Eigenstructure Variability of the Multiple-Source Multiple-Sensor Covariance Matrix with Contaminated Gaussian Data

ALIREZA MOGHADDAMJOO, MEMBER, IEEE

Abstract—**Several methods of current interest for counting and locating signal sources using data from a passive array depend on the accuracy of estimating the eigenstructure of the covariance matrix of the array's data vectors. When errors in the measured data vectors are Gaussian, conventional covariance estimation is optimal, but robust procedures are required for data with non-Gaussian additive contamination. Two different robust covariance estimators are compared by simulation with the conventional one for different degrees of contamination. Even in relatively good signal-to-noise ratios, however, closeness of signal sources in temporal, spatial frequency domain can cause inaccurate signal-related eigenvalue and eigenvector estimates. The degree of adversity for these problems is also shown by simulation.**

I. Introduction

In radar, sonar, and seismology, we are interested in estimating the directions of arrival and the spectral densities of radiating sources from measurements provided by a passive array of sensors. In this study we consider estimation of the eigenstructure of the covariance matrix of the received signal vectors. The utility of the eigenstructure has been detailed in a number of recent papers, for example, Wax *et al.* [1], based on original work by Schmidt [2]. Theoretical discussion on the variability of the eigenstructure is given in [3]–[5], and considerable theoretical work is being done on asymptotic statistics of the eigenstructure for non-Gaussian error by Krishnaiah *et al.* [6]. The work by Kaveh and Barabell [7], [8] is one of the original works relating asymptotic statistics of the eigenstructure to the eigenassisted methods in resolving plane waves. Eigenstructure variability is an important point; we relate it to the method of estimation of the covariance matrix.

Three different covariance matrix estimation methods are considered in this paper. The first method is the conventional sample covariance estimation; the second one is robust, based on rank correlation; and the last one is again robust, but based on the weighted M-estimate [9]. The last two methods are described in the Appendix. The reason for considering robust methods is that real world data are often contaminated with noise densities which are non-Gaussian and heavily tailed. Other robust schemes for eigenstructure estimation have been reported in [10]. Therein, the vectors are estimated directly (probably a method superior to ours), but the covariance matrices considered were not specifically related to the multiple source problem. Herein, we try techniques which have the covariance matrix as an intermediate quantity and use practical source configurations to illustrate the variability problems by simulation.

We find that in common situations the spread of signal-related eigenvalues is significant. By varying spatial frequencies, we show that it may often be difficult to separate signal (large) and noise (small) eigenvalues. Also we show that in some special cases, where signal eigenvalues are close to each other, the space spanned by signal related eigenvectors would not be substantially affected, although the variation of each individual signal eigenvector is very high.

II. Background and Simulation Specification

The number of sensors m is chosen to be 3, and they are equally spaced in line. For each target (signal source) configuration, 500 simulation samples of the resulting eigenvalues and eigenvectors have been found; each simulation sample corresponds to one N-sample data set and covariance estimate. The 500 samples of eigenvalues $\hat{\lambda}_j$ and eigenvectors $\hat{v}_j$ thus provide information of their variability under the conditions of the simulation. The variability may be measured and displayed by various means, such as by the normalized sample biases and variances of $\hat{\lambda}_j$ and by the magnitude of the difference between $\hat{v}_j$ and its true value, v_j, or the correlation of $\hat{v}_j$ with its true value. Selection of these parameters, especially those which correspond to the eigenvectors, should be based on their effectiveness in representing the actual variability of the estimated eigenstructure. The asymptotic statistics for the eigenvalues and eigenvectors of the sample covariance matrix for a complex Gaussian process, as derived in [4] and [7], are

$$E[\hat{\lambda}_j] = \lambda_j + O(N^{-1}) \tag{1}$$

$$E\left[(\hat{\lambda}_j - E[\hat{\lambda}_j])^2\right] = \frac{2\lambda_j^2}{N} + O(N^{-2}) \tag{2}$$

Manuscript received July 12, 1986; revised August 26, 1987.

The author is with the Department of Electrical Engineering and Computer Science, University of Wisconsin-Milwaukee, Milwaukee, WI 53201.

IEEE Log Number 8718018.

$$E[\hat{v}_i] = v_i + O(N^{-1}) \tag{3}$$

$$E\left[(\hat{v}_i - E[\hat{v}_i])^H (\hat{v}_i - E[\hat{v}_i])\right] = \frac{\lambda_i}{N} \sum_{\substack{k=1 \\ k \neq 1}}^{l} \frac{\lambda_k v_i^H v_i}{(\lambda_i - \lambda_k)^2} + O(N^{-2}). \tag{4}$$

Equations (1) and (2) are rearranged as follows:

$$E[\hat{\lambda} - \lambda_j] = O(N^{-1}) \tag{5}$$

$$E\left[\left\{(\hat{\lambda}_j - \lambda_j) - E[\hat{\lambda}_j - \lambda_j]\right\}^2\right] = E[(\hat{\lambda}_j - \lambda_j)^2] - (E[\hat{\lambda}_j - \lambda_j])^2 = \frac{2\lambda_j^2}{N} + O(N^{-2}). \tag{6}$$

By substituting (5) into (6), then solving for $E[(\hat{\lambda}_j - \lambda_j)^2]$ and dividing the result by λ_j^2, we obtain

$$E\left[\left(\frac{\hat{\lambda}_j - \lambda_j}{\lambda_j}\right)^2\right] = \frac{2}{N} + O(N^{-2}). \tag{7}$$

Since the first-order term is preserved in (7), the normalized difference of each eigenvalue from its true value would be a good measure of its variability. Similarly, we rearrange (3) and (4) as follows:

$$E[\hat{v}_i - v_i] = O(N^{-1}) \tag{8}$$

$$E\left[\left\{(\hat{v}_i - v_i) - E[\hat{v}_i - v_i]\right\}^H \cdot \left\{(\hat{v}_i - v_i) - E[\hat{v}_i - v_i]\right\}\right] = E\left[(\hat{v}_i - v_i)^H (\hat{v}_i - v_i)\right] - \left(E[(\hat{v}_i - v_i)]\right)^H \left(E[(\hat{v}_i - v_i)]\right) = \frac{\lambda_i}{N} \sum_{\substack{k=1 \\ k \neq 1}}^{l} \frac{\lambda_k v_i^H v_i}{(\lambda_i - \lambda_k)^2} + O(N^{-2}). \tag{9}$$

By substituting (8) into (9) we obtain

$$E\left[(\hat{v}_i - v_i)^H (\hat{v}_i - v_i)\right] = E[\|\hat{v}_i - v_i\|^2] = \frac{\lambda_i}{N} \sum_{\substack{k=1 \\ k \neq 1}}^{l} \frac{\lambda_k}{(\lambda_i - \lambda_k)^2} + O(N^{-2}) \tag{10}$$

where l is eigenvector dimension, $(\cdot)^H$ denotes Hermitian, and all eigenvectors are normalized in a way that $\hat{v}_i^H \hat{v}_i = 1$, $v_i^H v_i = 1$ for $i = 1, 2, \cdots, l$. Equation (10) shows that in $E[\|\hat{v}_i - v_i\|^2]$, the first-order term is preserved and $\|\hat{v}_i - v_i\|^2$ would be a good parameter to use in order to show the individual eigenvector variability. It is common to use the correlation of the eigenvectors $\hat{v}_i$ with their true values v_i in order to show its variability. This will work as well as $\|\hat{v}_i - v_i\|$ in the case of real data, but for complex data the variability of this correlation, which is defined as

$$\cos \alpha_i = |\hat{v}_i^H \cdot v_i|, \tag{11}$$

would not represent the variability of the corresponding eigenvector. To show this fact, we write (11) in the following form:

$$\cos^2 \alpha_i = [(\hat{v}_i^H \cdot v_i)(\hat{v}_i^H \cdot v_i)^*] = (\hat{v}_i^H \cdot v_i \cdot \hat{v}_i^T \cdot v_i^*) \tag{12}$$

where $(\cdot)^*$ denotes conjugate of its argument. Let $\delta_i = \hat{v}_i - v_i$. By substitution of $\hat{v}_i = v_i + \delta_i$, we obtain

$$\cos^2 \alpha_i = (v_i + \delta_i)^H v_i (v_i + \delta_i)^T \cdot v_i^* = (v_i^H v_i + \delta_i^H v_i)(v_i^T v_i^* + \delta_i^T v_i^*). \tag{13}$$

Since $v_i^H v_i = (v_i^T v_i^*)^T = v_i^T v_i^* = 1$ we can simplify (13) further:

$$\cos^2 \alpha_i = (1 + \delta_i^H v_i)(1 + \delta_i^T v_i^*) = 1 + \delta_i^H v_i + \delta_i^T v_i^* + \delta_i^H v_i \delta_i^T v_i^* = 1 + \delta_i^H v_i + (\delta_i^H v_i)^* + \delta_i^H v_i (v_i^H \delta_i)^T. \tag{14}$$

Since $v_i^H \delta_i$ is a complex number, its transpose is equal to itself, and hence,

$$\cos^2 \alpha_i = 1 + 2 \text{ Real } [\delta_i^H v_i] + (v_i^H \delta_i \delta_i^H v_i)^*. \tag{15}$$

Using the following result in [7]:

$$E[\delta_i] = -\frac{\lambda_i}{2N} \sum_{\substack{k=1 \\ k \neq i}}^{l} \frac{\lambda_k}{(\lambda_i - \lambda_k)^2} v_i + O(N^{-2}) \tag{16}$$

and (10) we obtain

$$E[\cos^2 \alpha_i] = 1 - \frac{\lambda_i}{N} \sum_{\substack{k=1 \\ k \neq i}}^{l} \frac{\lambda_k \text{ Real } (v_i^H v_i)}{(\lambda_i - \lambda_k)^2} + \frac{\lambda_i}{N} \sum_{\substack{k=1 \\ k \neq i}}^{l} \frac{\lambda_k}{(\lambda_i - \lambda_k)^2} + O(N^{-2}). \tag{17}$$

Since $v_i^H v_i = 1$, (17) simplifies to

$$E[\cos^2 \alpha_i] = 1 + O(N^{-2}). \tag{18}$$

In this equation, the term proportional to $1/N$, which has the major effect on the variability of eigenvector, is cancelled and, hence, cos α would not be a good parameter to use for complex data.

The parameters $e(\hat{\lambda}_i)$ and $e(\hat{v}_i)$ defined as

$$e(\hat{\lambda}_i) = (\hat{\lambda}_i - \lambda_i)/\lambda_i \tag{19}$$

and

$$e(\hat{v}_i) = (\hat{v}_i - v_i)^H (\hat{v}_i - v_i) \tag{20}$$

are used to display the eigenstructure variability. Estimates of the probability distribution, histogram, and (MSE) are also used. In order to show the variability of the estimated signal space (space spanned with signal related eigenvectors), especially in the cases that signal eigenvectors are mixing, the error of the map of $\hat{v}_i$ onto the

true signal space is also considered

$$e[\text{map}(\hat{v}_i)] = \left\| \hat{v}_i - \sum_{v_j \in \text{true signal eigenvectors}} v_j^H \hat{v}_i v_j \right\| \quad (21)$$

where

$$\|\cdot\| = \sqrt{(\cdot)^H(\cdot)}.$$

As it is a major intent to gauge the eigenstructure variability in the face of contaminated Gaussian noise (heavily tailed), the simulated, independent, white noise samples have density

$$f(n_i) = (1 - \epsilon)\, G(0, \sigma_1^2) + \epsilon G(0, \sigma_2^2) \quad (22)$$

where $G(0, \sigma^2)$ is the Gaussian density with mean zero and variance σ^2, and ϵ is the proportion of contaminating samples at level σ_2. We always use $\sigma_2^2 = 9 \cdot \sigma_1^2$ in our simulations.

Different numbers of signal sources with different temporal and spatial frequencies are used. The source signals are narrow-band Gaussian, thus, the nth sample of the received process at the qth tap of the ith sensor is

$$r_{iq}(nT) = \sum_{k=1}^{K} p_k(nT) \exp\left[-j2\pi(\eta_k q + \gamma_k i)\right] + n_{iq}(nT)$$
$$i = 0, 1, \cdots, m-1$$
$$q = 0, 1, \cdots, h-1 \quad (23)$$

where K is the number of sources, and $p_k(nT)$ are the complex amplitudes of the plane waves with $E[|p_k(nT)|^2] = P_k$. These amplitudes are jointly circular Gaussian and jointly independent and independent of n_i. T is the sampling period which is much larger than the time delay between taps. η_k are normalized temporal frequencies.

$$\eta_k = f_k T \quad (24)$$

where f_k are centered frequencies of the plane waves, γ_k are the normalized spatial frequencies which relates directly to the direction of arrivals α_k

$$\gamma_k = \frac{Df_k}{C} \sin \alpha_k, \quad (25)$$

C is the velocity of propagation, D is the sensor spacing, and α_k is the angle of incident, with respect to broadside, for the kth source, n_{iq} is a complex, zero mean, circular vector with independent elements and contaminated Gaussian distribution with ϵ degrees of contamination. m is the total number of sensors and h is the total number of delays (including zero delay) after each sensor.

The covariance matrix with this configuration is

$$R = E[r(nT)\, r^H(nT)] = ASA^H + \sigma_n^2 I \quad (26)$$

where

$$r(nT) = [r_{00}(nT), r_{01}(nT), \cdots, r_{(m-1),(h-1)}(nT)]^T \quad (27)$$

$$\sigma_n^2 = E[|n_i(nT)|^2] \quad i = 0, 1, \cdots, m-1. \quad (28)$$

S is the covariance matrix of the source signal vector

$$S = \begin{bmatrix} P_1 & & & \\ & P_2 & \cdots & 0 \\ & & & \\ 0 & & & P_k \end{bmatrix} \quad (29)$$

$$A = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ \exp(-j2\pi\eta_1) & \exp(-j2\pi\eta_2) & \cdots & \exp(-j2\pi\eta_k) \\ \exp[-j2\pi(2\eta_1)] & \exp[-j2\pi(2\eta_2)] & \cdots & \exp[-j2\pi(2\eta_k)] \\ \vdots & & & \\ \exp\{-j2\pi[\eta_1(h-1) + \gamma_1(m-1)]\} & \exp\{-j2\pi[\eta_2(h-1) + \gamma_2(m-1)]\} & \cdots & \exp\{-j2\pi[\eta_k(h-1) + \gamma_k(m-1)]\} \end{bmatrix} \quad (30)$$

and I is an $m \cdot h$ by $m \cdot h$ identity matrix.

We use $N = 100$ for producing each sample of R, and the sampling intervals in both space and time would be appropriate as long as

$$0 < \eta_k = f_k \cdot T < 0.5 \quad (31)$$

and

$$-0.5 < \gamma_k = \frac{Df_k}{C} \sin \alpha_k < 0.5. \quad (32)$$

For generation of noise in different cases, we wrote our own codes and used IMSL Math/Library subroutine for generation of independent Gaussian random number sequences. Table I displays the above variables for each simulation.

III. Simulation Results

Simulations are based on 3 sensors each followed by 2 delay taps yielding 9×9 covariance matrix. In this study we compare three different methods of covariance matrix estimation by calculating histogram and CDF's (Cumulative Distribution Function) of $\hat{\lambda}_i$, $e(\hat{\lambda}_i)$ in (19), $e(\hat{v}_i)$

TABLE I
SIMULATION VARIABLES

Number of Sources	K:	1, 2, and 3
Number of Sensors	m:	3
Number of Taps/Sensor	h:	3
Contamination Fraction	ϵ:	0.0, 0.05, and 0.1[a]
Number of Samples/R Element	N:	100
Number of R Samples	L:	500
Uncontaminated SNR[b]	SNR:	2, 4, and 8
R-Element Estimators in Plots	Solid Line ——	(Conventional, Method A)
	Dash Pattern --------	(Rank Correlation, Method B)
	Dash Pattern ---	(Weighted M-Estimate, Method C)

[a]The variance of the contamination noise σ_2^2 is 9 times variance of uncontaminated noise σ_1^2. In this case, the overall variance of noise is $\sigma_n^2 = (1 - \epsilon)\,\sigma_1^2 + \epsilon\sigma_2^2$.
[b]This is SNR when we set $\epsilon = 0$ (when we assume noise variance is σ_1^2).

TABLE II
MSE COMPARISON

Parameter	MSE (Method A)	MSE (Method B)	MSE (Method C)	MSE (Theory)
$e(\hat{\lambda}_1)$	1.237×10^{-2}	1.67×10^{-2}	2.1×10^{-2}	0.02
$e(\hat{\lambda}_2)$	8.631×10^{-3}	1.476×10^{-2}	1.141×10^{-2}	0.02
$e(\hat{\lambda}_3)$	1.523×10^{-2}	1.973×10^{-2}	1.636×10^{-2}	0.02
$e(\hat{\lambda}_{\text{noise}})$	0.307	0.191	0.316	—
$e(\hat{v}_1)$	8.354×10^{-2}	0.123	0.147	0.114
$e(\hat{v}_2)$	0.572	0.656	0.579	0.354
$e(\hat{v}_3)$	0.142	0.168	0.182	0.215
$e(\hat{v}_{\text{noise}})$	1.524	1.541	1.51	—

in (20), $e[\text{map}\,(\hat{v}_i)]$ in (21), $\hat{\eta}_k$, $\hat{\gamma}_k$, and $e(\hat{\beta}_k)$ defined as

$$e(\hat{\beta}_k) = \sqrt{(\hat{\eta}_k - \eta_k)^2 + (\hat{\gamma}_k - \gamma_k)^2}. \qquad (33)$$

Estimation of $\hat{\eta}_k$ and $\hat{\gamma}_k$ in these simulations is based on the MUSIC algorithm. All of these parameters have been calculated for the three covariance estimation methods: A—conventional, B—rank correlation, and C—weighted M-estimation.

For low contamination fraction, $\epsilon \leq 0.05$, Method A, the conventional, is always better than the robust methods, and Method C is better than Method B. But for $\epsilon = 0.1$ and any SNR (signal-to-noise ratio P/σ_n^2), Method B, the rank correlation, is better than Methods A and C, and A is better than C.

To compare mean square error (MSE) of $e(\hat{\lambda})$ and $e(\hat{v})$ to the predictions in (7) and (10), respectively, we conduct a simulation in which we chosed three equipowered signal sources with SNR = 10, $\eta_1 = \eta_2 = \eta_3 = \gamma_1 = 0.25$, $\gamma_2 = 0.0$, $\gamma_3 = -0.37$, and $\epsilon = 0.0$. True signal eigenvalues in this case are $\lambda_1 = 129$, $\lambda_2 = 83$, $\lambda_3 = 61$, and all noise eigenvalues are equal to 1. Calculated MSE's of $e(\hat{\lambda}_i)$ and $e(\hat{v}_i)$ are compared to the prediction in (7) and (10) in Table II.

Calculated MSE's are in agreement with the predictions in (7) and (10), except for $e(\hat{v}_2)$ where calculated MSE's are almost twice what we would expect from theory. In all cases, Method A shows lower MSE than the other two methods. Because it is difficult to interpret which of these methods yields the overall best result (just based on average bias, variance, and mean square error) over all situations, we have produced plots of estimates of distribution and density (histogram) functions of $\hat{\lambda}_i$, $e(\hat{\lambda}_i)$, $e(\hat{v}_i)$, $e[\text{map}\,(\hat{v}_i)]$, $\hat{\eta}_k$, $\hat{\gamma}_k$, and $e(\hat{\beta}_k)$, and we visualize the behavior of different procedures. In these plots, the scales of the abscissas are varied for plotting convenience.

Situation 1: Here we use one signal source with $\eta_1 = 0.35$, $\gamma_1 = 0.2$, and different SNR. Fig. 1 shows the CDF's (estimates of the probability distribution functions) of $e(\hat{\lambda}_1)$, $e(\hat{v}_1)$, and $e(\hat{\beta}_1)$, for SNR = 2 and $\epsilon = 0.0$. Fig. 2 shows the same CDF's for the case that SNR = 2 and $\epsilon = 0.1$. Fig. 3 shows the CDF's of $e(\hat{\lambda}_1)$, $e(\hat{v}_1)$, and $e(\hat{\beta}_1)$, for SNR = 8 and $\epsilon = 0.1$. Fig. 4 shows the same CDF's and histograms of $\hat{\eta}_1$ and $\hat{\gamma}_1$ for the case that SNR = 8 and $\epsilon = 0.05$. By careful study of these plots, we reach the following conclusion. For zero contamination fraction, $\epsilon = 0.0$, Method A and Method C are almost the same and always better than the robust Method B in estimating signal eigenvalues, eigenvectors, and frequencies; Method C is better than Method B in estimating signal eigenvalues and eigenvectors but not in signal frequencies. For low contamination fraction, $\epsilon = 0.05$, Methods A, B, and C are almost equal in estimating signal eigenvalues and eigenvectors, however, $\hat{\lambda}_1$ has negative bias in Method B, but very slight positive in Methods A and C. In this case for the signal frequencies [$\hat{\eta}_1$ and $\hat{\gamma}_1$ or $e(\hat{\beta}_1)$], Method B works better than Meth-

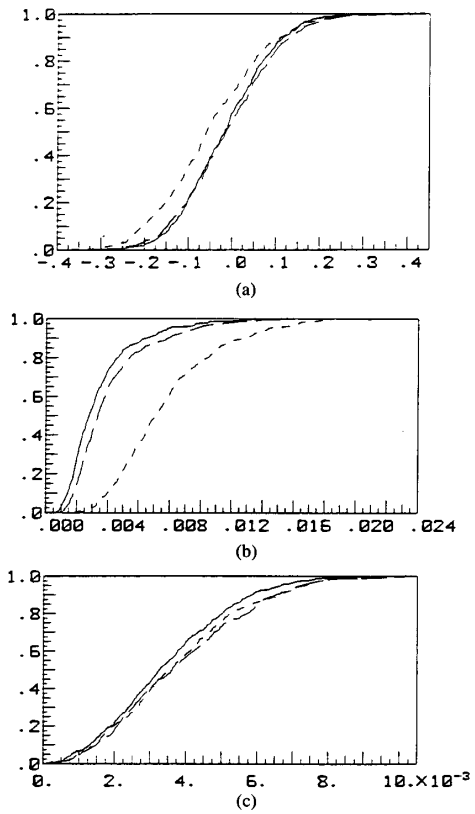


Fig. 1. One source with $\eta_1 = 0.35$, $\gamma_1 = 0.2$, $\epsilon = 0.0$, and SNR = 2. (a) CDF plot of $e(\hat{\lambda}_1)$, (b) CDF plot of $e(\hat{v}_1)$, (c) CDF plot of $e(\hat{\beta}_1)$. Solid line —— (Method A), dash pattern ---- (Method B), dash pattern --- (Method C).

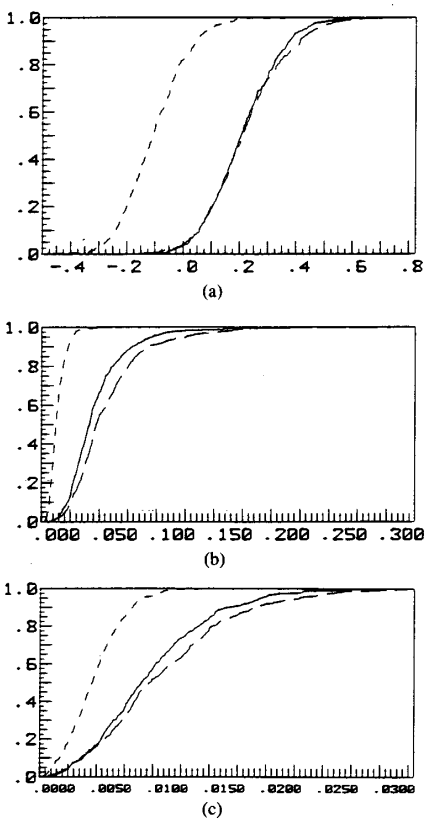


Fig. 2. One source with $\eta_1 = 0.35$, $\gamma_1 = 0.2$, $\epsilon = 0.1$, and SNR = 2. (a) CDF plot of $e(\hat{\lambda}_1)$, (b) CDF plot of $e(\hat{v}_1)$, (c) CDF plot of $e(\hat{\beta}_1)$. Solid line —— (Method A), dash pattern ---- (Method B), dash pattern --- (Method C).

ods A and C. For high contamination fraction, $\epsilon = 0.1$, Method B has a better overall performance and Method C is very close to Method A.

Estimates of the eigenvalues cluster about the true values and almost always have biases. Estimates of $e(\hat{v}_1)$ and $e(\hat{\beta}_1)$ show that, in general, errors in the estimates of signal eigenvectors and frequencies will increase if SNR decreases or ϵ increases.

Situation 2: In order to determine the behavior of estimators in multisource situations under different conditions, here we use two equipowered signal sources with $\eta_1 = \eta_2 = \gamma_1 = 0.25$, SNR = 4, and different second source normalized spatial frequency, γ_2, and variable contamination fraction, ϵ. In this situation we simulate cases with distinct signal eigenvalues. Fig. 5 shows the CDF's of $e(\hat{\lambda}_1)$, $e(\hat{\lambda}_2)$, $e(\hat{v}_1)$, $e(\hat{v}_2)$, $e[\text{map } (\hat{v}_1)]$, $e[\text{map } (\hat{v}_2)]$, $e(\hat{\beta}_1)$, and $e(\hat{\beta}_2)$ for the case that $\gamma_2 = 0.05$ and $\epsilon = 0.0$ (true noise eigenvalues are all equal to 1 and source eigenvalues are $\lambda_1 = 56.42$ and $\lambda_2 = 17.58$). Fig. 6 shows CDF's of the same parameters and histograms of $\hat{\eta}_1$, $\hat{\eta}_2$, $\hat{\gamma}_1$, and $\hat{\gamma}_2$ for the case that $\gamma_2 = 0.05$ and $\epsilon = 0.1$ (true uncontaminated source eigenvalues are the same as before and all contaminated eigenvalues are slightly bigger than uncontaminated ones by $\epsilon\sigma_2^2$ or 0.9 in this case). Study of these plots shows that Method A, conventional, is superior in the case that $\epsilon = 0$, and Method C is better than B in this case. But in the situation where $\epsilon = 0.1$, Method B works better than A and C, however, the error in general is higher than the uncontaminated case.

Estimates of the sources' eigenvalues again have biases, but this bias for Method B in the contaminated case, $\epsilon = 0.1$, is comparatively small. Estimates of parameters corresponding to signal eigenvectors show that errors in each individual eigenvector are much higher than the error of the map of the corresponding eigenvector onto the true signal space. This is an indication of mixing signal eigenvectors, especially in the contaminated case, without any significant effect on the whole signal space. Estimates of parameters corresponding to signal frequencies show that increase of contamination somewhat degrades estimates of signal frequencies in Methods A and C, but improves the corresponding estimates in Method B.

Variations of estimates in this situation with respect to the variation of SNR is exactly the same as Situation 1 (plots are not presented).

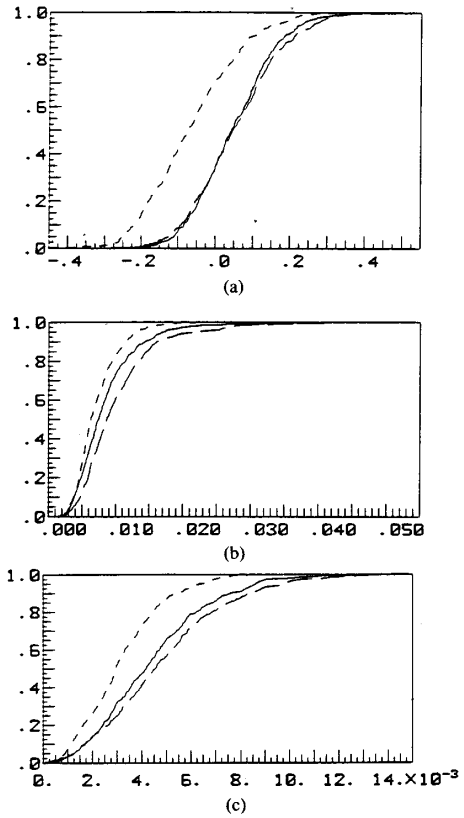


Fig. 3. One source with $\eta_1 = 0.35$, $\gamma_1 = 0.2$, $\epsilon = 0.1$, and SNR = 8. (a) CDF plot of $e(\hat{\lambda}_1)$, (b) CDF plot of $e(\hat{v}_1)$, (c) CDF plot of $e(\hat{\beta}_1)$. Solid line —— (Method A), dash pattern ---- (Method B), dash pattern --- (Method C).

Situation 3: In this situation we are trying to show that, in the two-source situation, when both signal eigenvalues are almost equal, various estimates of parameters will vary. In this situation we have SNR = 4. $\eta_1 = 0.1$, $\eta_2 = 0.4$, $\gamma_1 = 0.45$, $\gamma_2 = 0.1165$, and $\epsilon = 0.0$ (true noise eigenvalues are all equal to 1 and source eigenvalues are $\lambda_1 = \lambda_2 = 37$). Fig. 7 shows histograms of $e(\hat{v}_1)$, $e(\hat{v}_2)$, $e[\text{map}(\hat{v}_1)]$, $e[\text{map}(\hat{v}_2)]$, $\hat{\eta}_1$, $\hat{\eta}_2$, $\hat{\gamma}_1$, and $\hat{\gamma}_2$. From these plots we conclude that as long as signal eigenvectors are significantly different from noise ones, the MUSIC algorithm will work efficiently, because although estimates of both signal eigenvectors are completely different from their true values, the error between the space that they span and the true signal space is negligible. This is in agreement with the theoretical findings in [4] and [7]. Variations of all parameter estimates with respect to changes in SNR and ϵ are as in Situations 1 and 2 (plots are not presented).

Situation 4: In this situation the resolving characteristic of the system is tested by simulating three close equipowered source signals. In this simulation selected parameters are $\epsilon = 0$, SNR = 4, $\eta_1 = \eta_2 = \eta_3 = \gamma_3 = 0.25$, $\gamma_1 = 0.0$, and $\gamma_2 = 0.125$. Fig. 8 shows CDF's of $e(\hat{v}_1)$, $e(\hat{v}_2)$, $e(\hat{v}_3)$, $e[\text{map}(\hat{v}_1)]$, $e[\text{map}(\hat{v}_2)]$, $e[\text{map}(\hat{v}_3)]$,

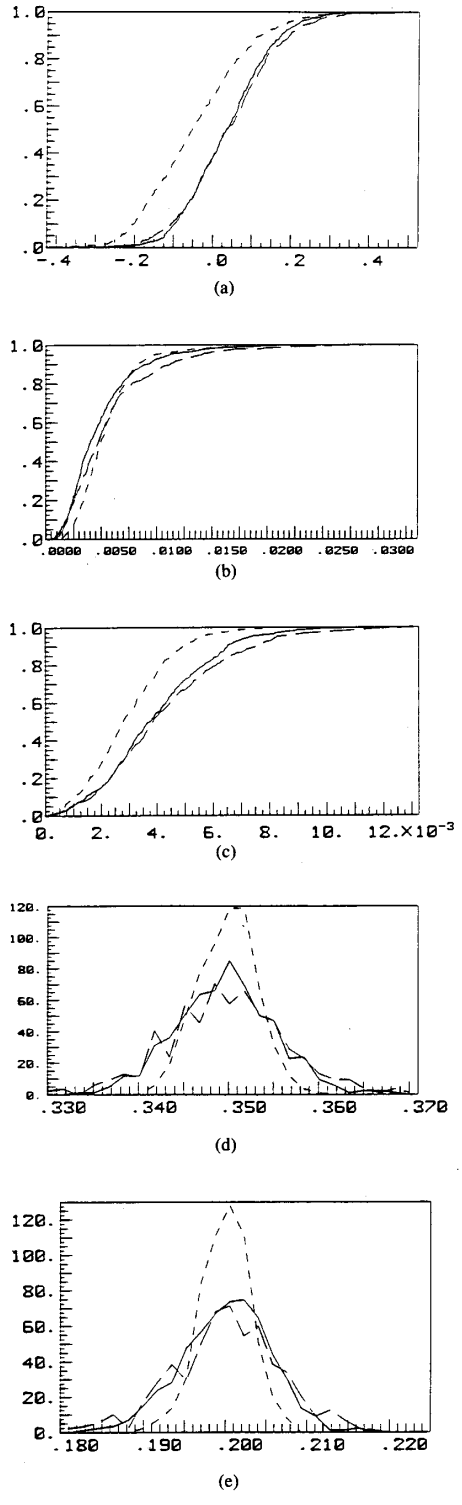


Fig. 4. One source with $\eta_1 = 0.35$, $\gamma_1 = 0.2$, $\epsilon = 0.05$, and SNR = 8. (a) CDF plot of $e(\hat{\lambda}_1)$, (b) CDF plot of $e(\hat{v}_1)$, (c) CDF plot of $e(\hat{\beta}_1)$, (d) histogram of $\hat{\eta}_1$, (e) histogram of $\hat{\gamma}_1$. Solid line —— (Method A), dash pattern ---- (Method B), dash pattern --- (Method C).

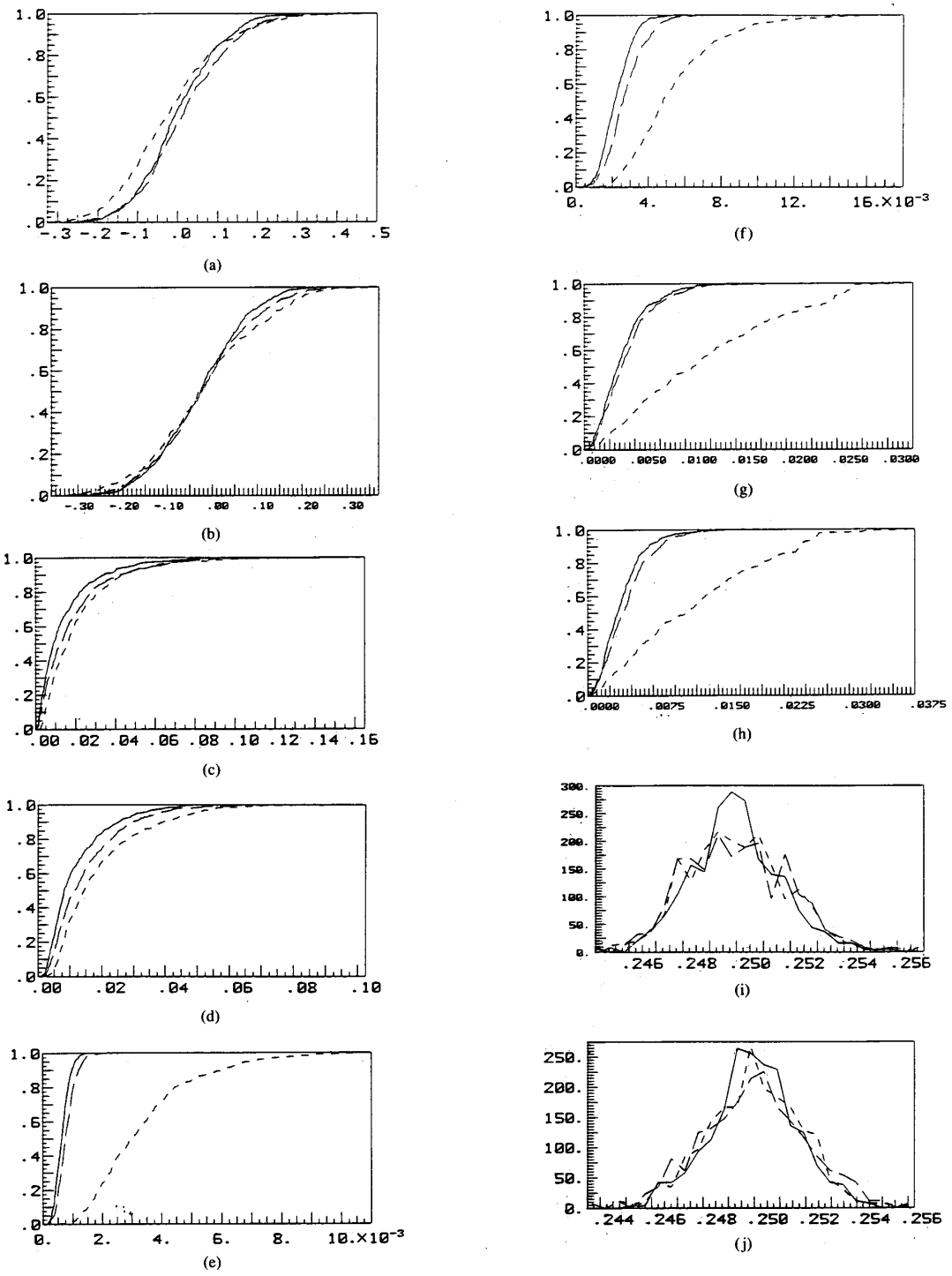


Fig. 5. Two sources with $\eta_1 = \eta_2 = \gamma_1 = 0.25$, $\gamma_2 = 0.05$, $\epsilon = 0.0$, and SNR = 4. (a) CDF plot of $e(\hat{\lambda}_1)$, (b) CDF plot of $e(\hat{\lambda}_2)$, (c) CDF plot of $e(\hat{v}_1)$, (d) CDF plot of $e(\hat{v}_2)$, (e) CDF plot of $e[\text{map}(\hat{v}_1)]$, (f) CDF plot of $e[\text{map}(\hat{v}_2)]$, (g) CDF plot of $e(\hat{\beta}_1)$, (h) CDF plot of $e(\hat{\beta}_2)$, (i) histogram of $\hat{\eta}_1$, (j) histogram of $\hat{\eta}_2$.

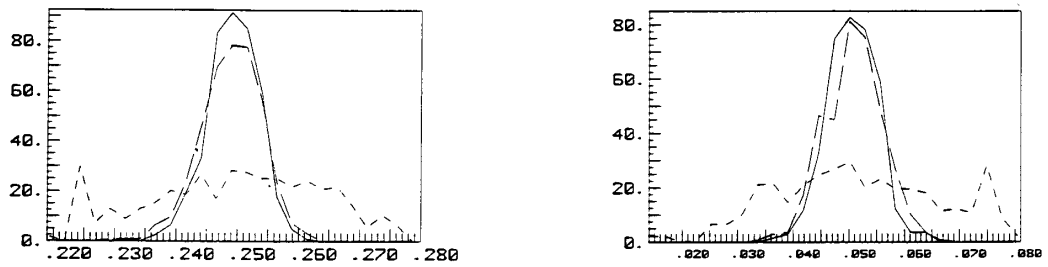


Fig. 5. (*Continued.*) (k) Histogram of $\hat{\gamma}_1$, (l) histogram of $\hat{\gamma}_2$. Solid line —— (Method A), dash pattern ---- (Method B), dash pattern --- (Method C).

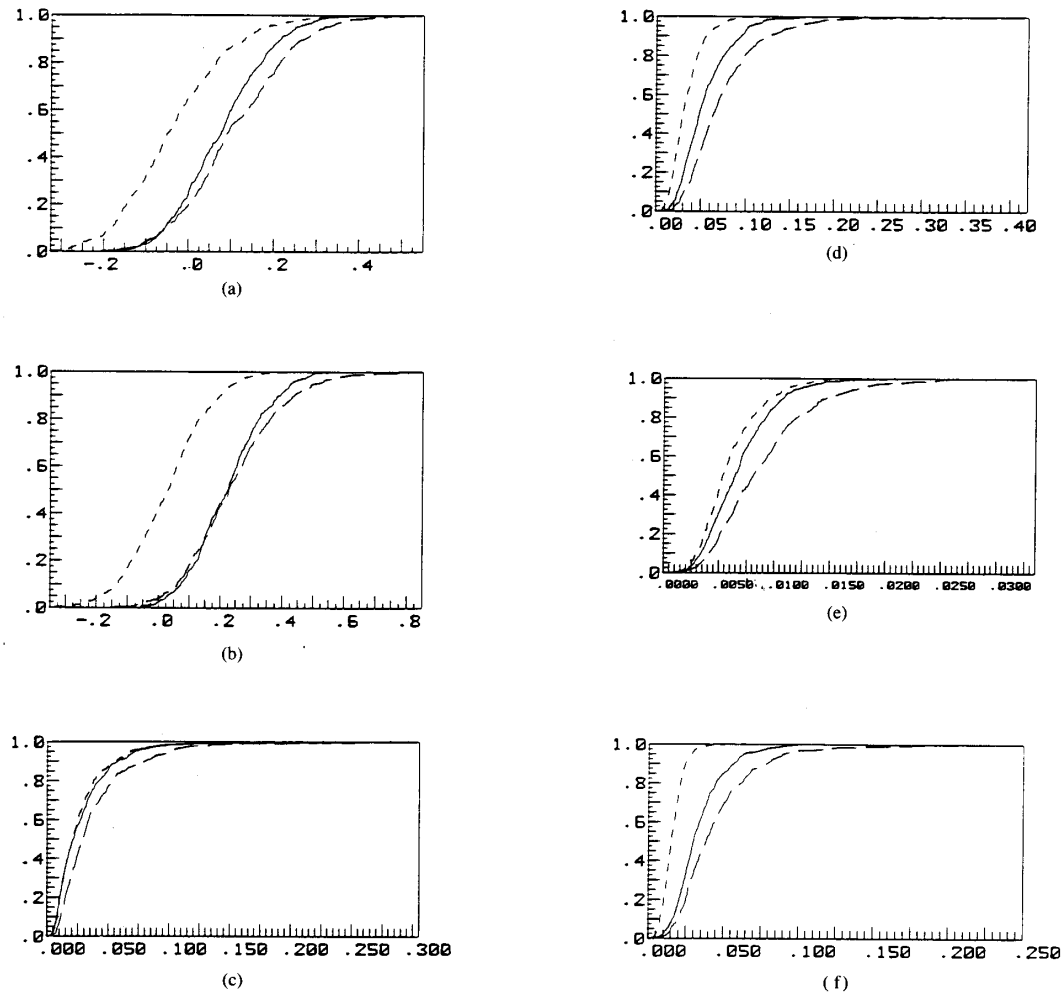


Fig. 6. Two sources with $\eta_1 = \eta_2 = \gamma_1 = 0.25$, $\gamma_2 = 0.05$, $\epsilon = 0.1$, and SNR = 4. (a) CDF plot of $e(\hat{\lambda}_1)$, (b) CDF plot of $e(\hat{\lambda}_2)$, (c) CDF plot of $e(\hat{v}_1)$, (d) CDF plot of $e(\hat{v}_2)$, (e) CDF plot of $e[\text{map}(\hat{v}_1)]$, (f) CDF plot of $e[\text{map}(\hat{v}_2)]$.

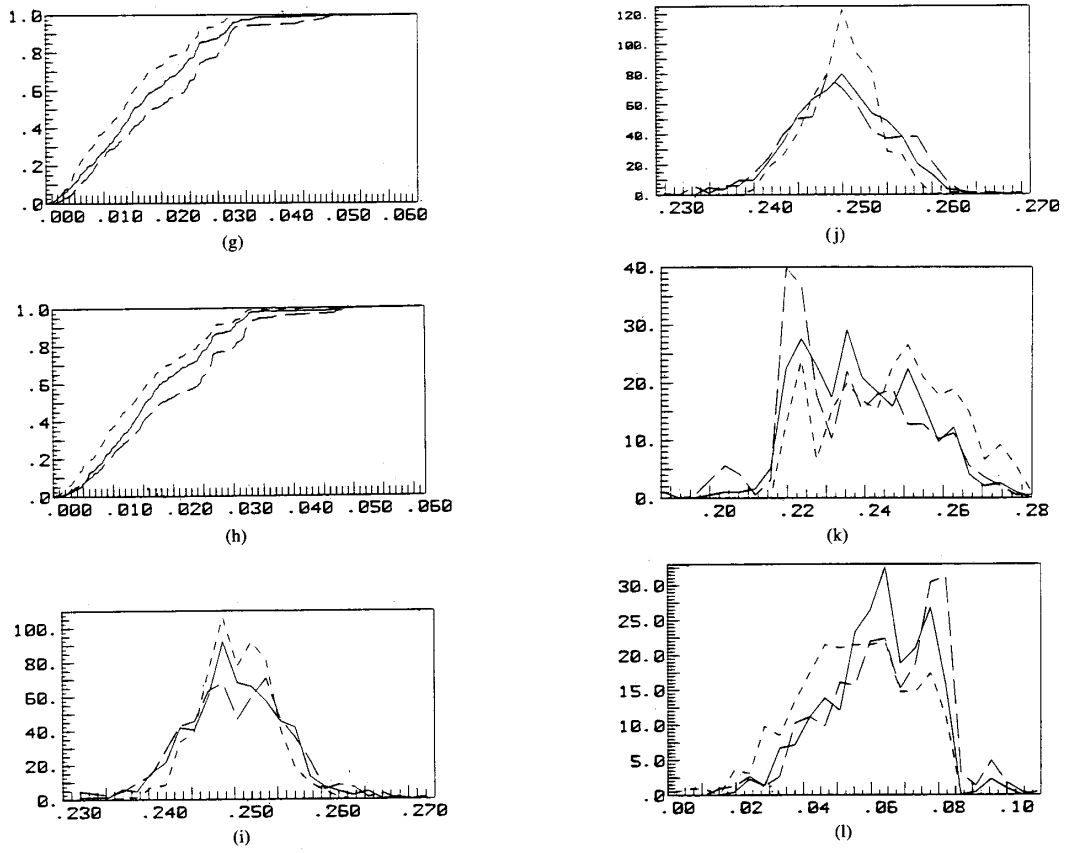


Fig. 6. (*Continued.*) (g) CDF plot of $e(\hat{\beta}_1)$, (h) CDF plot of $e(\hat{\beta}_2)$, (i) histogram of $\hat{\eta}_1$, (j) histogram of $\hat{\eta}_2$, (k) histogram of $\hat{\gamma}_1$, (l) histogram of $\hat{\gamma}_2$. Solid line ——— (Method A), dash pattern ---- (Method B), dash pattern --- (Method C).

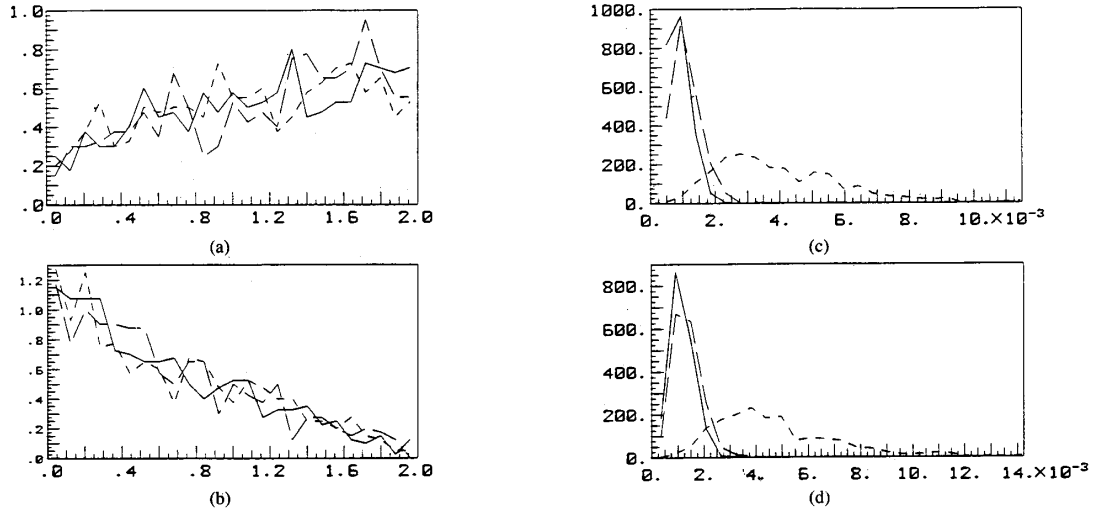


Fig. 7. Two sources with $\eta_1 = 0.1$, $\eta_2 = 0.4$, $\gamma_1 = 0.45$, $\gamma_2 = 0.1165$, $\epsilon = 0.0$, and SNR = 4. (a) Histogram of $e(\hat{v}_1)$, (b) histogram of $e(\hat{v}_2)$, (c) histogram of $e[\text{map}(\hat{v}_1)]$, (d) histogram of $e[\text{map}(\hat{v}_2)]$.

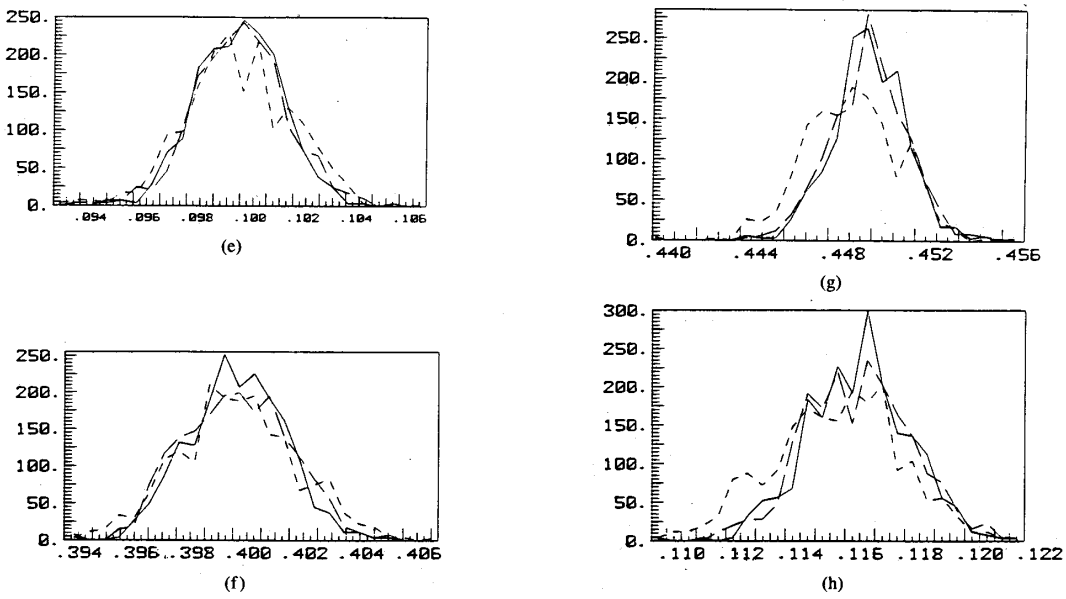


Fig. 7. (*Continued.*) (e) Histogram of $\hat{\eta}_1$, (f) histogram of $\hat{\eta}_2$, (g) histogram of $\hat{\gamma}_1$, (h) histogram of $\hat{\gamma}_2$. Solid line —— (Method A), dash pattern ---- (Method B), dash pattern --- (Method C).

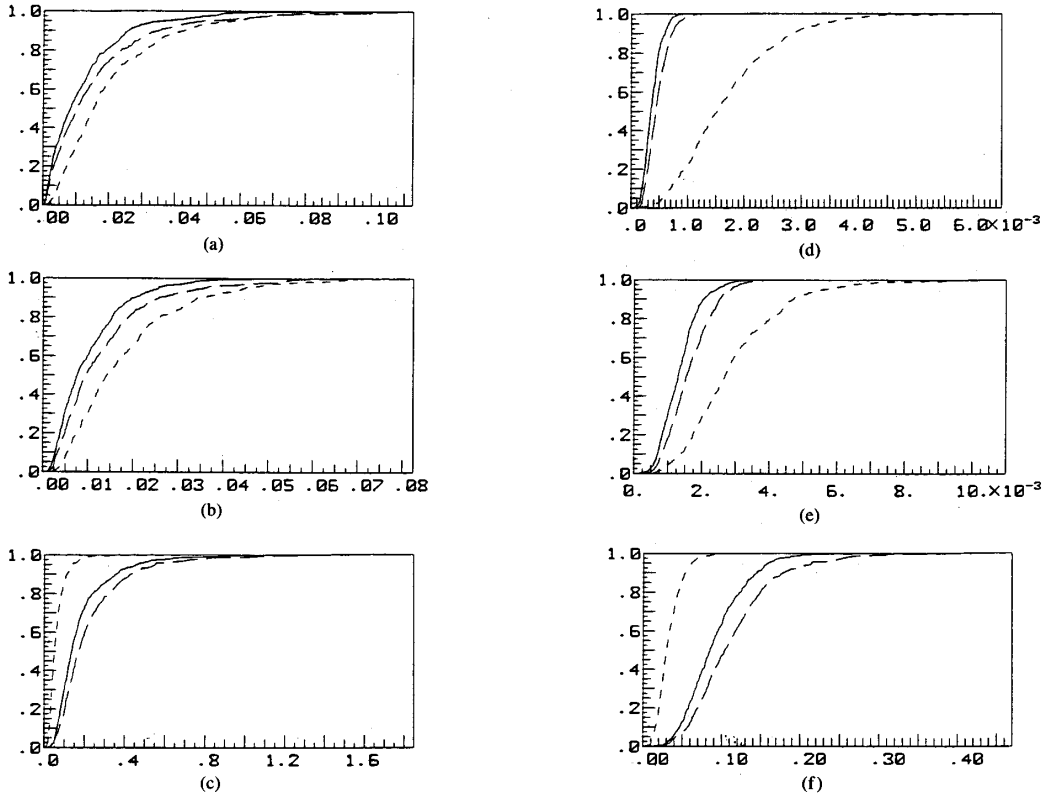


Fig. 8. Three sources with $\eta_1 = \eta_2 = \eta_3 = \gamma_3 = 0.25$, $\gamma_1 = 0.0$, $\gamma_2 = 0.125$, $\epsilon = 0.0$, and SNR = 4. (a) CDF plots of $e(\hat{v}_1)$, (b) CDF plot of $e(\hat{v}_2)$, (c) CDF plot of $e(\hat{v}_3)$, (d) CDF plot of $e[\text{map}(\hat{v}_1)]$, (e) CDF plot of $e[\text{map}(\hat{v}_2)]$, (f) CDF plot of $e[\text{map}(\hat{v}_3)]$.

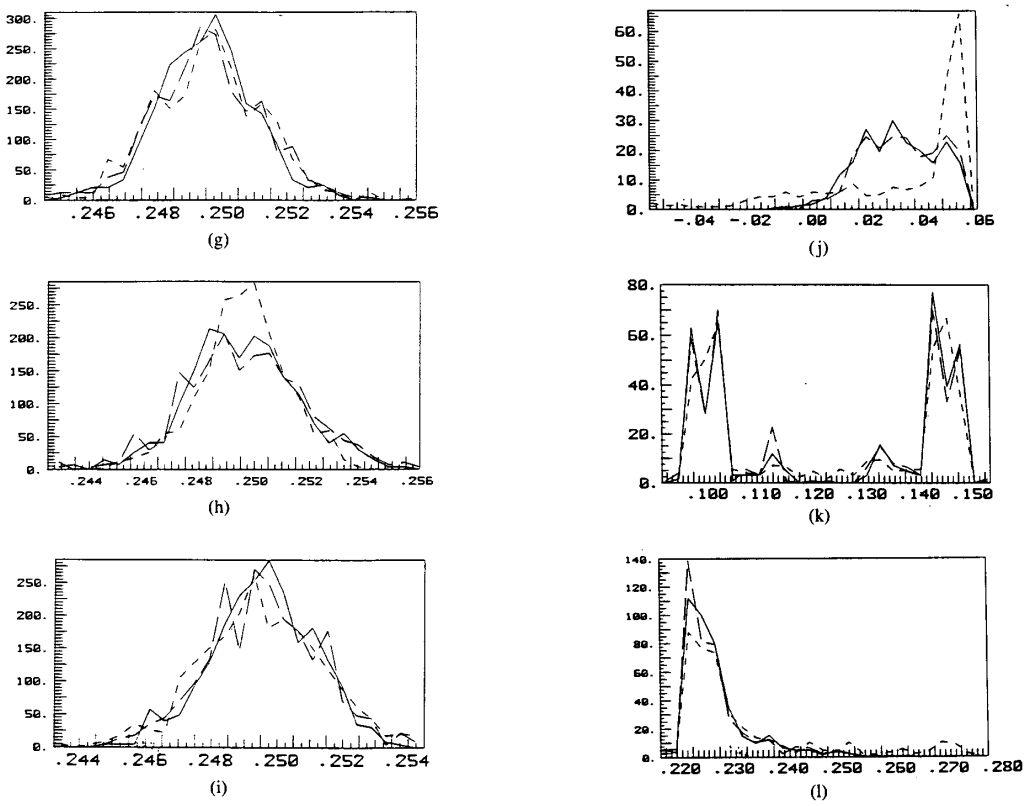


Fig. 8. (*Continued.*) (g) Histogram of $\hat{\eta}_1$, (h) histogram of $\hat{\eta}_2$, (i) histogram of $\hat{\eta}_3$, (j) histogram of $\hat{\gamma}_1$, (k) histogram of $\hat{\gamma}_2$, (l) histogram of $\hat{\gamma}_3$. Solid line ——— (Method A), dash pattern ---- (Method B), dash pattern --- (Method C).

$\hat{\eta}_1$, $\hat{\eta}_2$, $\hat{\eta}_3$, $\hat{\gamma}_1$, $\hat{\gamma}_2$, and $\hat{\gamma}_3$, in the case that $m = h = 3$ (true noise eigenvalues are all equal to 1 and signal eigenvalues are 1.59243, 25, and 84.4076). Fig. 9 shows the same CDF's and histograms, but in the case that $m = 6$ and $h = 2$ (true noise eigenvalues are all equal to 1 and signal eigenvalues are 21.6863, 60.3138, and 65.0001). Careful consideration of these plots shows that we do have mixing in signal eigenvectors if one of the signal eigenvalues becomes very close to another signal eigenvalue or to the noise eigenvalues, however, those with large eigenvalues mix within the true signal space but the one with small eigenvalue, 1.59, will be mixed with noise eigenvectors, and henceforth, we are unable to locate all three sources correctly. As a matter of fact, if we use all three signal eigenvectors to estimate the null spectrum, we are able to detect three sources, but our estimates would be completely off from the true values and hence unreliable. If we increase the number of sensors, and hence increase the resolution of the system, the estimate of the signal space improves substantially. However, estimates of the first two eigenvectors in the latter case are mixed up because of the closeness of the signal eigenvalues λ_1 and λ_2.

Variations of all estimates with respect to changes in SNR and ϵ are again the same as before (plots are not presented).

IV. Major Eigenvectors Variation with Bearing, Noise Free Data

The results in the previous section show that there is a large spread in the signal eigenvectors if their corresponding eigenvalues are bunched or if there is at least one signal eigenvalue close to the noise-related eigenvalues. In the former case, the MUSIC algorithm will not be hurt because the space spanned by the estimated signal eigenvectors is almost the same as the true signal space, but in the latter case one of the signal eigenvectors is mixing with noise-related eigenvectors which introduces a significant error in the estimated signal space and hence bearing.

Fig. 10 shows signal eigenvalues versus second source normalized spatial frequency, γ_2, where other parameters are kept constant. In this case, the number of sources is 2, $\eta_1 = \eta_2 = \gamma_1 = 0.25$, and SNR = 8. From this plot we realize that it is impossible to locate two sources correctly, in this condition, whenever $0.17 < \gamma_2 < 0.33$.

Fig. 11 shows signal eigenvalues versus third source normalized spatial frequency, γ_3, when other parameters are kept constant. In this case, the number of sources is 3, $\eta_1 = \eta_2 = \gamma_1 = 0.25$, $\gamma_2 = 0$, and SNR = 8. This plot shows that the estimate of the signal space will be

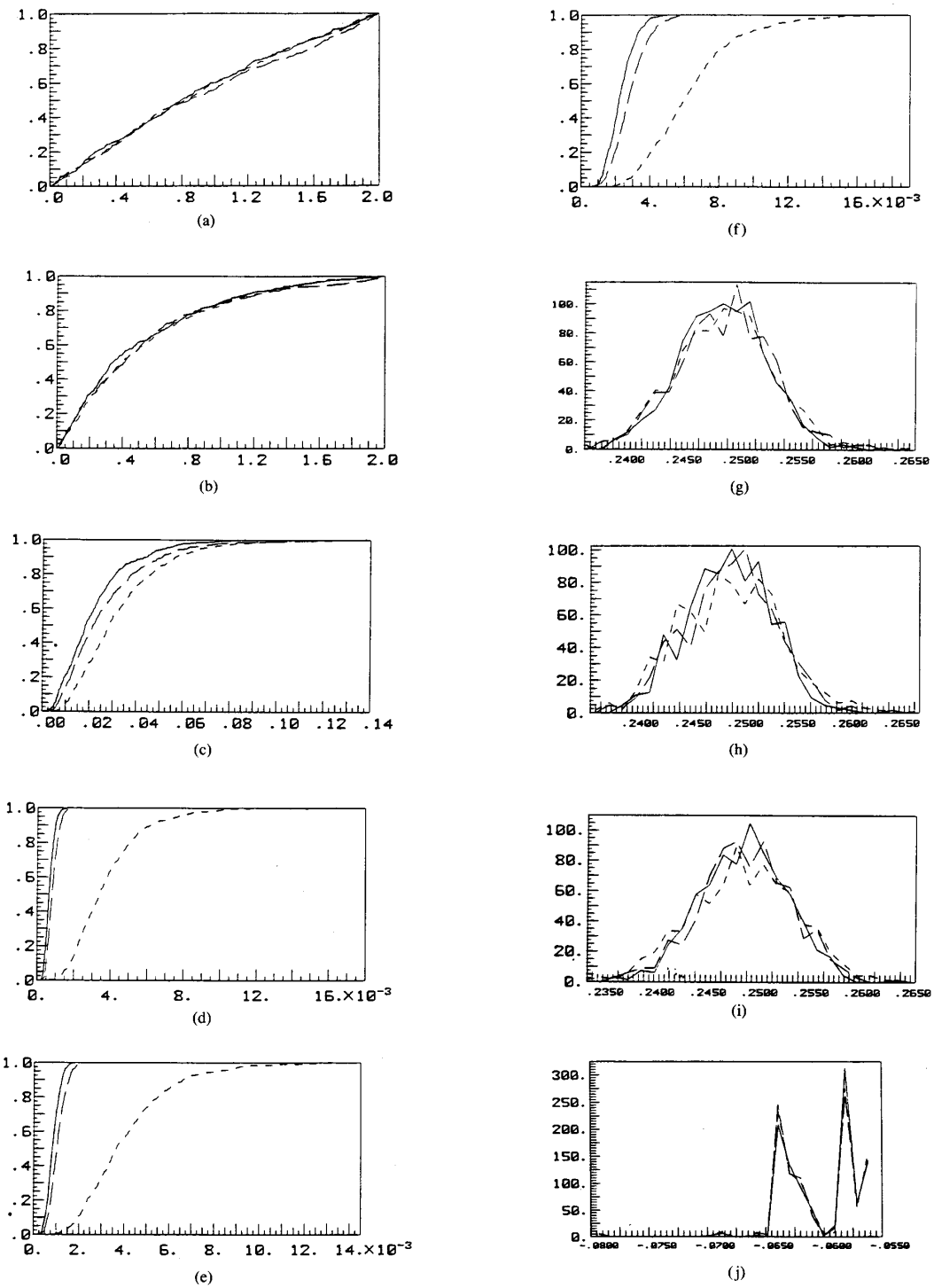
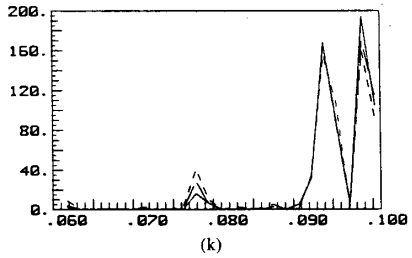
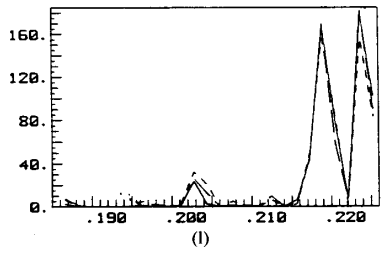


Fig. 9. Three sources with $\eta_1 = \eta_2 = \eta_3 = \gamma_3 = 0.25$, $\gamma_1 = 0.0$, $\gamma_2 = 0.125$, $\epsilon = 0.0$, and SNR = 4, $m = 6$, $h = 2$. (a) CDF plot of $e(\hat{v}_1)$, (b) CDF plot of $e(\hat{v}_2)$, (c) CDF plot of $e(\hat{v}_3)$, (d) CDF plot of $e[\text{map}(\hat{v}_1)]$, (e) CDF plot of $e[\text{map}(\hat{v}_2)]$, (f) CDF plot of $e[\text{map}(\hat{v}_3)]$, (g) histogram of $\hat{\eta}_1$, (h) histogram of $\hat{\eta}_2$, (i) histogram of $\hat{\eta}_3$, (j) histogram of $\hat{\gamma}_1$,



(k)



(l)

Fig. 9. (*Continued.*) (k) histogram of $\hat{\gamma}_2$. (l) histogram of $\hat{\gamma}_3$. Solid line —— (Method A), dash pattern ---- (Method B), dash pattern --- (Method C).

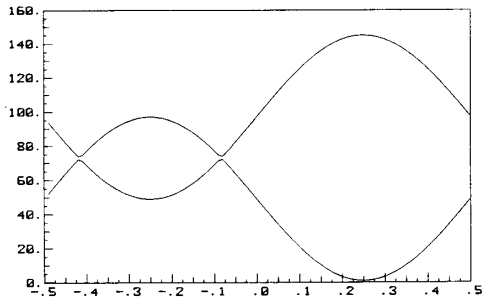


Fig. 10. Signal eigenvalues λ_1 and λ_2 versus γ_2, where number of sources is 2, $\eta_1 = \eta_2 = \gamma_1 = 0.25$, and SNR = 8.

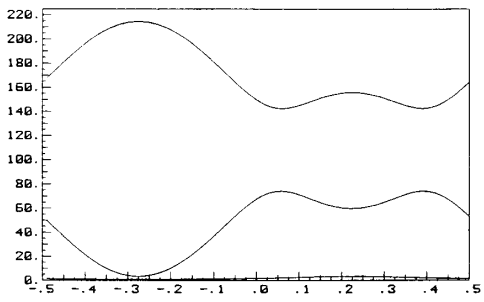


Fig. 12. Signal eigenvalues λ_1, λ_2 and λ_3 versus γ_3, where number of sources is 3, $\eta_1 = \eta_2 = \eta_3 = 0.25$, $\gamma_1 = -0.25$, $\gamma_2 = -0.3$, and SNR = 8.

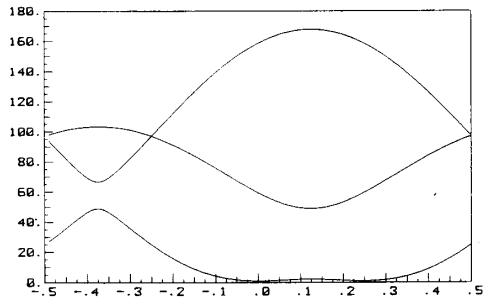


Fig. 11. Signal eigenvalues λ_1, λ_2 and λ_3 versus γ_3, where number of sources is 3, $\eta_1 = \eta_2 = \eta_3 = \gamma_2 = 0.25$, $\gamma_1 = 0$, and SNR = 8.

unreliable in this case whenever $-0.06 < \gamma_3 < 0.31$. This uncertainty will increase if the other two sources become closer to each other as in the case shown in Fig. 12. This effect will vanish if the source's temporal frequencies are completely distinct (plots are not presented).

V. Overall Conclusions

In summary, the following five conclusions are given.

1) In estimation of eigenvalues, all methods under any of the test conditions have bias. For low contamination fraction, $\epsilon < 0.05$, the conventional method gives the better result, but for high contamination, $\epsilon = 0.1$, the robust rank method works better than the others, and the robust weighted *M*-estimation method does not work better than the conventional one. This situation is true for any SNR.

2) Study of $e(\hat{v}_i)$ shows that in estimation of signal-related eigenvectors, as long as their corresponding eigenvalues are significantly different, the conventional method works very well in the low contamination, the robust rank method works better in the high contamination situations, and the robust weighted *M*-estimation method is worse in the comparison with the conventional one in all conditions. In the case where signal-related eigenvalues are bunched or close to noise-related eigenvalues, estimates of signal-related eigenvectors are mixed and they will be completely different from their true values.

3) By careful study of $e[\text{map}\ (\hat{v}_i)]$, we conclude that as long as signal-related eigenvalues are not close to noise-related eigenvalues, there is not any significant mixing between signal and noise eigenvectors, and the space spanned by the estimated signal eigenvectors is almost the same as the true signal space. In the case where one of the signal-related eignevalues is close to the noise eigenvalues, there is a significant mixing between its corresponding eigenvector and noise-related eigenvectors, and that will introduce a substantial error in the estimation of bearing for some sources. The only remedy to this problem is to increase the overall array signal-to-noise ratio by increasing the number of sensors and filtering noise as much as possible. In the case of high contamination fraction, ϵ, it is advisable to use the robust rank method.

4) Effects of variation of SNR and contamination fraction can be summarized as follows. In general, errors in estimation of all parameters decrease as SNR increases and inversely if ϵ, the contamination fraction, increases error in estimates will increase rapidly. This conclusion holds in all test conditions.

5) Source-bearing configuration will greatly effect the spread of the estimate of signal space, space spanned by the estimate of signal-related eigenvectors, and in some

situations, particularly common frequency, some source eigenvalues will be indistinguishable from those related to noise. This situation becomes worse as the number of sources increases.

Appendix

Robust Estimator of R Using Nonparametric Statistics (Method B)

One basic approach in estimating a robust covariance matrix R is to estimate each off-diagonal element of R separately by a robust correlation coefficient r. This elementwise bivariate approach may be contrasted with "multivariate" ones, which manipulate all of the variables simultaneously. Nonparametric estimators for the correlation coefficient are the Spearman, r_s, and the Kendall, r_k. The Spearman rank correlation coefficient [11], r_{sil}, is calculated using the ranks as the paired measurements on the two real variables X_i and X_l [for a given set of paired data $\{X_{ik}, X_{lk}; k = 1, 2, \cdots, N\}$, it is obtained by ranking the X_i's among themselves, and also the X_l's, i.e., $X_i(k)$ and $X_l(k)$ are the ranks of X_{ik} and X_{lk} in the pair (X_{ik}, X_{lk}), respectively] in the formula for the estimator of r. Thus,

$$r_{sil} = \frac{N \sum_{k=1}^{N} X_i(k)\, X_l(k) - \sum_{k=1}^{N} X_i(k) \sum_{k=1}^{N} X_l(k)}{\left[N \sum_{k=1}^{N} (X_i(k))^2 - \left(\sum_{k=1}^{N} X_i(k)\right)^2\right]^{1/2} \left[N \sum_{k=1}^{N} (X_l(k))^2 - \left(\sum_{k=1}^{N} X_l(k)\right)^2\right]^{1/2}}.$$

The above expression can be algebraically reduced to the simpler expression [11]

$$r_{sil} = 1 - \frac{6 \sum_{k=1}^{N} d_i^2(k)}{N(N^2 - 1)}$$

where $d_i(k) = X_i(k) - X_l(k)$.

The rank correlation coefficient of the above expression replaces the ρ_{ij} terms in the expression cov $[X_i, X_l] = \rho_{ij} \sigma_i \sigma_l$ for the correlation matrix. In this case, median deviation estimate of standard deviation is used for σ_i in the expression of covariance matrix. Hence, the method would be robust to outliers and suitable for contaminated heavily tailed Gaussian noise.

In the case of analytic data, all vectors are complex and it is necessary to estimate four real covariance matrix in order to formulate the complex covariance matrix. In this case, we assume that measurements on the two complex variables $X_i + jY_i$ and $X_l + jY_l$ are available. Then using the following definition:

$$\text{cov}\,[z, w] = E\left[\left(w - E(w)\right)\left(z - E(z)\right)^*\right]$$

we obtain

$$\begin{aligned}\text{cov}\,[X_i + jY_i, X_l + jY_l] &= \text{cov}\,[X_i, X_l] + \text{cov}\,[Y_i, Y_l] \\ &\quad + j\left\{\text{cov}\,[X_i, Y_l] - \text{cov}\,[Y_i, X_l]\right\}\end{aligned}$$

where each one of these four real covariances can be estimated using the rank correlation method. Thus, a robust rank estimator for R, in general, is formulated.

Robust Estimator of R Using Weighted M-Estimation (Method C) [9]

In this method, the Mahalanobis squared distance

$$d_i^2 = (\boldsymbol{X}_i - \overline{\boldsymbol{X}})^H R^{-1} (\boldsymbol{X}_i - \overline{\boldsymbol{X}})$$

is used to detect atypical multivariate vectors of observations and calculate their corresponding weights. In this equation, $\boldsymbol{X}_i$ is a random vector and $\overline{\boldsymbol{X}}$ is a measure of the mean. From the applied viewpoint, M-estimators can be considered as a simple modification of classical estimators; they give full weight to observations assumed to come from the main body of the data, but reduced weight or influence to observations from the tails of the contaminating distribution. In practice, the influence of observations with unduly large Mahalanobis distances is downweighted.

The equations used here to define robust estimators of means and covariances are as follows:

$$\overline{\boldsymbol{X}} = \sum_{i=1}^{n} w_i \boldsymbol{X}_i \Big/ \sum_{i=1}^{n} w_i$$

and

$$\hat{R} = \sum_{i=1}^{n} w_i^2 (\boldsymbol{X}_i - \overline{\boldsymbol{X}})(\boldsymbol{X}_i - \overline{\boldsymbol{X}})^H \Big/ \left(\sum_{i=1}^{n} w_i^2 - 1\right)$$

where

$$w_i = w(d_i) = \omega(d_i)/d_i$$

and

$$d_i = \left\{\left|(\boldsymbol{X}_i - \overline{\boldsymbol{X}})^H \hat{R}^{-1} (\boldsymbol{X}_i - \overline{\boldsymbol{X}})\right|\right\}^{1/2}.$$

The solutions for $\overline{\boldsymbol{X}}$ and $\hat{R}$ are iterative.

The two-parameter form of ω used here is

$$\begin{aligned}\omega(d) &= d \quad \text{if } d \le d_0 \\ &= d_0 \quad \text{if } d > d_0\end{aligned}$$

$$d_0 = \sqrt{\nu} + \sqrt{2}$$

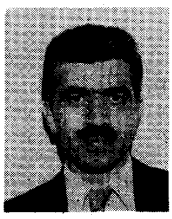
where ν is the dimension of $\boldsymbol{X}_i$. For comprehensive discussion on this method refer to [9].

ACKNOWLEDGMENT

The author wishes to acknowledge the discreet criticism and comprehensive comments of reviewers which were of enormous help throughout the work discussed here.

REFERENCES

[1] M. Wax, T. Shan, and T. Kailath, "Spatio-temporal spectral analysis by eigenstructure methods," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-32, Aug. 1984.
[2] R. O. Schmidt, Ph.D. dissertation, Stanford Univ., Stanford, CA, 1981.
[3] K. Fukunaga, *Introduction to Statistical Pattern Recognition.* New York: Academic, 1979.
[4] D. R. Brillinger, *Time Series, Data Analysis and Theory.* New York: Holt, Rinehart, and Winston, 1975.
[5] J. H. Wilkinson, *The Algebraic Eigenvalue Problem.* Oxford, England: Clarendon, 1965.
[6] C. Fang and P. R. Krishnaiah, "Asymptotic distribution on functions of the eigenvalues of some random matrices for nonnormal populations," *J. Multivar. Anal.*, vol. 12, pp. 39–63, 1982.
[7] M. Kaveh and A. J. Barabell, "The statistical performance of the MUSIC and the minimum-norm algorithms in resolving plane waves in noise," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-34, pp. 331–341, Apr. 1986.
[8] ——, "Corrections to 'The statistical performance of the MUSIC and the minimum-norm algorithms in resolving plane waves in noise,' " *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-34, June 1986.
[9] N. A. Campbell, "Robust procedures in multivariate analysis I: Robust covariance estimation," *Appl. Stat.*, vol. 29, no. 3, pp. 231–237, 1980.
[10] S. J. Devlin, R. Granadesikan, and J. R. Kettenring, "Robust estimation of dispersion matrices and principle components," *J. Amer. Stat. Assoc.*, vol. 76, pp. 354–362, June 1981.
[11] M. Hollander and D. A. Wolfe, *Nonparametric Statistical Methods.* New York: Wiley, 1973.



Alireza Moghaddamjoo (S'84–M'86) was born in Tehran, Iran, on March 16, 1953. He received the B.S. degree in electrical engineering from the University of Tehran, Tehran, Iran, in 1976, the M.S. degree in nuclear engineering from the Massachusetts Institute of Technology, Cambridge, in 1978, and the Ph.D. degree in electrical engineering from the University of Wyoming, Laramie, in 1986.

At the University of Wyoming he was a Research Assistant and a Western Research Institute Graduate Scholar. During his Ph.D. program he was involved in several funded projects, namely, robust Kalman tracking, robust step detection, covariance estimation and eigenstructure variability, and eigencoding for seismic data. He was with the Electronic Department of the Nuclear Research Center (NRC) in Tehran, Iran, from August 1978 to December 1983, and held a Lecturer position in the Department of Electrical Engineering of the University of Wyoming for the spring semester of 1986. He is currently an Assistant Professor in the Department of Electrical Engineering and Computer Science of the University of Wisconsin, Milwaukee, which he joined in 1986. His teaching and research are primarily in the areas of digital signal processing, adaptive filtering, pattern recognition, and image processing. His other professional experiences include collaboration and consulting activities with the Rush Medical School, Chicago, IL, AMOCO Research Center, Tulsa, OK, and UNOCAL Research and Technology, Los Angeles, CA.