Chapter 5

High-resolution ray-parameter spectra using eigenstructure methods

5.1 INTRODUCTION AND OVERVIEW

The velocity-estimation method presented in this thesis is based on the measurement of the reflections' moveouts by use of beam stacks. Beam stacks measure the data coherency along hyperbolic trajectories. The data coherency can be measured with different criteria; in Chapter 2 I presented the formulas for using stacking and semblance as coherency criteria and I used semblance for the field-data application presented in the last chapter. Semblance spectra are fast to compute and are robust, but their resolution is limited by the length of the stacking trajectory (Appendix A). This limitation in resolution can cause problems with the velocity-estimation procedure when there are two or more closely interfering wavefronts. The interfering wavefronts can be primaries from reflectors with conflicting dips, or a primary and an intrabed or peg-leg multiple. In these cases, the spectra may show fewer maxima than the number of the wavefronts, and the resulting estimation would be biased. In this chapter I introduce a new method for computing local spectra; this method has higher resolution than semblance (Biondi and Kostov, 1989). For velocity estimation, the potential advantage of the new method is that the trajectories used for the estimation of high-resolution spectra can be shorter than those used for the computation of conventional semblance spectra. The accuracy of the information on the moveouts of the reflections is increased, and thus the resolution of the velocity estimation is improved.

Figure 5.1 shows an examples of a marine CMP gather containing water-bottom multiples interfering with primary reflections. The interference is particularly evident in the

region indicated by the balloon. When the coherency of the data is measured by semblance, beam stacks can resolve the primaries reflections from the multiples only if long stacking trajectories are used. Figure 5.2 shows the results of beam stacking the CMP gather of Figure 5.1 with two different lengths for the stacking trajectories. Using stacking trajectories 700 m long (left panel) the primaries and the multiples reflections are not resolved. Only when staking trajectory 1500 m long are used (right panel) the semblance peaks corresponding to the primaries reflections are clearly separated from the semblance peaks corresponding to the multiples.

The resolution of frequency spectra of time series is improved with the introduction of a model for the data (Capon, 1969; Pisarenko, 1972; Burg, 1975). Similarly, the spectral method presented in this chapter achieves high resolution because a "sparse" model is assumed for the data. The data are modeled as a superposition of a finite number of wavefronts and additive noise. When the wavefronts are narrow-band, the spatial covariance matrix of the data has a particular structure, and the properties of its eigenstructure can be exploited to efficiently compute high-resolution spectra. (Bienvenu and Kopp, 1983; Schmidt, 1986). Eigenstructure methods can be applied to broad-band data, as seismic reflections: the data is decomposed in narrow-band components and the results obtained for the different frequency bands are subsequently combined (Wax et al., 1984; Wang and Kaveh, 1985). The redundancy of information provided by the different frequency components increases the reliability of the spectral estimation. On the other hand, the reliability of the spectra is adversely affected by the nonstationarity of seismic reflections. The ray parameter of the reflections can vary rapidly with time and thus only few time samples can be used for the estimation of the spatial covariance matrix.

In the next section I present the eigenstructure algorithms' application to narrow-band signals. Then I compare the new methods with the classical stacking method and show a geometric explanation for the superiority of the former. Section 5.4 contains an extension of the narrow-band method to the general case of a wide-band signal. In this section I discuss the particular features of the seismic problem, and the trade-off between resolution and robustness in the application of the method to seismic signals. The last section illustrates, with a field-data example, the application of the high-resolution methods to local slant stacks.

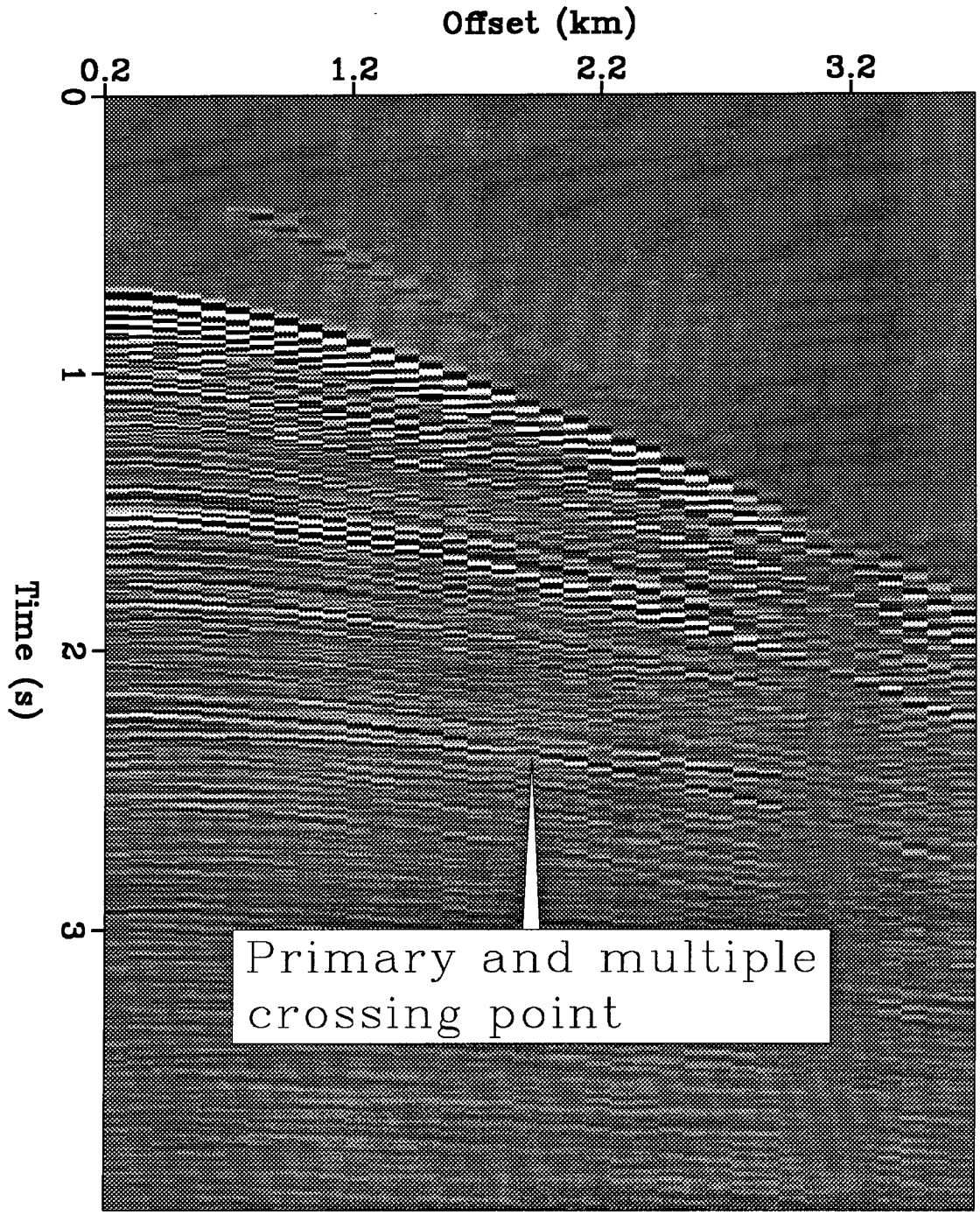

FIG. 5.1. A marine CMP gather recorded offshore Southern California. Water-bottom multiples interfere with primaries reflections. The interference is particularly evident in the region indicated by the balloon.

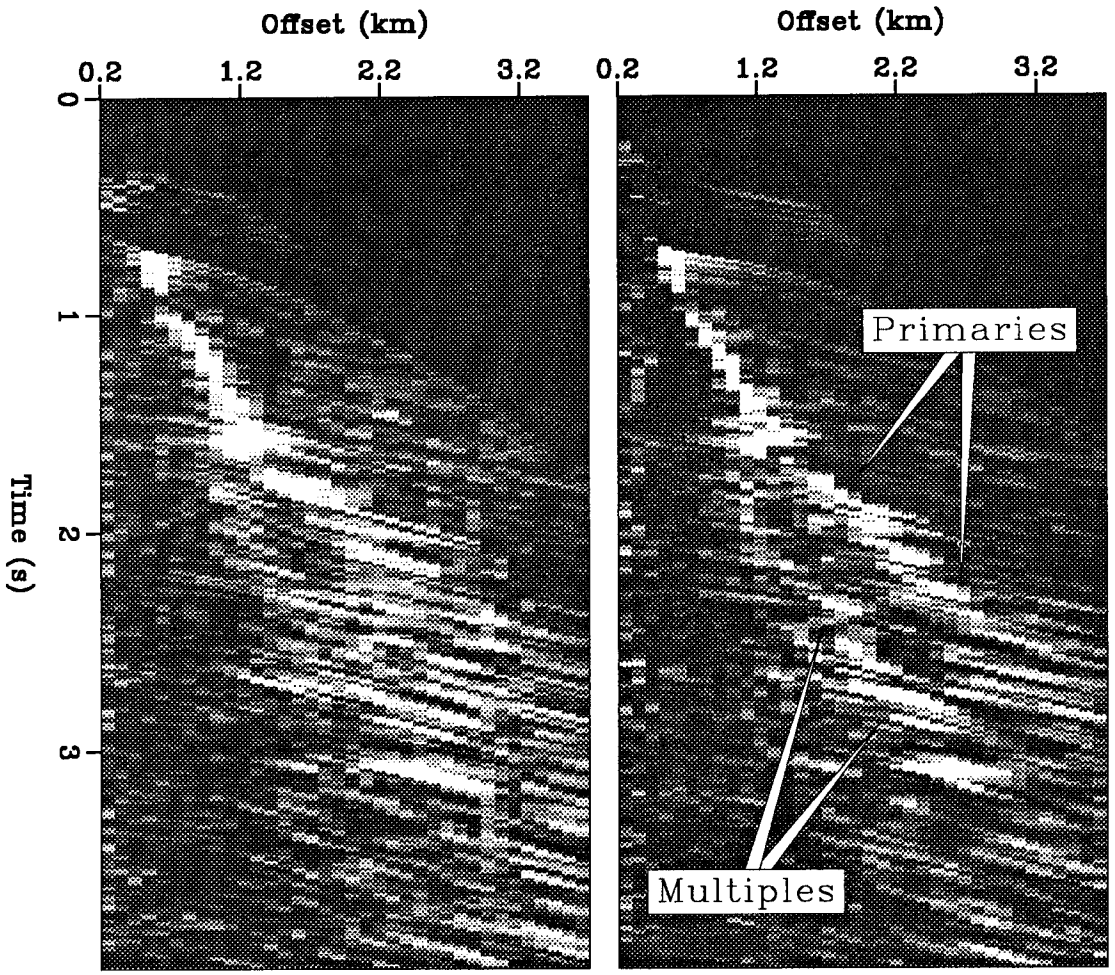

FIG. 5.2. Semblance panels obtained by beam stacking the CMP gather shown in Figure 5.1. The two panels were computed using two different lengths for the stacking trajectories: 700 m (left) and 1500 m (right). The primaries are well resolved from the multiples only by use of the longest stacking trajectories.

5.2 THE NARROW-BAND METHOD

Eigenstructure spectra can achieve higher resolution than stacking spectra because they are based on a model for the data. Seismic data are modeled as a superposition of W broad-band wavefronts and additive noise. The data $\mathrm{Data}(t,x)$, recorded at time t and surface location x, are

$$\mathrm{Data}(t,x) = \sum_{w=1}^{W} s_w[t - \mathrm{T}_w(x;\theta_w)] \; + \; n(t,x), \tag{5.1}$$

where $n(t,x)$ is the additive noise, assumed to be uncorrelated with the sources $s_w(t)$. The traveltimes of the wavefronts are denoted by $\mathrm{T}_w(x;\theta_w)$. When the wavefronts are well approximated with plane waves, and local slant stacks are computed, the traveltimes are expressed as

$$\mathrm{T}_w(x;\theta_w) = \bar{t} + \theta_w(x - \bar{x}), \tag{5.2}$$

where θ_w is the ray parameter of the w wavefront and $(\bar{t},\bar{x})$ is a reference point. When hyperbolic beam stacks are used [equation (2.8)] the traveltimes are

$$\mathrm{T}_w(x;\theta_w) = \sqrt{\bar{t}^2 - \bar{x}\bar{t}\theta + \frac{x^2}{\bar{x}}\theta\bar{t}}. \tag{5.3}$$

The eigenstructure spectra can also be applied for the computation of velocity spectra as well as ray-parameter spectra (Biondi and Kostov, 1989). In this case θ_w is the stacking slowness, and the time traveltimes are given by the conventional hyperbolic function

$$\mathrm{T}_w(x;\theta_w) = \sqrt{\bar{t}^2 + \theta_w^2 x^2}, \tag{5.4}$$

where $\bar{t}$ is the zero-offset traveltime.

The high-resolution method that I present in this chapter is based on the properties of the eigenstructure of the covariance matrix of narrow-band signals. The data modeled with equation (5.1) is broad-band, but it can be easily transformed in a narrow-band, analytical signal, by use of a bandpass filter and of 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). For a narrow-band signal, the time shifts are complex exponentials, and the data belonging to the band with central angular pulsation ω can be expressed as

$$\begin{aligned}\mathrm{Data}_\omega(t,x) &= \sum_{w=1}^{W} s_{w,\omega}[t - \mathrm{T}_w(x;\theta_w)] \;+\; n_\omega(t,x) = \\ \sum_{w=1}^{W} s_{w,\omega}(t)\, e^{j\omega \mathrm{T}_w(x;\theta_w)} + n_\omega(t,x) &= \sum_{w=1}^{W} \sqrt{M} s_{w,\omega}(t)\, \frac{e^{j\omega \mathrm{T}_w(x;\theta_w)}}{\sqrt{M}} + n_\omega(t,x),\end{aligned} \tag{5.5}$$

where the normalization factor M is equal to the number of receivers. 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's window is expressed as

$$\mathbf{D} = \mathbf{A}(\mathbf{\Theta})\mathbf{S} + \mathbf{N}, \tag{5.6}$$

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

$$\mathbf{a}(\theta_w) = \frac{1}{\sqrt{M}} \begin{pmatrix} e^{j\omega \mathrm{T}_w(x_1;\theta_w)} \\ \cdot \\ \cdot \\ e^{j\omega \mathrm{T}_w(x_m;\theta_w)} \\ \cdot \\ \cdot \\ e^{j\omega \mathrm{T}_w(x_M;\theta_w)} \end{pmatrix}. \tag{5.7}$$

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

$$\mathbf{R}_d = E[\mathbf{D}\mathbf{D}^H] = \mathbf{A}(\mathbf{\Theta})E[\mathbf{S}\mathbf{S}^H]\mathbf{A}(\mathbf{\Theta})^H + \mathbf{R}_n = \mathbf{A}(\mathbf{\Theta})\mathbf{R}_s\mathbf{A}(\mathbf{\Theta})^H + \mathbf{R}_n, \tag{5.8}$$

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

For simplicity, I assume that the noise is spatially uncorrelated and has equal power at all the receivers. Equation (5.8) becomes then

$$\mathbf{R}_d = \mathbf{A}(\mathbf{\Theta})\mathbf{R}_s\mathbf{A}(\mathbf{\Theta})^H + \sigma_n^2\mathbf{I}. \tag{5.9}$$

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).

5.2.1 Properties of the eigenstructure of the data covariance matrix

When the sources covariance matrix $\mathbf{R}_s$ is not singular (i.e. there are no pairs of fully correlated sources) and there are no linearly dependent steering vectors $\mathbf{a}_w(\theta_w)$, the matrices $\mathbf{R}_s$ and $\mathbf{A}(\mathbf{\Theta})$ have rank W (here I assume that $M \geq W$), and therefore the matrix $\mathbf{A}(\mathbf{\Theta})\mathbf{R}_s\mathbf{A}(\mathbf{\Theta})^H$ also has rank W.

Let $[\lambda_1 \geq \lambda_2 \ ... \ \geq \lambda_M]$ be the eigenvalues, and let $[\mathbf{E}_1, \mathbf{E}_2, ...\mathbf{E}_M]$ be the eigenvectors of the data covariance matrix $\mathbf{R}_d$. Equation (5.9) implies that the eigenvalues of $\mathbf{R}_d$ are of the form $\lambda = \mu^2 + {\sigma_n}^2$, with μ^2 the eigenvalues of $\mathbf{A}(\mathbf{\Theta})\mathbf{R}_s\mathbf{A}(\mathbf{\Theta})^H$. 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:

$$\lambda_{W+1} = \lambda_{W+2} = ... = \lambda_M = \sigma_n^2 \tag{5.10}$$

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

$$\mathbf{a}(\theta_w)^H \mathbf{E}_m = 0, \qquad m = W+1, W+2, ..., M. \tag{5.11}$$

The subspace $\mathbf{E}_n = [\mathbf{E}_{W+1}, \mathbf{E}_{W+2}, ..., \mathbf{E}_M]$ spanned by the eigenvectors corresponding to the minimal eigenvalues is called the "noise subspace". The noise subspace is orthogonal to the "signal subspace" $\mathbf{E}_s = [\mathbf{E}_1, \mathbf{E}_2, ..., \mathbf{E}_W]$ 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 wavefronts' ray-parameter.

5.2.2 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 along offset, the maximum-likelihood estimate of the covariance matrix is

$$\hat{\mathbf{R}}_d = \frac{1}{T} \ \mathbf{D}\mathbf{D}^H. \tag{5.12}$$

Property (1) in equation (5.10) does not hold exactly true for the eigenvalues $[\hat{\lambda}_1 \geq \hat{\lambda}_2 \ldots \geq \hat{\lambda}_M]$ 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 overparametrization of the model.

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

$$AIC(W) = -2\log\left(\frac{\prod_{m=W+1}^{M}\hat{\lambda}_m^{1/(M-W)}}{\frac{1}{M-W}\sum_{m=W+1}^{M}\hat{\lambda}_m}\right)^{(M-W)T} + 2W(2M-W). \tag{5.13}$$

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

$$MDL(W) = -2\log\left(\frac{\prod_{m=W+1}^{M}\hat{\lambda}_m^{1/(M-W)}}{\frac{1}{M-W}\sum_{m=W+1}^{M}\hat{\lambda}_m}\right)^{(M-W)T} + W(2M-W)\log(T). \tag{5.14}$$

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.

5.2.3 Estimation of the wavefront ray parameters

Property (2) in equation (5.11) ensures that the data steering vectors $\mathbf{a}(\theta_w)$ 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. Further, a set of distinct ray 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, $\hat{\mathbf{E}}_s = [\hat{\mathbf{E}}_1, \hat{\mathbf{E}}_2, ..., \hat{\mathbf{E}}_W]$, and $\hat{\mathbf{E}}_n = [\hat{\mathbf{E}}_{W+1}, \hat{\mathbf{E}}_{W+2}, ..., \hat{\mathbf{E}}_M]$, where the $\hat{\mathbf{E}}_m$ are the eigenvectors of $\hat{\mathbf{R}}_d$ (Anderson, 1963). The ray-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 ray-parameter spectrum are:

$$P_{s_1}(\theta) = \sum_{m=1}^{W} \mid \mathbf{a}(\theta)^H \hat{\mathbf{E}}_m \mid^2, \tag{5.15}$$

and

$$P_{s_2}(\theta) = \frac{1}{1 - \sum_{m=1}^{W} \mid \mathbf{a}(\theta)^H \hat{\mathbf{E}}_m \mid^2} \tag{5.16}$$

when the signal subspace is used, or

$$P_{n_1}(\theta) = \frac{1}{\sum_{m=W+1}^{M} \mid \mathbf{a}(\theta)^H \hat{\mathbf{E}}_m \mid^2} \tag{5.17}$$

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 of the others, and therefore is equivalent for the estimation

of the ray parameter θ.

Another useful ray-parameter spectrum is

$$P_{n_2}(\theta) = \frac{1}{\sum_{m=W+1}^{M} \frac{1}{\hat{\lambda}_m} \mid \mathbf{a}(\theta)^H \hat{\mathbf{E}}_m \mid^2}. \tag{5.18}$$

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.

5.2.4 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 (5.10) and (5.11) do not hold true any more. 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 5.3. The covariance matrix $\hat{\mathbf{R}}_d^k$ can be defined for each subarray k, and the *spatially smoothed* covariance matrix is defined as

$$\hat{\mathbf{R}}_d^K = \frac{1}{K} \sum_{k=1}^{K} \hat{\mathbf{R}}_d^k. \tag{5.19}$$

Theoretically, use of the spatial smoothing technique can be justified only for planar wavefronts [equation (5.2)]. However, if the number of subarrays K is small the spatial smoothing technique can be successfully applied to spherical wavefronts as well [equations (5.3) and (5.4)].

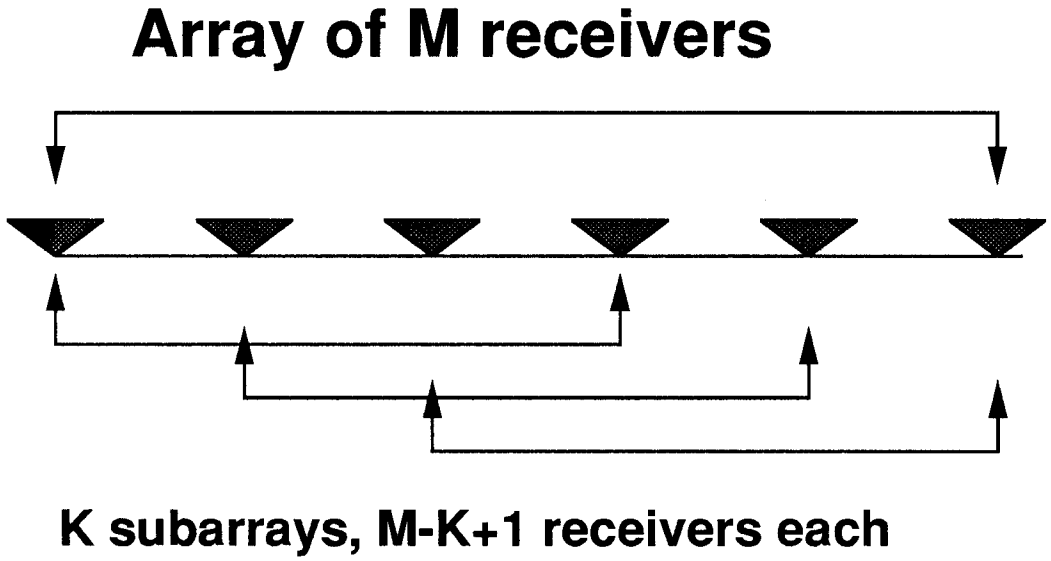

FIG. 5.3. Subdivision of the original array of M receivers into K subarrays of $M-K+1$ receivers for application of "spatial smoothing".

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

5.3 COMPARISON WITH THE STACKING METHOD

For narrow-band data, the ray-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

$$P_{Stack}(\theta) = \frac{1}{T}\sum_{t=1}^{T}|\frac{1}{M}\sum_{m=1}^{M}\text{Data}\,[t+\text{T}(x,\theta),x]\,|^2. \tag{5.20}$$

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

$$P_{Stack}(\theta) = \frac{1}{T}||\mathbf{a}^H(\theta)\mathbf{D}||^2_{L_2} = \frac{1}{T}[\mathbf{a}^H(\theta)\mathbf{D}]\ \ [\mathbf{a}^H(\theta)\mathbf{D}]^H = \tag{5.21}$$

$$\frac{1}{T}\mathbf{a}^H(\theta)[\mathbf{D}\mathbf{D}^H]\mathbf{a}(\theta) = \mathbf{a}(\theta)^H\,\hat{\mathbf{R}}_d\,\mathbf{a}(\theta). \tag{5.22}$$

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

$$P_{Stack}(\theta) = \mathbf{a}(\theta)^H\,\hat{\mathbf{R}}_d\,\mathbf{a}(\theta) = \sum_{m=1}^{M} \hat{\lambda}_m |\, \mathbf{a}(\theta)^H\hat{\mathbf{E}}_m\,|^2. \tag{5.23}$$

The stacking spectrum, equation (5.23), differs from the eigenstructure spectrum, equation (5.15), in two ways. First, and most important, is the weighting by eigenvalues in the summation in equation (5.23). The weights in equation (5.23) reward those test steering vectors that are close to eigenvectors corresponding to the largest eigenvalues. On the contrary the spectrum in equation (5.15) is a projection of the test steering vectors onto the signal subspace; it does not privilege any direction in this subspace. Second, the ranges of the summations differ. In the stacking method, the summation is carried over all the eigenvectors. In the eigenstructure method, it 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.

5.3.1 Geometric interpretation

The superior resolution achieved by the eigenstructure spectrum can be easily explained by geometric considerations, as shown in Figure 5.4. I 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 $\mathbf{\Omega}_M$ spanned by the M orthonormal steering vectors $[\mathbf{a}(\theta = 0),\ \mathbf{a}(\theta = \pm\Delta\theta),\ \mathbf{a}(\theta = \pm 2\Delta\theta), \ldots,\ \mathbf{a}(\theta = +\frac{M}{2}\Delta\theta)]$, 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 $\mathbf{\Omega}_M$. This manifold intersects the signal subspace only in $\mathbf{a}(\theta_1)$ and $\mathbf{a}(\theta_2)$.

The eigenstructure spectrum in equation (5.15) is a projection of the test steering vectors onto the signal subspace; it thus has two maxima exactly at θ_1 and θ_2, independently 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

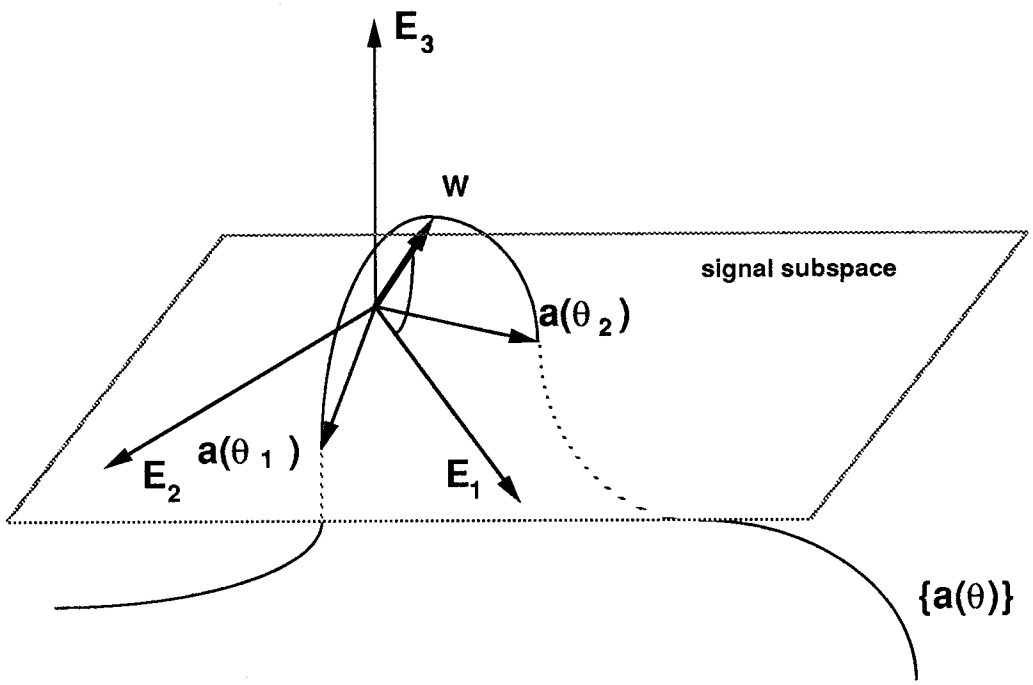

FIG. 5.4. Geometric interpretation of the stacking spectrum and of the eigenstructure spectrum for two incident plane waves. The first two eigenvectors, $\mathbf{E}_1$ and $\mathbf{E}_2$, define the "signal subspace". The data steering vectors, $\mathbf{a}(\theta_1)$ and $\mathbf{a}(\theta_2)$, lie in 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 $\mathbf{a}(\theta_2)$.

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),

$$\cos\alpha = \mathbf{a}^H(\theta_1)\mathbf{a}(\theta_2) = \frac{\sin(\omega\Delta x M[(\theta_2-\theta_1)/2])}{M\sin(\omega\Delta x[(\theta_2-\theta_1)/2])} = \frac{\sin(2M\beta)}{M\sin(2\beta)}, \tag{5.24}$$

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

The eigenvalues of the data covariance matrix are

$$\lambda_1 = S(1+\cos\alpha), \quad \lambda_2 = S(1-\cos\alpha), \quad \text{and } \lambda_m = 0 \quad \text{for } m > 2. \tag{5.25}$$

The difference between the first two eigenvalues increases as the angle between the data

steering vectors decreases. The first two eigenvectors are

$$E_1 = \frac{\mathbf{a}(\theta_1) + \mathbf{a}(\theta_2)}{\sqrt{2 + 2\cos\alpha}} \quad \text{and} \quad \mathrm{E}_2 = \frac{\mathbf{a}(\theta_1) - \mathbf{a}(\theta_2)}{\sqrt{2 - 2\cos\alpha}}. \tag{5.26}$$

The test steering vector $\mathbf{w} = \mathbf{a}[(\theta_1 + \theta_2)/2]$ is closer to the first eigenvector than either $\mathbf{a}(\theta_1)$ or $\mathbf{a}(\theta_2)$ are. Therefore, the smaller the angle α between the data steering vectors, the larger the weight given by the stacking method [equation (5.23)] 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 ray-parameter equal to $(\theta_1 + \theta_2)/2$.

When the two plane waves are correlated in time, 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 sources covariance matrix $\mathbf{R}_s$, is real, the eigenvalues of the data covariance matrix are:

$$\lambda_1 = (S + S_{12})(1 + \cos\alpha), \quad \lambda_2 = (S - S_{12})(1 - \cos\alpha), \quad \text{and} \quad \lambda_m = 0 \text{ for } m > 2, \tag{5.27}$$

while the eigenvectors are the same as they are for uncorrelated sources [equation (5.26)]. 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 5.5 shows a stacking spectrum (solid line) superimposed upon an eigenstructure spectrum (dashed line) computed using equation (5.16). The two spectra are normalized to one. The data are two monochromatic plane waves with frequency 25 Hz, and ray parameters $\theta_1 = .2$ s/Km and $\theta_2 = .25$ s/Km. The array has 40 receivers spaced 10 m, 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 5.6 compares the spectra when the sources are fully correlated. In this case neither method resolves the two correlated signals (upper Figure 5.6). On the other hand, after spatial smoothing is applied, as presented in section 5.2.4, the eigenstructure method resolves the two wavefronts (lower Figure 5.6). This figure compares the spectra obtained

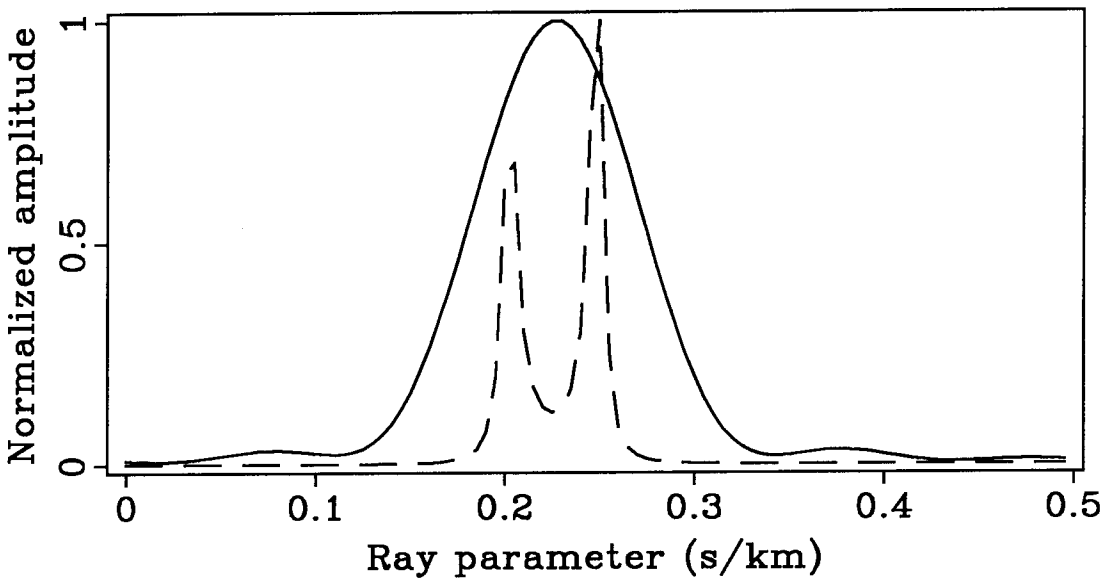

FIG. 5.5. Stacking spectrum (solid line) and eigenstructure spectrum (dashed line) for two monochromatic plane waves. The ray parameters are $\theta_1 = .2$ s/Km and $\theta_2 = .25$ s/Km. The angle α between the two data steering vectors is 50°. Only the eigenstructure method resolves the two wavefronts.

when spatial smoothing is applied by dividing the original array in 11 subarrays. The spatially smoothed covariance matrix is used for the computation of both spectra.

In this section I considered the simple case of two plane waves and noiseless data. I did not consider the effect of the data's length limitations on the quality of the estimate of the data covariance matrix. The reader interested in a more general discussion of the statistical performances of eigenstructure method, will find relevant material in Kaveh and Barabell (1986) and Wang and Kaveh (1986).

5.4 THE WIDE-BAND METHOD

The eigenstructure method described in the previous section can be applied to narrow-band data. Seismic data are wide-band, but they can 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 ray-parameters remain the same for all frequencies. The frequency components can be combined at different stages of the processing: averages can

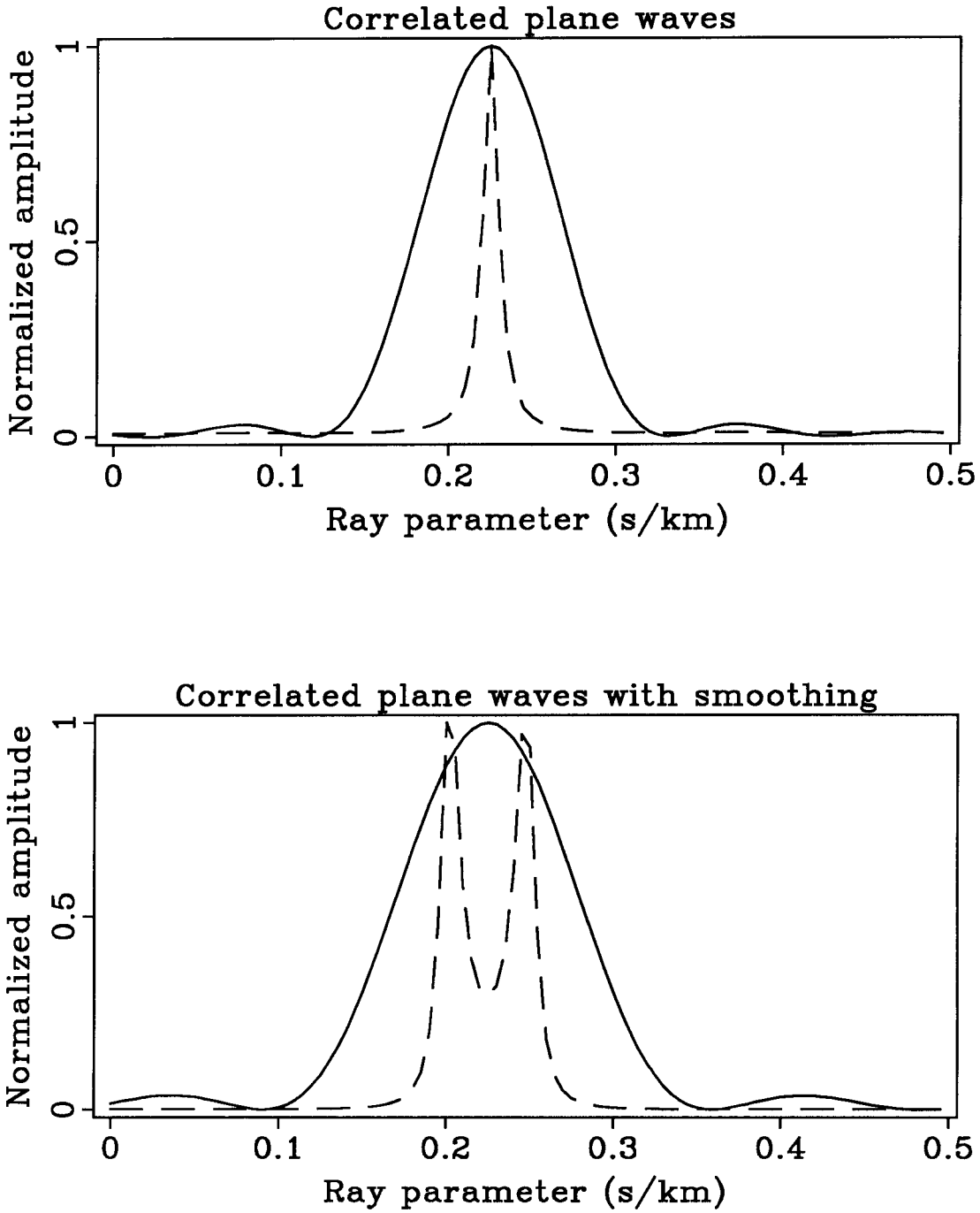

FIG. 5.6. Stacking spectrum (solid line) and eigenstructure spectrum (dashed line) for two correlated monochromatic plane waves. When the plane waves are fully correlated neither method resolves the two wavefronts (upper figure). The eigenstructure spectrum resolves the two wavefronts when spatial smoothing is applied (lower figure).

be made of either the final estimates of the spectra (Wax et al., 1984), or the estimates of the covariance matrices of the different components. (Wang and Kaveh, 1985). In the latter case, the frequency components must be linearly 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 my computations, I combine both methods described above: I average correlation matrices for nearby frequencies, and average spectra from different frequency bands.

First, the data are time-corrected according to the moveout $\mathrm{T}(x,\bar{\theta})$, where $\bar{\theta}$ is an estimate of the ray 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 ray 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 ray parameter $\theta + d\theta$, is

$$\Delta\phi = (\omega_1 - \omega_2)\mathrm{T}(x, \bar{\theta} + d\theta). \tag{5.28}$$

After the time correction, the same phase difference becomes

$$\Delta\phi = (\omega_1 - \omega_2)\mathrm{T}(x, d\theta). \tag{5.29}$$

A bank of band-pass filters are used to decompose the time corrected data into a few, thus 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, I select the value that minimizes the sum of the MDL criteria [equation (5.14)] for each of the F

frequency bands,

$$\sum_{f=1}^{F} MDL_f(W) = \sum_{f=1}^{F} \left[-2\log \left(\frac{\prod_{m=W+1}^{M} \hat{\lambda}_m(f)^{1/(M-W)}}{\frac{1}{M-W}\sum_{m=W+1}^{M} \hat{\lambda}_m(f)} \right)^{(M-W)T} + W(2M-W)\log(T) \right]. \tag{5.30}$$

Once the number of wavefronts W is determined, the parameter spectra of all the bands are averaged to obtain the ray parameter spectrum. For example, the wide-band version of equation (5.15) is

$$P_{s1}(\theta) = \frac{1}{F}\sum_{f=1}^{F}\sum_{m=1}^{W} \mid \mathbf{a}(\theta,f)^H \hat{\mathbf{E}}_m(f) \mid^2. \tag{5.31}$$

If the signal to noise ratio (S/N) in each band is known, the terms summed in equation (5.30) and (5.31) could be weighted accordingly so as to increase the quality of the estimates.

The estimation of the covariance matrix from noisy data 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 ray parameters significantly different from $\bar{\theta}$, one can perform spatial dip filtering in conjunction with partial stacking.

5.5 APPLICATION TO LOCAL SLANT STACKS

The advantages of the eigenstructure method are confirmed by its application to the marine CMP gather shown in Figure 5.1. The data sampling rate is 2 ms in time and 100 m in the offset dimension. I estimate local spectra at the intersection of a primary reflection and a water-bottom multiple. In Figure 5.1 the region of interference between the two reflections is indicated by a balloon.

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 trade-off between ray-parameter

resolution and spatial resolution of the local estimates, I 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 5.7 shows the normalized ray-parameter spectra estimated from the original gather with frequencies from 15 Hz to 100 Hz. The spectra shown in the lower figure are estimated with a subarray of 8 geophones, centered at the intersection of the two reflections; while the spectra shown in upper figure are estimated with a subarray of 6 geophones. As reference values I used the reflections' ray parameters as estimated by hand from a plot of the gather (dotted line); even if carefully measured, these values are affected by some error. To compute the eigenstructure spectrum (dashed line), I first applied a time correction corresponding to the ray parameter $\bar{\theta}$=.32 s/Km, then I decomposed the data into 10 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, independently of the subarray length. The maxima 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, I compute the spectra after having band-passed the gather with different high-frequency cut-off values. Figure 5.8 shows the local spectra computed from a gather with frequencies of 15 Hz to 70 Hz. The eigenstructure spectrum is computed after a time correction, corresponding to the ray parameter $\bar{\theta}$=.32 s/Km, is applied, and after the data are decomposed into 6 frequency bands. For this bandwidth, the stacking spectrum (solid line in the lower figure) fails to resolve the two reflections when a subarray of 6 geophones is used. The eigenstructure spectrum resolves both reflections well, independently of the subarray length.

Figure 5.9 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 $\bar{\theta}$=.32 s/Km, is applied, and after the data is decomposed into 4 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 6 geophones is used.

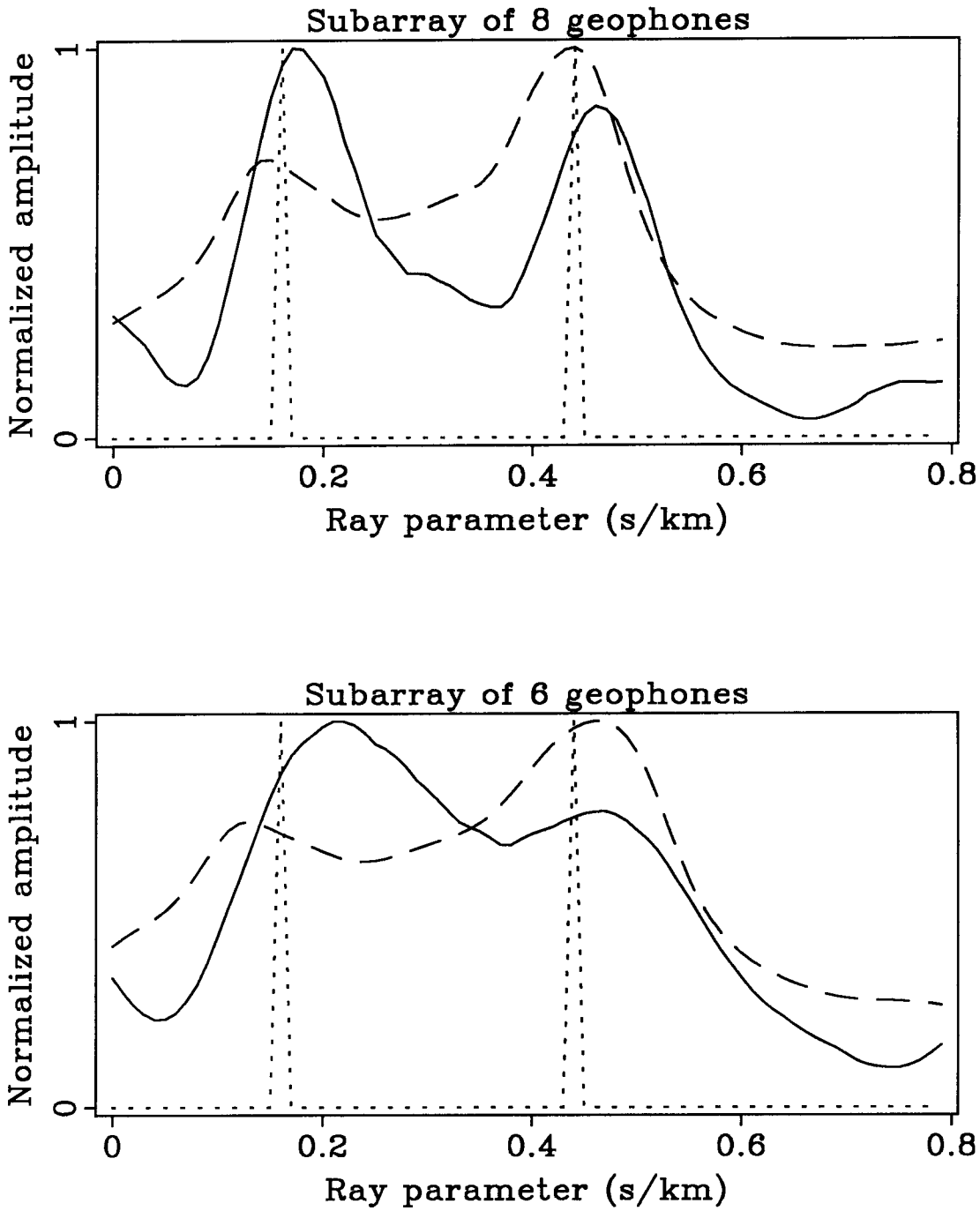

FIG. 5.7. Normalized, local ray parameter spectra. The dashed lines are the eigenstructure spectra; the solid lines are the stacking spectra. The dotted lines show two spikes at the values of the ray parameter estimated by hand. The data have frequencies from 15 Hz to 100 Hz. A subarray of 8 geophones is used for the upper figure and a subarray of 6 geophones is used for the lower figure. The maxima of the spectra agree well with the ray-parameter values picked by hand.

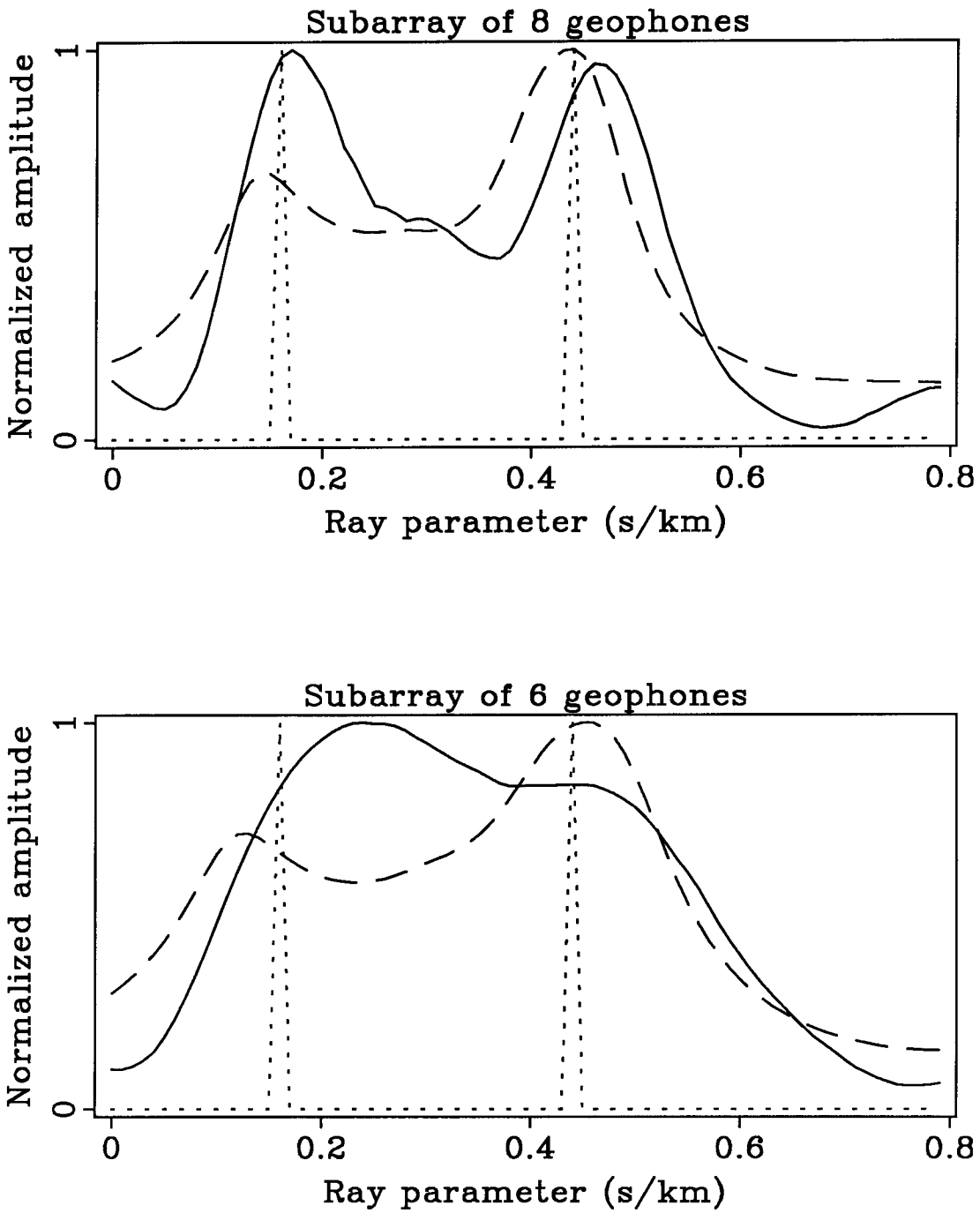

FIG. 5.8. The stacking spectra (solid lines) and the eigenstructure spectra (dashed lines) computed when the data frequencies extend from 15 Hz to 70 Hz. The dotted lines show two spikes at the values of the ray parameter estimated by hand. A subarray of 8 geophones is used for the upper figure and a subarray of 6 geophones is used for the lower figure. The stacking spectrum cannot resolve the two reflections when the shorter subarray is used (lower figure).

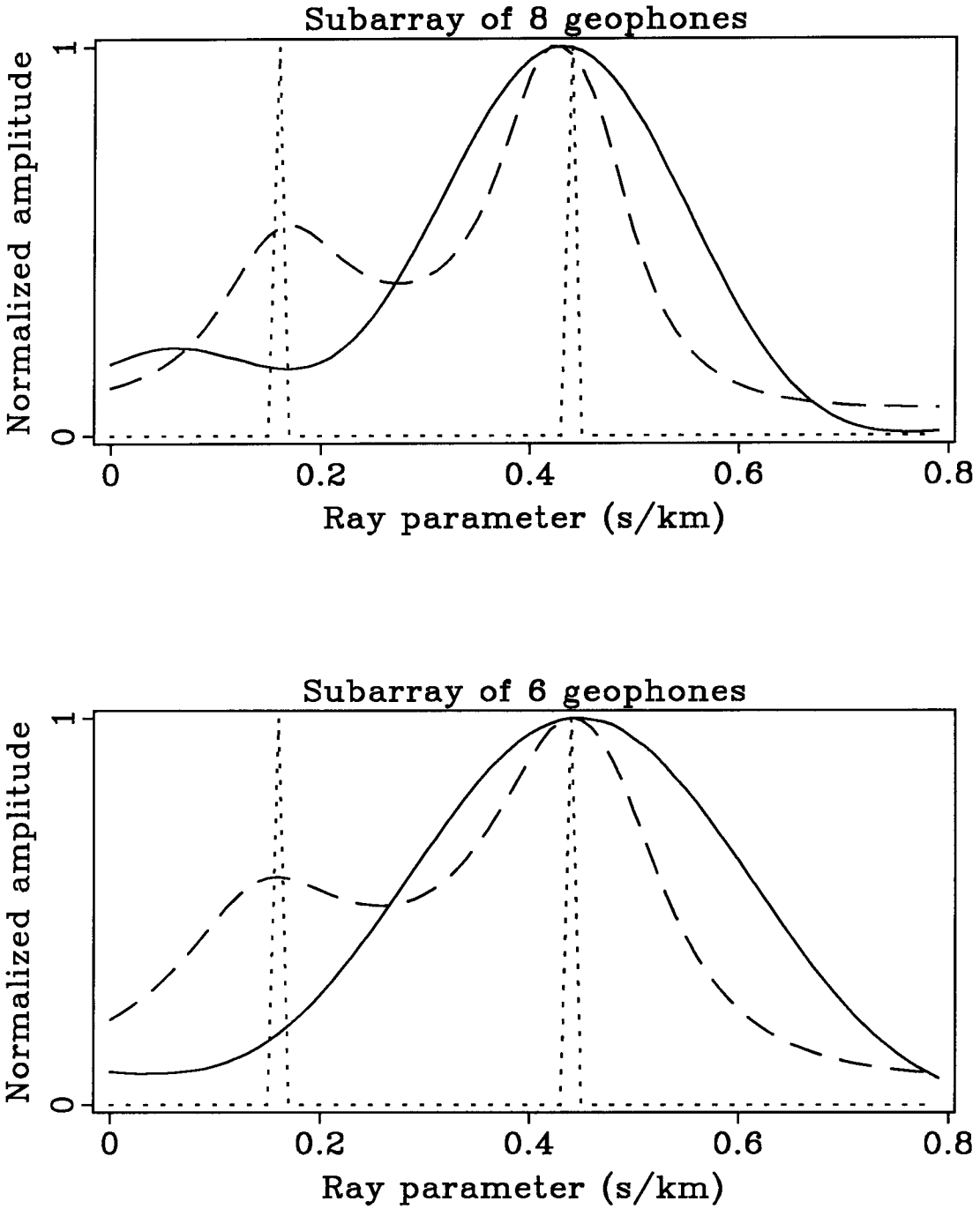

FIG. 5.9. The stacking spectra (solid lines) and the eigenstructure spectra (dashed lines) computed when the data frequencies extend from 15 Hz to 33 Hz. The dotted lines show two spikes at the values of the ray parameter estimated by hand. A subarray of 8 geophones is used for the upper figure and a subarray of 6 geophones is used for the lower figure. The stacking spectrum fails to resolve the reflections with either subarray.

5.6 CONCLUSIONS

The methods for spectral estimation presented in this chapter 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.

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 data samples in time. Taking advantage of the full bandwidth of the seismic data, I 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 has to 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.

The field-data example demonstrates the gain in lateral resolution that can be achieved when the estimation of local spectra uses the eigenstructure method.