) \log(T) \right] \quad (23)$$

Once the number of wavefronts W is determined, the parameter spectra of all the bands are averaged to obtain the shape-parameter spectrum. For example, the wide-band version of equation (13a) is

$$P_{s_1}(\theta) = \frac{1}{F} \sum_{f=1}^F \sum_{m=1}^W |\mathbf{a}(\theta, f) * \hat{\mathbf{E}}_m(f)|^2. \quad (24)$$

If the signal-to-noise ratio (S/N) in each band is known, the terms summed in equations (23) and (24) could be weighted accordingly so as to increase the quality of the estimates.

The estimation of the covariance matrix from noisy data

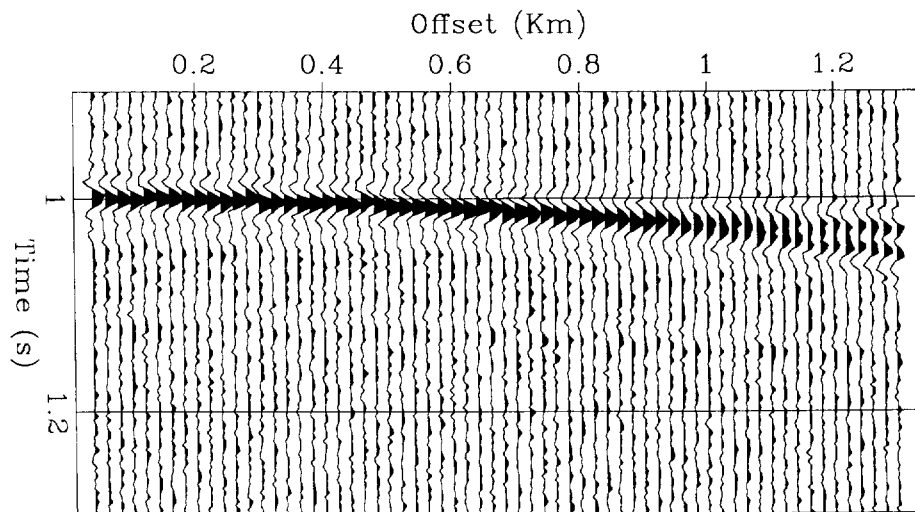


FIG. 4. Synthetic CMP gather used in comparing the performance of the eigenstructure method with that of the stacking method. Two hyperbolic reflections are recorded with stacking slownesses $\theta_1 = 0.2$ s/km and $\theta_2 = 0.25$ s/km. The S/N of the data is 2. Random time shifts were applied to the traces. Reflection amplitudes increase linearly with offset.

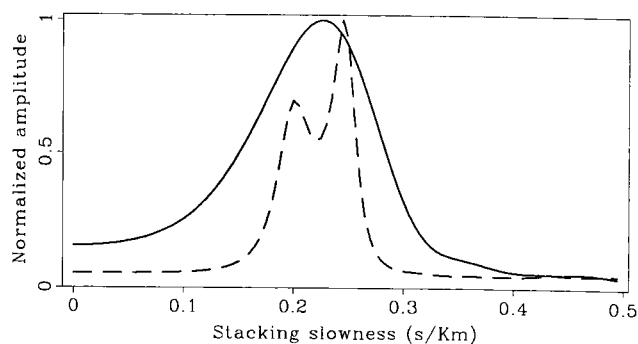


FIG. 5. The stacking spectrum (solid line) and the eigenstructure spectrum (dashed line) for the data shown in Figure 4. The eigenstructure spectrum has well resolved the two reflections; the stacking spectrum has not.

can be improved by *partial stacking*, that is, summation of adjacent moveout-corrected traces. A further advantage of partial stacking is the reduction of the dimensions of the covariance matrix and therefore also of the computational cost of the procedure. To avoid the aliasing of wavefronts with shape parameters significantly different from θ , one can perform spatial dip filtering in conjunction with partial stacking.

APPLICATION TO VELOCITY ANALYSIS

The wide-band method can be applied directly to a common-midpoint (CMP) gather for the estimation of high-resolution velocity spectra. Velocity spectra are slightly different from and more complicated than ray-parameter

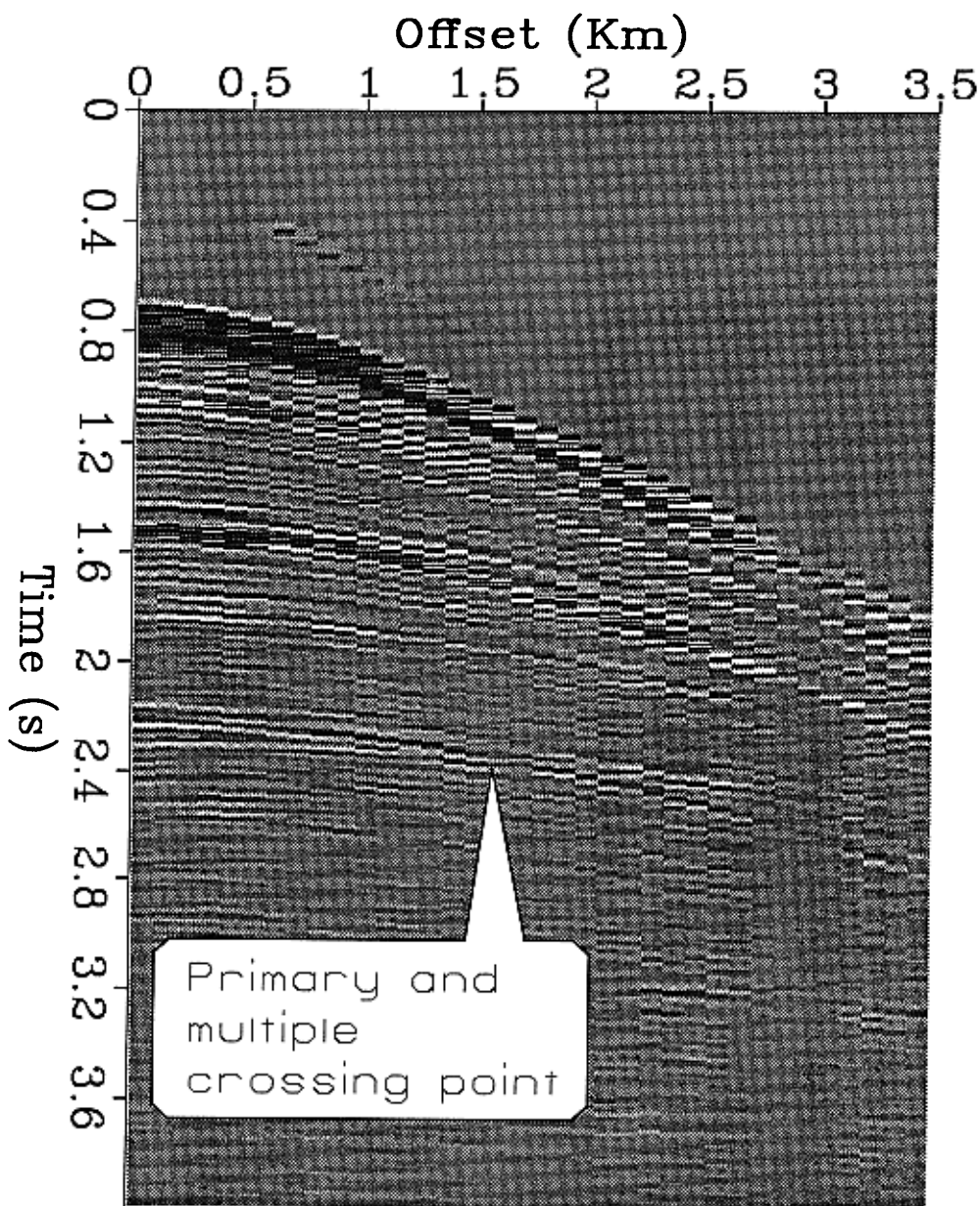


FIG. 6. A marine CMP gather recorded offshore southern California. We estimate local spectra at the intersection of a primary reflection and a water-bottom multiple. The region of interest is indicated by the balloon.

spectra. The properties of the steering vectors with hyperbolic delays of equation (3) are different from those of the steering vectors with linear delays [equation (2)]. In particular, the data steering vectors with hyperbolic delays are never orthogonal.

Figure 4 shows the synthetic CMP gather used in our comparison of the performances of the eigenstructure method and of the usual stacking method. Two hyperbolic reflections are recorded by an array of 64 geophones, spaced 20 m apart, with nearest offset of 40 m. The stacking slownesses are $\theta_1 = 0.2$ s/km and $\theta_2 = 0.25$ s/km; the zero-offset time is 1 s. The sources are wide-band, from 10 Hz to 50 Hz, correlated with correlation coefficient $\rho = 0.87$. The S/N is 2, with Gaussian white noise added. For testing the new method in more realistic situations, we applied to the traces random time shifts, drawn from a Gaussian distribution with standard deviation of 1 ms, and also increased the reflection amplitudes linearly with offset.

Figure 5 shows the stacking spectrum and the eigenstructure spectrum for the data shown in Figure 4; the spectra were computed at the correct zero-offset time of 1 s and normalized to one. For the computation of the eigenstructure spectrum, time corrections corresponding to a stacking slowness $\bar{\theta} = 0.225$ s/km were applied; partial stacking reduced the number of traces to eight. The data were then decomposed into six frequency bands. The results from the different frequency bands were combined following equations (23) and (24). The eigenstructure spectrum has well resolved the two reflections, while the stacking spectrum has not.

APPLICATION TO LOCAL SLANT STACKS

Using shape-parameter spectra, we implicitly assume that the data can be modeled by a simple propagation model as either plane waves or hyperbolic reflections. These simple models well predict the data locally but they are not accurate in modeling data recorded over a larger area (a cable length). Thus, for instance, traveltimes curves are well approximated by straight lines only in the Fresnel zone; also, they are not exactly hyperbolic when the velocity is not constant.

There are thus two conflicting needs: short arrays to maintain the accuracy of the model and the spatial resolution

of the estimates and long arrays to provide the shape-parameter resolution. Without decreasing the resolution of the estimates, shorter arrays can be used in the eigenstructure spectrum than in the stacking spectrum. Therefore, high-resolution spectra can prove crucial to the success of a local-spectra estimation method.

Field-data example

The advantages of the eigenstructure method are confirmed by its application to the marine CMP gather shown in Figure 6. The data were recorded offshore southern California. The data sampling rate is 2 ms in time and 32 m in the offset dimension. We estimate local spectra at the intersection of a primary reflection and a water-bottom multiple.

The broad frequency spectrum of the data, extending up to 100 Hz, allows the comparison of stacking spectra and eigenstructure spectra for different frequency bands. As expected, the stacking spectrum cannot resolve the two reflections when the frequency of the data is too low. Furthermore, in order to study the tradeoff between ray-parameter resolution and spatial resolution of the local estimates, we varied the length of the subarray used in the estimation of local spectra. The results confirm that the eigenstructure method has a higher spatial resolution than the stacking method.

Figure 7 shows the normalized ray-parameter spectra estimated from the original gather with frequencies from 15 Hz to 100 Hz. The spectra shown in Figure 7a are estimated with a subarray of eight geophones, centered at the intersection of the two reflections; while the spectra shown in Figure 7b are estimated with a subarray of six geophones. As reference values, we used the reflection ray parameters estimated by hand from a plot of the gather; even if carefully measured, these values are affected by some error. To compute the eigenstructure spectrum (dashed line), we first applied a time correction corresponding to the ray parameter $\bar{\theta} = 0.32$ s/km; then we decomposed the data into ten frequency bands. When the stacking spectrum (solid line) was computed, semblance as a coherency measure along the offset direction was used. Both methods resolve the two reflections, independent of the subarray length. The maxima

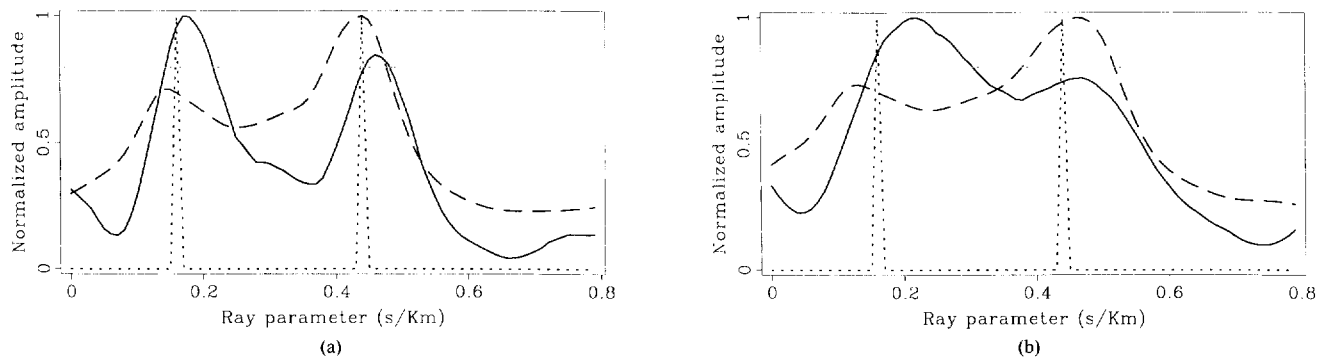


FIG. 7. Normalized, local ray-parameter spectra. The dashed lines are the eigenstructure spectra; the solid lines are the stacking spectra. The dotted lines show two spikes at the values of the ray parameter estimated by hand. The data have frequencies from 15 Hz to 100 Hz. (a) A subarray of eight geophones is used. (b) A subarray of six geophones is used. The maxima of the spectra agree well with the ray-parameter values picked by hand.

of the spectra agree well with the ray-parameter values estimated by hand.

The resolution of local spectra depends also on the frequency bandwidth of the data. Therefore, we computed the spectra after having band-passed the gather with different high-frequency cutoff values. Figure 8 shows the local spectra computed from a gather with frequencies from 15 Hz to 70 Hz. The eigenstructure spectrum is computed after a time correction, corresponding to the ray parameter $\bar{\theta} = 0.32$ s/km, is applied and after the data are decomposed into six frequency bands. For this bandwidth, the stacking spectrum (solid line in Figure 8b) fails to resolve the two reflections when a subarray of six geophones is used. The eigenstructure spectrum resolves both reflections well, independent of the subarray length.

Figure 9 shows the local spectra computed from a gather with frequencies from 15 Hz to 33 Hz. The eigenstructure spectrum is computed after a time correction, corresponding to the ray parameter $\bar{\theta} = 0.32$ s/km, is applied and after the data are decomposed into four frequency bands. In this extreme case, the stacking method (solid lines) cannot resolve the two reflections when either array aperture is used. By contrast, the eigenstructure spectra (dashed lines)

are successful in resolving both reflections even when a subarray of only six geophones is used.

CONCLUSIONS

The methods for seismic velocity analysis presented in this paper are based on the eigenstructure decomposition of the covariance matrix of the data; these methods succeed in resolving closely interfering reflections better than do the usual stacking methods.

The data recorded by an array of geophones are modeled as the superposition of wavefronts; no specific assumptions about the geology or the source wavelet are made. This general model leads to a family of eigenstructure spectral estimators, which includes the stacking method; thus the comparison of spectral estimation methods can be made in a unified framework.

The new eigenstructure methods for velocity analysis are closely related to methods that are well-known in other areas, such as 1-D signal processing or radar and sonar signal processing. High-resolution spectral estimates are rather sensitive to the estimates of covariance matrices, which in seismic applications need to be obtained from only a few

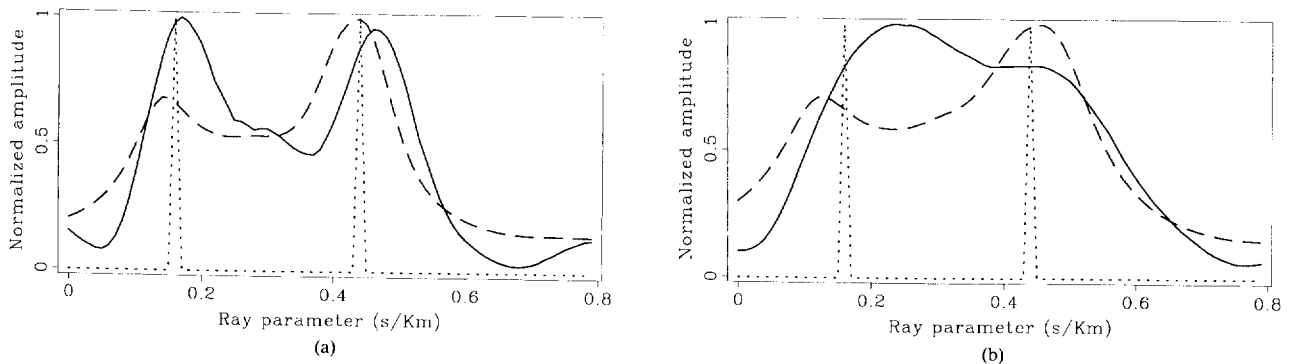


Fig. 8. The stacking spectra (solid lines) and the eigenstructure spectra (dashed lines) computed when the data frequencies extend from 15 Hz to 70 Hz. The dotted lines show two spikes at the values of the ray parameter estimated by hand. (a) A subarray of eight geophones is used. (b) A subarray of six geophones is used. The stacking spectrum cannot resolve the two reflections when the shorter subarray is used (b).

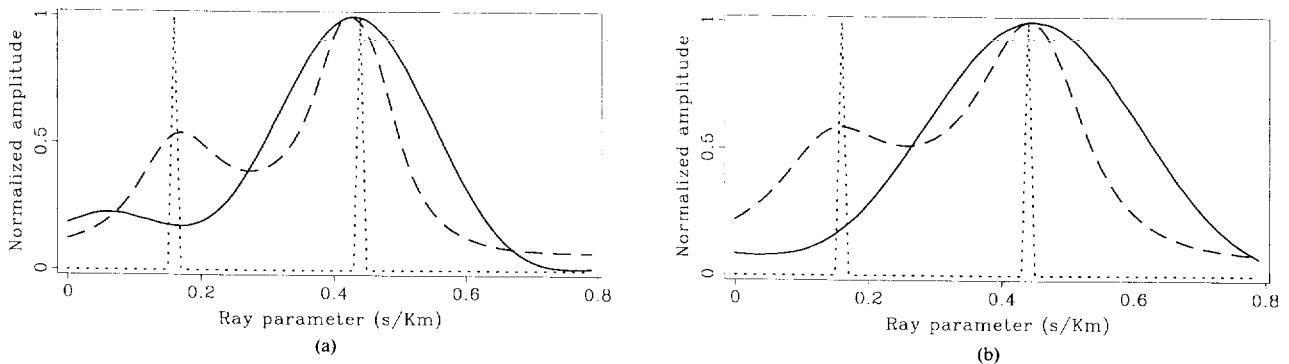


Fig. 9. The stacking spectra (solid lines) and the eigenstructure spectra (dashed lines) when the data have frequencies from 15 Hz to 33 Hz. The dotted lines show two spikes at the values of the ray parameter estimated by hand. (a) A subarray of eight geophones is used. (b) A subarray of six geophones is used. The stacking spectrum fails to resolve the reflections with either subarray.

data samples in time. Taking advantage of the full bandwidth of the seismic data, we have suggested methods for the robust estimation of covariance matrices from short segments of data and for the combination of spectral estimates from different frequency bands.

For eigenstructure methods, unlike for stacking, the number of wavefronts impinging at the array must be estimated. This step requires that a decision be made based on a statistical criterion; this is to some extent subjective. However, once the covariance matrix and its eigenstructure decomposition are computed, the cost of computing several spectral estimates—for different numbers of incident wavefronts or for different spectral norms—is negligible.

Often in seismic processing, velocity analysis is followed by suppression of some events, for instance water-bottom multiples, according to differences in apparent velocities. The concepts of eigenstructure decomposition could be applied also to the problem of wavefront separation; the eigenstructure decomposition already computed for the velocity analysis could be used.

The field-data example and the synthetic examples show the utility of the high-resolution methods for some practical applications. In particular, the field-data example demonstrates the gain in lateral resolution that can be achieved when the estimation of local spectra uses the eigenstructure method.

ACKNOWLEDGMENTS

We would like to thank Professor Fabio Rocca for challenging us to improve our understanding of the subject by asking questions that are simple in the formulation, but hard to answer.

We are grateful to Professor Jon Claerbout and the sponsors of the Stanford Exploration Project for the financial support. We also would like to thank British Petroleum and Dr. Paul Fowler for providing the field data we used.

REFERENCES

- Akaike, H., 1973, Information theory and an extension of the maximum likelihood principle: *Proc. 2nd Int. Symp. Inform. Theory*, suppl. Problems of Control and Inform. Theory, 267-281.
- Anderson, T. W., 1963, Asymptotic theory for principal component analysis: *Ann. J. Math. Stat.*, **34**, 122-148.
- Bienvenu, G., and Kopp, L., 1983, Optimality of high resolution array processing using the eigensystem approach: *IEEE Trans. Acoust. Speech, Signal Processing*, **31**, 1235-1248.
- Burg, J. P., 1975, Maximum entropy spectral analysis: Ph.D. thesis, Stanford Univ.
- Capon, J., 1969, High-resolution frequency-wavenumber spectral analysis: *IEEE Proceedings*, **57**, 1408-1418.
- Claerbout, J. F., 1976, Fundamentals of geophysical data processing: McGraw-Hill Book Co., 20-23.
- Kaveh, M., and Barabell, A., 1986, The statistical performance of the MUSIC and the minimum-norm algorithms in resolving plane waves in noise: *IEEE Trans. Acoust. Speech, Signal Processing*, **34**, 331-341.
- Key, S. C., Kirilin, L., and Smithson, S. B., 1987, Seismic velocity analysis using maximum-likelihood weighted eigenvalue ratios: 57th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded abstracts, 461-464.
- Kostov, C., and Biondi, B., 1987, Improving resolution of slant stacks using beam stacks: 57th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded abstracts, 792-794.
- Mars, J., Glangeaud, F., Lacoume, J. L., Fourmann, J. M., and Spitz, S., 1987, Separation of seismic waves: 57th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded abstracts, 489-492.
- Pisarenko, V. F., 1972, On the estimation of spectra by means of non-linear functions of the covariance matrix: *Geophys. J. Roy. Astr. Soc.*, **28**, 511-531.
- Rissanen, J., 1978, Modeling by shortest data description: *Automatica*, **14**, 465-471.
- Schmidt, R., 1986, Multiple emitter location and signal parameter estimation: *IEEE Trans. Antennas Propagat.*, **34**, 276-280.
- Sguazzero, P., and Vesnaver, A., 1987, A comparative analysis of algorithms for stacking velocity estimation, in Bernabini, M., Carrion, P., Jacovitti, G., Rocca, F., Freitel, S., Eds., *Deconvolution and inversion*: Blackwell Scientific Publ., 267-286.
- Shan, T. J., Paulraj, A., and Kailath, T., 1987, On smoothed rank profile tests in eigenstructure methods for detection-of-arrival estimation: *IEEE Trans. Acoust. Speech, Signal Processing*, **35**, 1377-1385.
- Shan, T. J., Wax, M., and Kailath, T., 1985, On spatial smoothing for direction-of-arrival estimation of coherent signals: *IEEE Trans. Acoust. Speech, Signal Processing*, **33**, 806-811.
- Sword, C. H., 1987, Tomographic determination of interval velocities from reflection seismic data: The method of controlled directional reception: Ph.D. thesis, Stanford Univ.
- Taner, M. T., and Koehler, F., 1969, Velocity spectra-digital computer derivation and applications of velocity functions: *Geophysics*, **34**, 859-881.
- Wang, H., and Kaveh, M., 1985, Coherent signal-subspace processing for the detection and estimation of angles of arrival of multiple wide-band sources: *IEEE Trans. Acoust. Speech, Signal Processing*, **33**, 823-831.
- , 1986, On the performance of signal-subspace processing Part I: Narrow band systems: *IEEE Trans. Acoust. Speech, Signal Processing*, **34**, 1201-1209.
- Wax, M., and Kailath, T., 1985, Detection of signals by information theory criteria: *IEEE Trans. Acoust. Speech, Signal Processing*, **33**, 387-392.
- Wax, M., Shan, T. J., and Kailath, T., 1984, Spatio-temporal spectral analysis by eigenstructure methods: *IEEE Trans. Acoust. Speech, Signal Processing*, **32**, 817-827.