MUSIC, Maximum Likelihood, and Cramér–Rao Bound: Further Results and Comparisons

PETRE STOICA AND ARYE NEHORAI, SENIOR MEMBER, IEEE

Abstract—**The problem of determining the direction of arrival of narrow-band plane waves using sensor arrays and the related one of estimating the parameters of superimposed signals from noisy measurements have received significant attention in the recent signal processing literature. A number of results have been presented recently in an article by the authors of this paper on the statistical performance of the multiple signal characterization (MUSIC) and the maximum likelihood (ML) estimators for the above problems. This companion paper extends the results of our previous article in several directions. First, it establishes that in the class of weighted MUSIC estimators, the unweighted MUSIC achieves the best performance (i.e., the minimum variance of estimation errors), in large samples. Next, it derives the covariance matrix of the ML estimator and presents detailed analytic studies of the statistical efficiency of MUSIC and ML estimators. These studies include performance comparisons of MUSIC and MLE with each other as well as with the ultimate performance corresponding to the Cramér–Rao bound (CRB). Finally, the paper contains some numerical examples which provide a more quantitative study of performance for the problem of finding two directions with uniform linear sensor arrays.**

I. Introduction

FOR a number of signal processing applications (for the details of which we refer to [1]–[3] and the references therein), the relevant problem is estimation of the parameters in the following model:

$$y(t) = A(\theta)\, x(t) + e(t) \qquad t = 1, 2, \cdots, N \quad (1.1a)$$

where $\{y(t)\} \in \mathbf{C}^{m \times 1}$ are the observed data vectors, $\{x(t)\} \in \mathbf{C}^{n \times 1}$ are unknown vectors and $e(t) \in \mathbf{C}^{m \times 1}$ is an additive noise. The matrix $A(\theta) \in \mathbf{C}^{m \times n}$ and the vector $\theta \in \mathbf{R}^{n \times 1}$ are given by

$$A(\theta) = [a(\omega_1) \ \cdots \ a(\omega_n)] \quad (1.1b)$$

$$\theta = [\omega_1 \ \cdots \ \omega_n]^T \quad (1.1c)$$

with $\{\omega_i\}$ being the real-valued unknown parameters. The form of $a(\omega)$ varies from application to application. For most of this paper, we will not specify the form of $a(\omega)$ to confer generality on our results.

In this paper we will be concerned with the problem of estimating the parameter vector θ in (1.1). The dimension n of $x(t)$ is assumed to be known. Techniques for estimating n are well documented in the literature (see, e.g., [4]–[6]), and will not be discussed here. Concerning $x(t)$, estimation of these "parameters" is immediate once an estimate of θ is available (see, e.g., [1], [3], [7], [8]).

In [1] we presented a number of results on the performance of the MUltiple Signal Characterization (MUSIC) and the maximum likelihood (ML) estimates of the parameters θ. We also derived the Cramér–Rao bound (CRB) on the covariance matrix of any unbiased estimate of θ, and provided a comparison of the MUSIC performance to the ultimate performance corresponding to the CRB. In the present (companion) paper, we continue our study in [1] and establish a number of additional results on MUSIC, ML method, and CRB, which complete the picture of the estimation methods in question and, in particular, of their performances. For convenience, the presentation of the contributions of this paper and of the way in which they complete the results of [1], is in the next section.

II. Preliminaries and Outlook

First we introduce some assumptions on the model (1.1), that are considered to hold throughout the paper.

A. Basic Assumptions

A1: The vectors $a(\omega)$ corresponding to $(n + 1)$ different values of ω, are linearly independent (which implies $m > n$).

A2: The noise $\{e(t)\}$ is Gaussian distributed, and

$$Ee(t)\, e^*(s) = \sigma I \delta_{t,s}$$

$$Ee(t)\, e^T(s) = 0 \qquad (\text{all } t \text{ and } s)$$

where the superscripts T and $*$ denote, respectively, the transpose and the conjugate transpose of the quantity in question, and $\delta_{t,s}$ is the Kronecker delta ($\delta_{t,s} = 1$ for $t = s$, and 0 for $t \neq s$).

A3: The covariance matrix of $x(t)$

$$P = Ex(t)\, x^*(t) \quad (2.1)$$

is positive definite. Furthermore, $x(t)$ and $e(s)$ are uncorrelated for all t and s, and $N > m$.

The assumptions above are standard in the literature on estimation of the parameters θ in (1.1). For some comments on these assumptions, see [1].

Next, we set some notation.

Manuscript received October 4, 1988; revised January 9, 1990. The work of A. Nehorai was supported by the Air Force Office of Scientific Research under Grant AFOSR-90-0164.

P. Stoica is with the Department of Automatic Control, Polytechnic Institute of Bucharest, R-77 206 Bucharest, Romania.

A. Nehorai is with the Department of Electrical Engineering, Yale University, New Haven, CT 06520.

IEEE Log Number 9038956.

B. Basic Notation

A_{ij}	the i, j element of a general matrix $A \in C^{k \times p}$,
$A \odot B$	the Hadamard product of the matrices A, $B \in C^{k \times p}$, defined by $[A \odot B]_{ij} = A_{ij} B_{ij}$,
$A \geq B$	the difference $A - B$ of the Hermitian positive semidefinite matrices A and B, is positive semidefinite,
$d(\omega_i)$	$da(\omega)/d\omega \mid_{\omega = \omega_i}$,
$D(\theta)$	$[d(\omega_1) \cdots d(\omega_n)]$,
$R = Ey(t) \cdot y^*(t)$	$APA^* + \sigma I$ (the covariance matrix of $y(t)$), (2.2)
$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_m$	the eigenvalues of R, in decreasing order,
$s_1, \cdots, s_n$	the orthonormal eigenvectors of R, associated with $\lambda_1, \cdots, \lambda_n$ (the "signal eigenvectors"),
$g_1, \cdots, g_{m-n}$	the orthonormal eigenvectors of R, associated with $\lambda_{n+1}, \cdots, \lambda_m$ (the "noise eigenvectors"),
S	$[s_1 \cdots s_n]$,
G	$[g_1 \cdots g_{m-n}]$,
$\hat{R}$	$1/N \sum_{t=1}^{N} y(t) y^*(t)$ (the sample covariance matrix of $y(t)$), (2.3)
$\hat{s}_1, \cdots, \hat{s}_n$; $\hat{g}_1, \cdots, \hat{g}_{m-n}$	the orthonormal eigenvectors of $\hat{R}$, arranged in the decreasing order of the associated eigenvalues,
$\hat{S}$, $\hat{G}$	the matrices S and G made of $\{\hat{s}_k\}$ and, respectively, $\{\hat{g}_k\}$.

We are now able to present some of the results in [1] that are relevant to this paper. These results concern the statistical performance (for large N) of the following two estimators of θ:

- The *MUSIC estimator* ([5], [9], [14]), which is given by the locations of the n smallest values of the following function:

$$f(\omega) = a^*(\omega) \hat{G}\hat{G}^* a(\omega) = a^*(\omega)[I - \hat{S}\hat{S}^*] a(\omega). \quad (2.4)$$

- The deterministic (or conditional) *ML estimator* (introduced in [15] and also studied in [3], [7], [8], [13]), which is given by the minimizer of the following function:

$$F(\theta) = \text{tr}\left[I - A(A^*A)^{-1}A^*\right]\hat{R} \quad (2.5)$$

where "tr" stands for "trace." Note that, for notational convenience, we will simply write A instead of $A(\theta)$, D instead of $D(\theta)$, etc., whenever there is no possibility of confusion.

C. Review of Some Results in [1]

1) MUSIC Covariance Matrix: The MUSIC estimation errors $\{\hat{\omega}_i - \omega_i\}$ are asymptotically (for large N) jointly Gaussian distributed with zero means and the following covariance matrix:

$$C_{\text{MU}} = \frac{\sigma}{2N}(H \odot I)^{-1} \text{Re}\left\{H \odot (A^*UA)^T\right\} \cdot (H \odot I)^{-1} \quad (2.6)$$

where Re (x) denotes the real part of x,

$$H = D^*GG^*D = D^*\left[I - A(A^*A)^{-1}A^*\right]D \quad (2.7)$$

and U is implicitly defined by

$$A^*UA = P^{-1} + \sigma P^{-1}(A^*A)^{-1}P^{-1}. \quad (2.8)$$

The diagonal elements of C_{MU} give the variances of estimation errors. An equivalent but less compact expression for these variance elements has been presented in [14] for the special case of uncorrelated sources (i.e., diagonal P matrices).

2) CRB Covariance Matrix: The conditional (also called deterministic) Cramér–Rao lower bound on the covariance matrix of any (asymptotically) unbiased estimator of θ is, for large N, given by

$$C_{\text{CR}} = \frac{\sigma}{2N}\left\{\text{Re}\left[H \odot P^T\right]\right\}^{-1}. \quad (2.9)$$

A formula for the CRB matrix which is valid for any value of N, was also derived in [1]. However, the commonly used estimators for θ, such as MUSIC and MLE, are complicated nonlinear functions of the data vector and their statistical behavior for "small" N appears difficult to establish. In the study of the statistical efficiency of these estimators we will thus use the (asymptotic) formula (2.9) for the CRB.

3) Relationship Between MUSIC and MLE: The MUSIC estimator is a large-sample (for $N >> 0$) realization of the MLE if and only if the covariance matrix P is *diagonal*.

4) Statistical Efficiency of MUSIC: For diagonal P, it holds that

$$[C_{\text{MU}}]_{ii} \geq [C_{\text{CR}}]_{ii} \quad (2.10)$$

where *equality* holds in the limit as m increases, if and only if

$$a^*(\omega)\, a(\omega) \to \infty \quad \text{as } m \to \infty. \quad (2.11)$$

For nondiagonal P, $[C_{\text{MU}}]_{ii}$ is strictly greater than $[C_{\text{CR}}]_{ii}$.

Note that the condition (2.11) on $a(\omega)$ is satisfied in several important applications of the model (1.1), such as direction finding with uniform linear arrays, and estimation of the parameters of undamped exponential signals (see [1]–[9], [15] for details).

Finally, we outline the present paper and its contributions.

D. Outline of the Paper

There are a number of directions in which the results of [1] should be extended in order to obtain a fairly complete image of the topic under discussion:

i) A conceptually simple generalization of the MUSIC estimator is obtained by considering the following function instead of (2.4):

$$f_W(\omega) = a^*(\omega)\,\hat{G}W\hat{G}^*a(\omega) \tag{2.12}$$

where W is a positive semidefinite (weighting) matrix. Minimization of (2.12) with respect to ω gives the so-called weighted MUSIC estimators (see [2], [10]). The statistical behavior of the weighted MUSIC estimators needs to be established, and their performance compared to that of the nonweighted MUSIC.

ii) The study in [1] of the statistical (in)efficiency of MUSIC was only concerned with the diagonal elements $[C_{\text{MU}}]_{ii}$ and $[C_{\text{CR}}]_{ii}$. It should be of interest to compare the whole covariance matrices C_{MU} and C_{CR}. *Inter alia*, this comparison will provide more exact results on the statistical efficiency of the MUSIC estimator.

iii) The covariance matrix of the ML estimator was not derived in [1] (except for the special case of diagonal P matrices, when $C_{\text{ML}} = C_{\text{MU}}$, see 3) above). Derivation of the covariance matrix C_{ML} of the MLE appears to be of significant interest. Among other things, the availability of an expression for C_{ML} will make it possible to study the statistical efficiency of the MLE, and to compare the performances of the MUSIC and MLE in the general case of nondiagonal P matrices.

The aim of this paper is to provide a number of results along the lines outlined above. More exactly, in Section III we establish the covariance matrix of the weighted MUSIC estimator and show that in the class of estimators which minimize (2.12), the nonweighted MUSIC possesses minimum asymptotic variance. Thus, we provide theoretical support to the empirically observed fact that use of a weighting matrix $W \neq I$ in (2.12) cannot improve the accuracy of the MUSIC estimator, in large samples (see, e.g., [2]).

In Section IV we prove that $C_{\text{MU}} \geq C_{\text{CR}}$. Furthermore, we show that for diagonal P and vectors $a(\omega)$ satisfying (2.11), the equality $C_{\text{MU}} = C_{\text{CR}}$ holds in the limit, as m increases. For nondiagonal P we show that the inequality is strict and that the difference $C_{\text{MU}} - C_{\text{CR}}$ may be substantial if the elements of $x(t)$ are nearly colinear. These results again provide theoretical support to a fact observed in simulations: the performance of MUSIC for diagonal P is excellent, but it degrades if P approaches a (nearly) singular matrix.

In Section V we derive the covariance matrix of the ML estimator. Also in that section we prove that $C_{\text{ML}} \geq C_{\text{CR}}$, and that for vectors $a(\omega)$ which satisfy (2.11), the equality $C_{\text{ML}} = C_{\text{CR}}$ holds in the limit as m increases. For vectors $a(\omega)$ which do not satisfy (2.11), the inequality $C_{\text{ML}} \geq C_{\text{CR}}$ is shown to be strict. These results were anticipated in [1] using simple examples and argumentation based on the general properties of the ML estimators. Here, we provide direct algebraic proofs of them.

In Section VI we compare the covariance matrices C_{MU} and C_{ML} of the MUSIC and ML estimators. For diagonal P we prove that $C_{\text{MU}} = C_{\text{ML}}$, thus rediscovering in another way the result of [1] which asserts that MUSIC and MLE coincide asymptotically (for large N) in this case. For nondiagonal P, we show that a generally valid order relation between C_{MU} and C_{ML} does not exist. Quite often, the MLE is expected to offer better performance, but in certain rare cases MUSIC can be superior to the MLE.

In Section VII we present the results of a numerical study of performance, whose aim is to provide a more quantitative comparison of C_{MU}, C_{ML}, and C_{CR} in the case of a two-direction finding application.

Finally, in the Appendices we include some useful results on the Hadamard matrix product, and proofs of the theorems in the paper.

III. Weighted and Nonweighted MUSIC

First, we derive the distribution of the weighted MUSIC estimator.

Theorem 3.1: The estimation errors $\{\hat{\omega}_k - \omega_k\}$ associated with the weighted MUSIC estimator which minimizes (2.12), are asymptotically (for large N) jointly Gaussian distributed with zero means and the following covariance matrix:

$$C_{W\text{MU}} = \frac{\sigma}{2N}(\bar{H} \odot I)^{-1}\,\text{Re}\left\{\tilde{H} \odot (A^*UA)^T\right\} \cdot (\bar{H} \odot I)^{-1} \tag{3.1}$$

where

$$\bar{H} = D^*GWG^*D \tag{3.2a}$$

$$\tilde{H} = D^*GW^2G^*D \tag{3.2b}$$

and all the other quantities have been defined before.

Proof: See Appendix B. ∎

Note that the expression (3.1) of $C_{W\text{MU}}$ holds for both data independent and data dependent W matrices. In the latter case W in (3.2) should be interpreted as the limit of the data dependent weighting matrix as $N \to \infty$ (see the proof in Appendix B).

It is also interesting to note that the previous analysis of the weighted MUSIC encompasses the direction estimators which do not make use of the orthogonality of $a(\omega_i)$ to all the columns of G but only to a certain vector in the range of G. A typical example of such an estimator is the minimum-norm algorithm introduced in [17]. This algorithm corresponds to minimizing the function (2.12) with $W = hh^*$, where h is an $(m - n)$ vector which is such that $\hat{G}h$ is the minimum (Euclidean) norm vector in the column space of $\hat{G}$, whose first component is equal to one. (Theorem 3.1 holds for many other estimators of this form under the weak assumption that h is such that none of the elements of the vector D^*Gh is equal to zero (this ensures that $(\bar{H} \odot I)$ in (3.1) is a nonsingular matrix).)

Next, we show that the minimum asymptotic variance in the class of weighted MUSIC estimators is achieved by the nonweighted MUSIC.

Theorem 3.2: The diagonal elements of $C_{W\text{MU}}$ are greater than the corresponding diagonal elements of C_{MU}

$= C_{WMU}|_{W=I}$, or

$$C_{WMU} \odot I \geq C_{MU} \odot I. \tag{3.3}$$

Proof: See Appendix C. ■

It is an open question whether the inequality (3.3) extends to the whole covariance matrices C_{WMU} and C_{MU}. However, the answer to this question cannot change the conclusion that use of a weighting matrix $W \neq I$ in (2.12) will, for large N, worsen the accuracy of the estimates, instead of improving it. It is perhaps worth stressing once more that the above result is asymptotically valid. In the case of small or medium-sized samples the optimal weighting matrix might be different from $W = I$.

IV. MUSIC and CRB

First, we prove that the covariance matrix C_{MU} is bounded from below by the CRB covariance matric C_{CR}. We present a simple proof of this expected result, which has the virtue of revealing some ways of doing more accurate comparisons between C_{MU} and C_{CR}.

Theorem 4.1: The covariance matrix of the MUSIC estimator can be decomposed additively as follows:

$$C_{MU} = \tilde{C}_{MU} + \bar{C}_{MU} \tag{4.1}$$

where (see (2.6), (2.8))

$$\tilde{C}_{MU} = \frac{\sigma}{2N}(H \odot I)^{-1} \operatorname{Re}[H \odot P^{-T}](H \odot I)^{-1} \tag{4.2}$$

$$\bar{C}_{MU} = \frac{\sigma^2}{2N}(H \odot I)^{-1} \cdot \operatorname{Re}\left\{H \odot [P^{-1}(A^*A)^{-1}P^{-1}]^T\right\}(H \odot I)^{-1}. \tag{4.3}$$

The matrices $\tilde{C}_{MU}$ and $\bar{C}_{MU}$ satisfy

$$\tilde{C}_{MU} \geq C_{CR} \tag{4.4}$$

$$\bar{C}_{MU} \geq 0. \tag{4.5}$$

Proof: The results follow immediately from Lemmas A.1 and A.2 in Appendix A. ■

Next, let us assume that P is *diagonal*. Then, we get

$$\tilde{C}_{MU} = \frac{\sigma}{2N}(H \odot I)^{-1}P^{-1} = C_{CR}. \tag{4.6}$$

If we additionally assume that $a(\omega)$ satisfies (2.11), then it is not difficult to see that C_{MU} tends to $\tilde{C}_{MU}$, as m increases. Indeed, under assumption (2.11), the matrix $(A^*A)^{-1}$ tends to zero as $m \to \infty$, which implies that $\bar{C}_{MU}$ tends to zero faster than $\tilde{C}_{MU}$ (compare the expressions of the two matrices); thus the contribution of $\bar{C}_{MU}$ in the decomposition (4.1) vanishes as m increases. As an example, consider the case of the steering vector

$$a(\omega) = [1 \ e^{i\omega} \ \cdots \ e^{i(m-1)\omega}]^T \tag{4.7}$$

which appears in direction estimation applications using narrow-band uniform linear arrays. For (4.7), we have $a^*(\omega)\,a(\omega) = m$ and, therefore, the condition (2.11) is satisfied. Some straightforward calculations show that the limits, for $m \to \infty$, of the matrices H and (A^*A) corresponding to (4.7), are given by (see [1])

$$\frac{1}{m}(A^*A) \to I \tag{4.8a}$$

$$\frac{1}{m^3}H \to \frac{1}{12}I. \tag{4.8b}$$

Thus, as m increases

$$\tilde{C}_{MU} \to \frac{6\sigma}{Nm^3} I \odot P^{-1} \tag{4.9}$$

and

$$\bar{C}_{MU} \to \frac{6\sigma^2}{Nm^4} I \odot P^{-2}. \tag{4.10}$$

In conclusion, under the assumptions above (P is diagonal and $a(\omega)$ satisfies (2.11)) C_{MU} approaches C_{CR} as m increases, which means that the MUSIC estimator is asymptotically (for large N and m) efficient in this case. This result was also obtained in [1] by different reasoning.

If either the matrix P is nondiagonal or the vector $a(\omega)$ does not satisfy (2.11), then C_{MU} cannot attain C_{CR}, and the MUSIC estimator is statistically inefficient. This is so since the equality in (4.4) cannot hold if P is not diagonal, and $\bar{C}_{MU}$ will be strictly bounded from zero if $a(\omega)$ does not satisfy (2.11).

For nondiagonal P, it is worth noting that the difference $C_{MU} - C_{CR}$ may be quite large if P is nearly singular. We illustrate this fact by considering once more the case of the steering vector $a(\omega)$ given by (4.7). The corresponding matrices C_{MU} and C_{CR} behave, for large m, as follows (see (4.8)–(4.10)):

$$C_{MU} \to \frac{6\sigma}{Nm^3}(I \odot P^{-1}) \tag{4.11}$$

$$C_{CR} \to \frac{6\sigma}{Nm^3}(I \odot P)^{-1}. \tag{4.12}$$

Thus, the difference matrix $C_{MU} - C_{CR}$ may indeed take very large values if P is nearly singular. As a simple example, for

$$P = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \qquad |\rho| < 1$$

we obtain from (4.11) and (4.12)

$$C_{MU} - C_{CR} \to \frac{6\sigma}{Nm^3}\frac{\rho^2}{1-\rho^2} I$$

which increases without bound as $|\rho|$ approaches one.

In the above analysis it was assumed that the signal-to-noise ratio (SNR) is finite (the SNR for the ith source/signal is defined as P_{ii}/σ). If the SNR's of all signals are very large then $\bar{C}_{MU}$ is much less than $\tilde{C}_{MU}$ and can be neglected in (4.1). If in addition P is diagonal then it follows from (4.6) that as SNR $\to \infty$, $C_{MU} \to \tilde{C}_{MU} = C_{CR}$. This observation confirms a result established in [14] by a different method.

V. MLE AND CRB

First, we derive the asymptotic (for large N) distribution of the ML estimator. Derivation of the MLE distribution from (2.5), by the usual technique of expanding the corresponding gradient equation $F'(\hat{\theta}) = 0$ in a Taylor series around the true parameter values, appears somewhat difficult. To obtain a simpler derivation we make use of a result proven in [1] which states that a large-sample (for $N >> 0$) realization of the MLE is given by the minimizer of the following function:

$$h(\theta) = \text{tr}\,[A^*\hat{G}\hat{G}^*AP]. \tag{5.1}$$

Observe that $h(\theta)$ depends on θ in a simpler way than $F(\theta)$ does.

Theorem 5.1: The estimation errors $\{\hat{\omega}_i - \omega_i\}$ of the ML estimator are asymptotically (for large N) jointly Gaussian distributed, with zero means and a covariance matrix given by

$$C_{ML} = \frac{\sigma}{2N}\left[\text{Re}\,(H \odot P^T)\right]^{-1} \cdot \left\{\text{Re}\left[H \odot (PA^*UAP)^T\right]\right\}\left[\text{Re}\,(H \odot P^T)\right]^{-1} \tag{5.2}$$

(all the quantities appearing in (5.2) have been defined previously).

Proof: See Appendix D. ■

Next, we show a simple relationship between the covariance matrix of the MLE and the CRB.

Theorem 5.2: The covariance matrix C_{ML} can be decomposed additively as follows:

$$C_{ML} = C_{CR} + \bar{C}_{ML} \tag{5.3}$$

where

$$\bar{C}_{ML} = \frac{\sigma^2}{2N}\left[\text{Re}\,(H \odot P^T)\right]^{-1} \cdot \left\{\text{Re}\left[H \odot (A^*A)^{-T}\right]\right\}\left[\text{Re}\,(H \odot P^T)\right]^{-1}. \tag{5.4}$$

Proof: Inserting the expression (2.8) of A^*UA into (5.2), we readily get the additive decomposition (5.3), (5.4). ■

The simple result of Theorem 5.2 has several immediate implications:

a) Since the matrix $\bar{C}_{ML}$ is positive semidefinite (cf. Lemma A.1), it follows that $C_{ML} \geq C_{CR}$ (as expected).

b) For vectors $a(\omega)$ which do not satisfy (2.11) and for finite SNR, $\bar{C}_{ML}$ is strictly bounded from zero and, therefore, the MLE cannot achieve the CRB.

c) For vectors $a(\omega)$ which satisfy (2.11), $\bar{C}_{ML}$ goes to zero faster than C_{CR} when m increases (compare the expressions of these two matrices). Thus, in this case the MLE is asymptotically (for large N and m) statistically efficient. Note, however, that for "small or medium" values of m (and finite SNR), $\bar{C}_{ML}$ is different from zero and, thus, the MLE is statistically inefficient. Also note that C_{CR} is inversely proportional to SNR, while $\bar{C}_{ML}$ is inversely proportional to $(\text{SNR})^2$. Thus, for a given finite m, the C_{ML} approaches the C_{CR} as the signal-to-noise ratio (SNR) increases. These properties of the MLE are similar to the analogous properties of the MUSIC but here they hold for arbitrary P matrices.

The results a)–c) above reinforce some conclusions of the analysis of the MLE, obtained in [1] by less direct argumentation. The statistical inefficiency of the MLE in the case of finite m (and finite SNR), appears to be a rather unusual result which leaves open the question whether estimators better than MLE could exist (see [16] for a study of this problem). In this light, comparison between the MLE and the (much) simpler computationally MUSIC estimator becomes of significant interest. This aspect is addressed in the next section.

VI. MUSIC AND MLE

For *diagonal P* matrices we simply have

$$C_{MU} = \frac{\sigma}{2N}(H \odot I)^{-1}\,\text{Re}\left[H \odot (A^*UA)^T\right] \cdot (H \odot I)^{-1} = C_{MU}. \tag{6.1}$$

The equality above between C_{ML} and C_{MU} was expected in view of the fact that, for diagonal P, the MLE and MUSIC are known to coincide as N increases [1] (we stress that the equality (6.1) holds for all $m > n$).

Next, consider the more general case of *nondiagonal P* matrices. If either the SNR is "large enough" or the vector $a(\omega)$ satisfies (2.11) and m is "sufficiently large," then we have (according to the analysis in Sections IV and V): $C_{ML} = C_{CR} < C_{MU}$. However, in all other cases $C_{ML} > C_{CR}$, which means that the inequality $C_{ML} \geq C_{MU}$ is not *a priori* excluded. In fact, this inequality can really hold true, as we show in the following.

From (4.1) and (5.3), we get

$$\frac{2N}{\sigma}(C_{MU} - C_{ML}) = \frac{2N}{\sigma}(\tilde{C}_{MU} - C_{CR}) + \frac{2N}{\sigma}(\bar{C}_{MU} - \bar{C}_{ML}). \tag{6.2}$$

The first term on the right-hand side of (6.2) is positive semidefinite and independent of σ, while the second term is proportional to σ. We thus conclude that a sufficient and necessary condition for the inequality $C_{MU} \geq C_{ML}$ to hold, is

$$\bar{C}_{MU} \geq \bar{C}_{ML}. \tag{6.3}$$

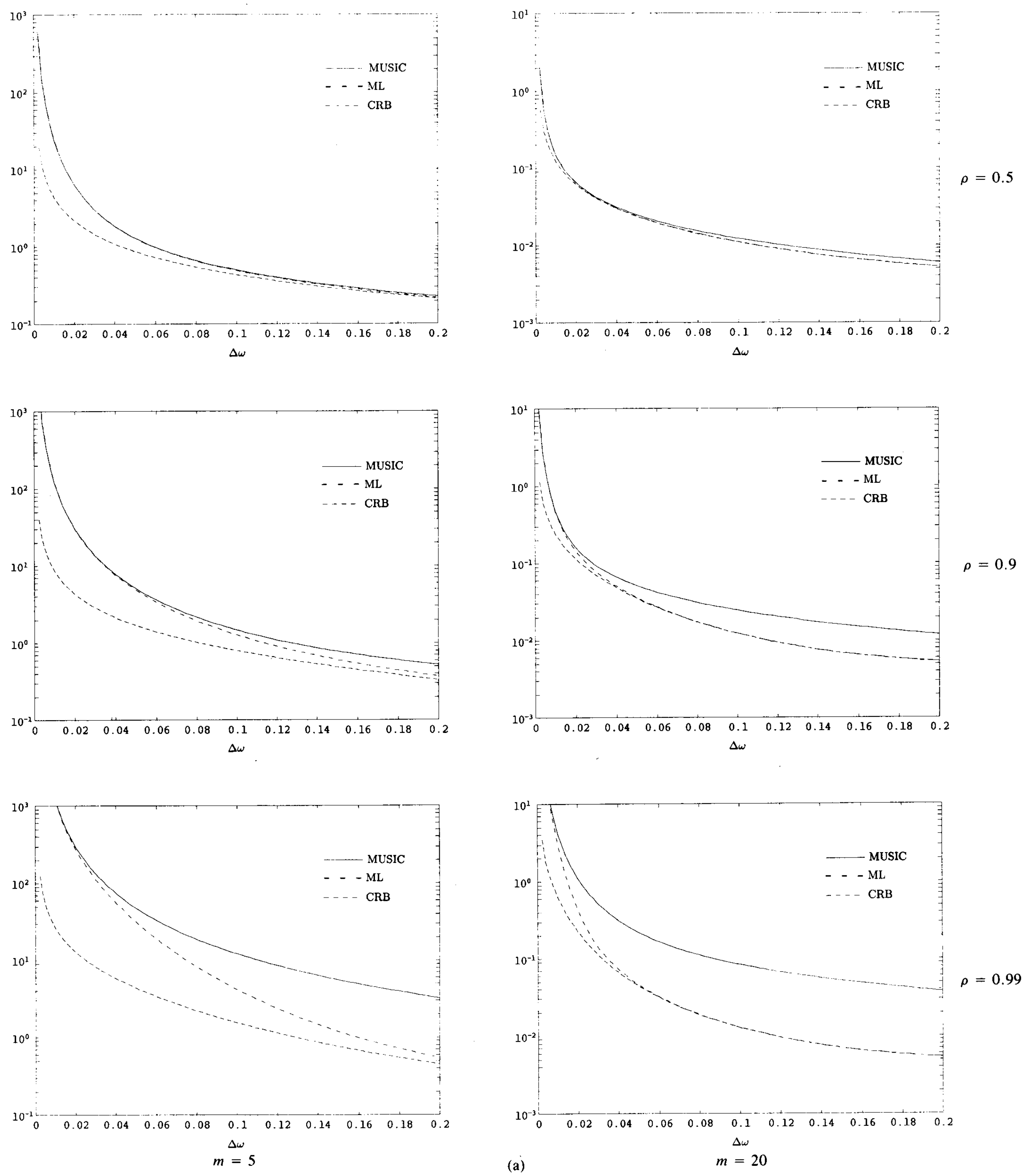
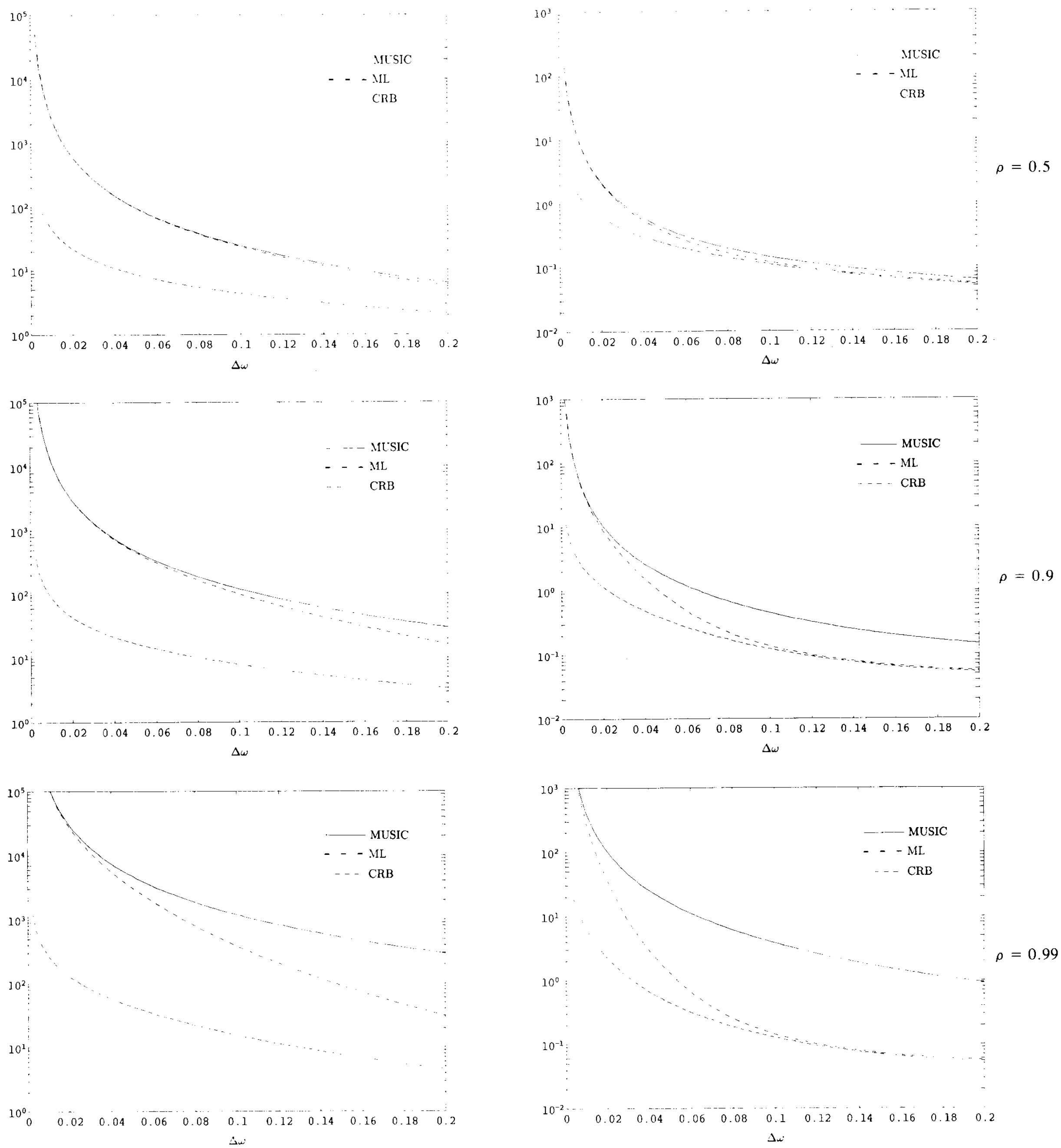


Fig. 1. Comparison of the normalized estimation error standard deviations of MUSIC, MLE, and CRB for the shown values of ρ, m, and SNR, and for varying $\Delta\omega$. (a) $\sigma = 0.01$ (SNR = 20 dB). (*Continued on next page.*)

The sufficiency of (6.3) is evident. The necessity follows from the observation that the second term in (6.2) is proportional to σ, while the first does not depend on σ. Thus, if (6.3) does not hold then one can choose σ sufficiently large such that $C_{MU} \geq C_{ML}$ does not hold either. In the following we provide a simple counterexample to (6.3).

Let $P = (A^*A)^{-1/2}$ (the (Hermitian) positive definite square root of the inverse matrix $(A^*A)^{-1}$).[1] Then, we

[1]Let H be a (Hermitian) positive definite matrix. Then we can write $H = Q^*\Lambda Q$, where Q is unitary and Λ is a positive definite real-valued diagonal matrix. Define $P = Q^*\Lambda^{1/2}Q$ and observe that $P^2 = H$. Thus, P is a (Hermitian) positive definite square root of H.



(b)

Fig. 1. (con't.) (b) $\sigma = 1$ (SNR = 0 dB).

obtain

$$\overline{C}_{\text{MU}} = \frac{\sigma^2}{2N} (H \odot I)^{-1}$$

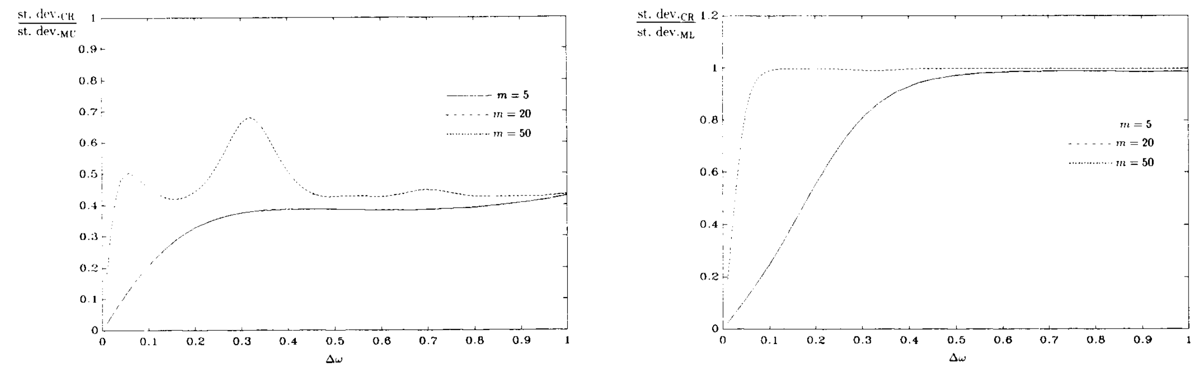
and

$$\overline{C}_{\text{ML}} = \frac{\sigma^2}{2N} [\text{Re}\,(H \odot P^T)]^{-1} [\text{Re}\,(H \odot P^T P^T)] \cdot [\text{Re}\,(H \odot P^T)]^{-1}$$

which satisfy the following inequality (cf. Lemma A.2 in Appendix A)

$$\overline{C}_{\text{ML}} \geq \overline{C}_{\text{MU}}.$$

Thus, for the above choice of the matrix P (which is special but is not peculiar), the MUSIC estimator will be more accurate than the MLE for "sufficiently large" values of



(a) (b)

Fig. 2. The ratios (7.1) for $\rho = 0.9$, $\sigma = 0.1$ (SNR = 10 dB), $m = 5$, 20, and 50, and varying $\Delta\omega$. (a) Standard deviation (CRB)/standard deviation (MUSIC). (b) Standard deviation (CRB)/standard deviation (ML).

σ! The above result makes a detailed study of the performances of MUSIC and MLE of considerable interest. In the next section we present a numerical study of performance for the case of a two-direction-finding application.

VII. MUSIC, MLE, and CRB: A Numerical Comparison

Consider the situation of two narrow-band plane waves impinging on a uniform linear array of m sensors. Then $n = 2$ and $a(\omega)$ is given by (4.7) (see [2], [6], [10]). Let

$$P = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \qquad |\rho| < 1.$$

The square roots of the variance elements $[C_{\mathrm{MU}}]_{11} = [C_{\mathrm{MU}}]_{22}$, $[C_{\mathrm{ML}}]_{11} = [C_{\mathrm{ML}}]_{22}$ and $[C_{\mathrm{CR}}]_{11} = [C_{\mathrm{CR}}]_{22}$ times N have been evaluated for several values of ρ, m, and σ, and for varying (ω_1, ω_2). Note that in the present case these variance elements depend on $\Delta\omega = |\omega_1 - \omega_2|$ only. Thus they have been evaluated for $\Delta\omega \in (0, \pi)$. However, only the variance values for $\Delta\omega \in (0, 0.2)$ [rad] will be shown since the case of practical interest is of closely spaced sources.

Fig. 1 shows the results obtained as outlined above. These results confirm that for highly correlated sources the performance of MUSIC degrades significantly compared to MLE and CRB; whereas for weakly correlated sources MUSIC and MLE provide similar performance that is close to the CRB for reasonably high SNR. Fig. 1 also shows the extent to which a degradation of estimation accuracy caused by a decrease in SNR or $\Delta\omega$ may be compensated for by an increase in m. This aspect is further studied in Fig. 2 where the ratios

$$\text{st. dev. (CRB)/st. dev. (MUSIC)}$$
$$\text{and st. dev. (CRB)/st. dev. (ML)} \qquad (7.1)$$

are plotted versus $\Delta\omega$, for $\rho = 0.9$, $\sigma = 0.1$ (SNR = 10 dB), and several values of m. It is seen from this figure that ML achieves CRB as m increases (the smaller $\Delta\omega$ (or SNR), the larger m required for MLE to achieve CRB). However, MUSIC variance for correlated sources ($\rho \neq 0$) cannot achieve CRB by increasing m and, in fact, MUSIC statistical efficiency does not necessarily increase when m increases.

VIII. Conclusions

The present paper and its companion [1] provide a fairly complete image of the performance and statistical efficiency of the MUSIC and conditional ML estimators. The main contributions of this paper were outlined in Section II and will not be repeated here. In this concluding section we only reemphasize the importance of the explicit expressions for the covariance matrices of the ML and MUSIC estimators and of the conditional CRB, derived in this paper and in [1]. Using these expressions we were able to perform general analytic comparisons between the methods under discussion, as well as rapid numerical evaluations of performance in specific situations. A main result which emerged from the present study concerns the fact that for small or medium values of m and SNR (which is the case of practical interest), the (conditional) MLE is not the most accurate estimator. This fact opens the possibility that other estimators which are computationally simpler and statistically more accurate than the MLE could exist. (Computationally, MUSIC is such a simpler estimator but is seldomly more accurate than MLE). The search for such estimators remains a research topic of considerable practical and theoretical interest (see [16]).

Appendix A
Some Useful Results on the Hadamard Product

Lemma A.1: Let $A, B \in \mathbf{C}^{n \times n}$ be two (Hermitian) positive semidefinite matrices. Then the matrix $A \odot B$ is positive semidefinite too.

Proof: This result (for the real-valued case) is attributed to Schur (see [11]). For completeness, we provide a simple proof of it.

Since the matrix B is positive semidefinite, it can be written as $B = W^*W$. Let w_k denote the kth column of W.

Then

$$(A \odot B)_{ij} = A_{ij} w_i^* w_j$$

and, therefore, we can write

$$A \odot B = \begin{bmatrix} w_1^* & & 0 \\ & \ddots & \\ 0 & & w_n^* \end{bmatrix} \begin{bmatrix} A_{11}I & \cdots & A_{1n}I \\ \vdots & & \vdots \\ A_{n1}I & \cdots & A_{nn}I \end{bmatrix} \cdot \begin{bmatrix} w_1 & & 0 \\ & \ddots & \\ 0 & & w_n \end{bmatrix} = \overline{W}^*(A \otimes I)\overline{W}$$

where $\otimes$ denotes the Kronecker product (see, e.g., [11], [12]), and

$$\overline{W} \triangleq \begin{bmatrix} w_1 & & 0 \\ & \ddots & \\ 0 & & w_n \end{bmatrix}.$$

Since the matrix $A \otimes I$ is positive semidefinite (by the properties of the Kronecker product, e.g., [11]), the proof is finished. ■

Lemma A.2: Let A, B, and C be (Hermitian) positive semidefinite matrices (B can be Hermitian only). Then, assuming that the inverses appearing below exist, it holds that

$$[\text{Re}\,(A \odot B)]^{-1}[\text{Re}\,(A \odot C)][\text{Re}\,(A \odot B)]^{-1} \geq \left\{\text{Re}\,[A \odot (BC^{-1}B)]\right\}^{-1}. \tag{A.1}$$

Proof: The inequality (A.1) is equivalent to

$$\text{Re}\,[A \odot (BC^{-1}B)] - [\text{Re}\,(A \odot B)][\text{Re}\,(A \odot C)]^{-1} \cdot [\text{Re}\,(A \odot B)] \geq 0$$

which, in turn, is equivalent to

$$\text{Re}\begin{bmatrix} A \odot BC^{-1}B & A \odot B \\ A \odot B & A \odot C \end{bmatrix} = \text{Re}\left\{\begin{bmatrix} A & A \\ A & A \end{bmatrix} \odot \begin{bmatrix} BC^{-1}B & B \\ B & C \end{bmatrix}\right\} \geq 0.$$

Since the matrices

$$\begin{bmatrix} A & A \\ A & A \end{bmatrix} = \begin{bmatrix} I \\ I \end{bmatrix} A [I \;\; I]$$

and

$$\begin{bmatrix} BC^{-1}B & B \\ B & C \end{bmatrix} = \begin{bmatrix} BC^{-1} \\ I \end{bmatrix} C [C^{-1}B \;\; I]$$

are both positive semidefinite, the assertion of the lemma follows from Lemma A.1 and the fact that Re $(H) \geq 0$ if $H \geq 0$. ■

Corollary: Let P be a (Hermitian) positive definite matrix. Then

$$(I \odot P^{-1}) \geq (I \odot P)^{-1}.$$

Proof: Set $A = B = I$ and $C = P^{-1}$ in Lemma A.2. ■

Appendix B
Proof of Theorem 3.1

To prove the theorem, we will use the following result proved in [1] (also, see [2]).

Lemma B.1: The orthogonal projections of $\{\hat{g}_i\}$ onto the column space of S are asymptotically (for large N) jointly Gaussian distributed with zero means and the following variances—covariances:

$$E(SS^*\hat{g}_i)(SS^*\hat{g}_j)^* = \frac{\sigma}{N} U\delta_{i,j}$$

$$E(SS^*\hat{g}_i)(SS^*\hat{g}_j)^T = 0 \qquad \text{for all } i, j$$

where the matrix U is defined (implicitly) by (2.8). ■

Now we turn to the proof of Theorem 3.1, which will closely follow the proof for the nonweighted case in [1]. As $\hat{\omega}_i$ is a minimum point of $f_W(\omega)$, we must have

$$f'_W(\hat{\omega}_i) = 2\,\text{Re}\,[a^*(\hat{\omega}_i)\,\hat{G}W\hat{G}^*d(\hat{\omega}_i)] = 0.$$

Since $\hat{\omega}_i$ is a consistent estimate of ω_i ([1], [2]) we can write (for large N)

$$\begin{aligned} 0 = f'_W(\hat{\omega}_i) &\simeq f'_W(\omega_i) + f''_W(\omega_i)(\hat{\omega}_i - \omega_i) \\ &= 2\,\text{Re}\,[a^*(\omega_i)\,\hat{G}W\hat{G}^*d(\omega_i)] \\ &\quad + 2\,\text{Re}\,[d^*(\omega_i)\,\hat{G}W\hat{G}^*d(\omega_i) \\ &\quad + a^*(\omega_i)\,\hat{G}W\hat{G}^*d'(\omega_i)](\hat{\omega}_i - \omega_i) \\ &\simeq 2\,\text{Re}\,[a^*(\omega_i)\,\hat{G}WG^*d(\omega_i)] \\ &\quad + 2[d^*(\omega_i)\,GWG^*d(\omega_i)](\hat{\omega}_i - \omega_i). \end{aligned} \tag{B.1}$$

Remark: The replacement of $\hat{G}^*$ by G^* in the above calculation is a subtle issue discussed in some detail in [1]. The complication is caused by the fact that $\hat{G}$ is unique (with probability one) whereas G is not (any multiplication of G on the right by a unitary matrix produces another set of noise eigenvectors, i.e., another G matrix). Thus, the problem is which G should be used instead of $\hat{G}$. Consider the orthogonal projection of $\hat{G}$ on the subspace spanned by the columns of G: $G(G^*\hat{G})$. The distribution of $G^*\hat{G}$ can be shown to be independent of the distribution of the other terms in (B.1) (see [1]). Thus, the matrix $G^*\hat{G}$ can be treated as given, and since it is asymptotically unitary (as can be readily verified, see [1]) it follows that asymptotically $G(G^*\hat{G})$ forms a matrix of noise eigenvectors. This matrix can thus be redenoted by G and used in (B.1) instead of $\hat{G}$. Note that the difference $\hat{G} - GG^*\hat{G} = (I - GG^*)\hat{G} = SS^*\hat{G}$ tends to zero as $N \to \infty$. ■

Next note that

$$a^*(\omega_i)\,\hat{G}WG^*d(\omega_i) = a^*(\omega_i)\,SS^*\hat{G}WG^*d(\omega_i) = \sum_{k=1}^{m-n}[\tilde{g}_k^* d(\omega_i)][a^*(\omega_i)SS^*\hat{g}_k] \quad (B.2)$$

where $[\,\tilde{g}_1 \cdots \tilde{g}_{m-n}] \triangleq GW$. Thus, from (B.1) and (B.2) we obtain (neglecting the higher order terms)

$$(\hat{\omega}_i - \omega_i) = -\text{Re}\left\{\sum_{k=1}^{m-n}[\tilde{g}_k^* d(\omega_i)][a^*(\omega_i)SS^*\hat{g}_k]\right\}\Big/ (\overline{H} \odot I)_{ii}. \quad (B.3)$$

The asymptotic zero-mean joint Gaussian distribution of $\{\hat{\omega}_i - \omega_i\}$ follows from (B.3) and Lemma B.1. To evaluate the covariance matrix of the distribution, we make use of the following simple result: for two scalar variables u and v, it holds that

$$\text{Re}\,(u) \cdot \text{Re}\,(v) = \tfrac{1}{2}[\text{Re}\,(uv) + \text{Re}\,(uv^*)].$$

Let

$$u = \sum_{k=1}^{m-n}[\tilde{g}_k^* d(\omega_i)][a^*(\omega_i)SS^*\hat{g}_k]$$

and let v be defined similarly to u but with ω_i replaced by ω_j. Using Lemma B.1, we obtain

$$Euv = \sum_{k=1}^{m-n}\sum_{p=1}^{m-n}[\tilde{g}_k^* d(\omega_i)][\tilde{g}_p^* d(\omega_j)] \cdot E\left\{[a^*(\omega_i)SS^*\hat{g}_k] \cdot [a^*(\omega_j)SS^*\hat{g}_p]^T\right\} = 0$$

and

$$\begin{aligned} Euv^* &= \sum_{k=1}^{m-n}\sum_{p=1}^{m-n}[d^*(\omega_j)\tilde{g}_p\tilde{g}_k^* d(\omega_i)] \cdot [a^*(\omega_i)\,E(SS^*\hat{g}_k)(SS^*\hat{g}_p)^*\,a(\omega_j)] \\ &= \frac{\sigma}{N}\sum_{k=1}^{m-n} d^*(\omega_j)\tilde{g}_k\tilde{g}_k^* d(\omega_i)\;a^*(\omega_i)Ua(\omega_j) \\ &= \frac{\sigma}{N}[d^*(\omega_j)GW^2G^*d(\omega_i)][a^*(\omega_i)Ua(\omega_j)]. \end{aligned}$$

Thus, we have

$$[C_{WMU}]_{ij} = \frac{\sigma}{2N}\text{Re}\,[\tilde{H}_{ij}(A^*UA)_{ji}]/[(\overline{H} \odot I)_{ii}(\overline{H} \odot I)_{jj}]$$

and the proof is complete.

Appendix C
Proof of Theorem 3.2

Since

$$C_{WMU} \odot I - C_{MU} \odot I = \frac{\sigma}{2N}\left\{[(\overline{H} \odot I)^{-1}(\tilde{H} \odot I)(\overline{H} \odot I)^{-1} - (H \odot I)^{-1}] \odot (A^*UA)^T\right\}$$

the inequality (3.3) will follow from Lemma A.1 if we can prove that

$$(I \odot \overline{H})^{-1}(I \odot \tilde{H})(I \odot \overline{H})^{-1} \geq (I \odot H)^{-1}$$

or, equivalently,

$$(I \odot \overline{H})(I \odot \tilde{H})^{-1}(I \odot \overline{H}) \leq (I \odot H)$$

or yet

$$\begin{bmatrix} I \odot H & I \odot \overline{H} \\ I \odot \overline{H} & I \odot \tilde{H} \end{bmatrix} = \begin{bmatrix} I & I \\ I & I \end{bmatrix} \odot \begin{bmatrix} H & \overline{H} \\ \overline{H} & \tilde{H} \end{bmatrix} = \left\{\begin{bmatrix} I \\ I \end{bmatrix}[I \quad I]\right\} \odot \left\{\begin{bmatrix} D^*G \\ D^*GW \end{bmatrix} \cdot [G^*D \quad WG^*D]\right\} \geq 0. \quad (C.1)$$

However, using Lemma A.1 once more we conclude readily that (C.1) holds, and the proof is finished.

Appendix D
Proof of Theorem 5.1

It follows from (5.1) and the consistency of the MLE [1] that, for large N, we can write (neglecting higher order terms)

$$0 = h'(\hat{\theta}) = h'(\theta) + h''(\theta)(\hat{\theta} - \theta). \quad (D.1)$$

Let p_{ij} denote the i, j element and P_i the i column of the matrix P. Some straightforward calculations then give

$$\frac{\partial h(\theta)}{\partial \omega_i} = \text{tr}\left[\frac{\partial A^*}{\partial \omega_i}\hat{G}\hat{G}^*AP\right] + \text{tr}\left[PA^*\hat{G}\hat{G}^*\frac{\partial A}{\partial \omega_i}\right] = 2\,\text{Re}\,\{d^*(\omega_i)\hat{G}\hat{G}^*AP_i\} \quad (D.2)$$

and

$$\frac{\partial^2 h(\theta)}{\partial \omega_i \partial \omega_j} = 2\,\text{Re}\,\{d^*(\omega_i)\,\hat{G}\hat{G}^*d(\omega_j)p_{ji}\} + 2\,\text{Re}\left\{\frac{\partial d^*(\omega_i)}{\partial \omega_j}\hat{G}\hat{G}^*AP_i\right\}. \quad (D.3)$$

The sample projector matrix $\hat{G}\hat{G}^*$ which appears in the expression above of the Hessian matrix $h''(\theta)$, can be replaced by the true projector matrix GG^*, without affecting the dominant term in (D.1). This observation implies that the second term in (D.3) can be neglected, since $G^*A = 0$ (as can readily be verified, see (2.2)). Similarly, we can replace the first $\hat{G}$ in (D.2) by G (see the remark after equation (B.1)). This replacement will only introduce a higher order term in (D.1), since $\hat{G}^*A = \hat{G}^*SS^*A$ tends to zero as $N \to \infty$.

Using the observations above, we get from (D.1)–(D.3)

$$\hat{\theta} - \theta = -\left[\mathrm{Re}\,(H \odot P^T)\right]^{-1} \mathrm{Re}\,(\mu) \qquad \text{(D.4)}$$

where the ith component of the vector μ is given by

$$\mu_i = \left[(AP)_i^* \hat{G} G^* d(\omega_i)\right] \qquad \text{(D.5)}$$

(here $(AP)_i$ denotes the ith column of the matrix AP).

Next, observe that μ_i is similar to the quantity (B.2) (with $W = I$) in Appendix B. This similarity and the calculations in Appendix B can be exploited to conclude readily that $\mathrm{Re}(\mu)$ is asymptotically (for large N) Gaussian distributed with zero mean and the following covariance matrix:

$$E\left[\mathrm{Re}(\mu) \cdot \mathrm{Re}(\mu^*)\right] = \frac{\sigma}{2N} \mathrm{Re}\left\{H \odot (PA^*UAP)^T\right\}. \qquad \text{(D.6)}$$

The assertion of the theorem now follows from (D.4) and (D.6).

References

[1] P. Stoica and A. Nehorai, "MUSIC, maximum likelihood, and Cramér-Rao bound," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, pp. 720–741, May 1989.

[2] K. Sharman and T. S. Durrani, "A comparative study of modern eigenstructure methods for bearing estimation—A new high performance approach," in *Proc. 25th IEEE Conf. Dec. Contr.* (Athens, Greece), Dec. 1986, pp. 1737–1742.

[3] Y. Bresler and A. Macovski, "Exact maximum likelihood parameter estimation of superimposed exponential signals in noise," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-34, pp. 1081–1089, Oct. 1986.

[4] M. Wax and T. Kailath, "Detection of signals by information theoretic criteria," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-33, pp. 387–392, Apr. 1985.

[5] G. Bienvenu and L. Kopp, "Optimality of high resolution array processing using the eigensystem approach," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-31, pp. 1235–1248, Oct. 1983.

[6] H. Wang and M. Kaveh, "On the performance of signal-subspace processing, Part I: Narrow-band systems," *IEEE Trans Acoust., Speech, Signal Processing*, vol. ASSP-34, pp. 1201–1209, Oct. 1986.

[7] R. Kumaresan and A. K. Shaw, "High resolution bearing estimation without eigen decomposition," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing* (Tampa, FL), Mar. 1985, pp. 576–579.

[8] R. Kumaresan, L. L. Scharf, and A. K. Shaw, "An algorithm for pole-zero modeling and spectral analysis," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-34, pp. 637–640, June 1986.

[9] R. O. Schmidt, "Multiple emitter location and signal parameter estimation," in *Proc. RADC Spectral Estimation Workshop* (Rome, NY), 1979, pp. 243–258.

[10] D. Johnson, "The application of spectral estimation methods to bearing estimation problems," *Proc. IEEE*, vol. 70, pp. 1018–1028, Sept. 1982.

[11] R. Bellman, *Introduction to Matrix Analysis*, 2nd ed. New York: McGraw-Hill, 1972.

[12] T. Söderström and P. Stoica, *System Identification*. London, U.K.: Prentice-Hall, 1989.

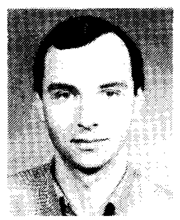
[13] D. Starer and A. Nehorai, "Maximum likelihood estimation of exponential signals in noise using a Newton algorithm," in *Proc. 4th ASSP Workshop Spectrum Estimation, Modeling*, (Minneapolis, MN), Aug. 1988, pp. 240–245.

[14] B. Porat and B. Friedlander, "Analysis of the asymptotic relative efficiency of the MUSIC algorithm," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-36, Apr. 1988, pp. 532–544.

[15] M. Wax, "Detection and estimation of superimposed signals," Ph.D. dissertation, Stanford University, Stanford, CA, 1985.

[16] P. Stoica and A. Nehorai, "Performance study of conditional and unconditional direction-of-arrival estimation," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 38, no. 10, pp. 1783–1795, Oct. 1990.

[17] R. Kumaresan and D. W. Tufts, "Estimating the angles of arrival of multiple plane waves," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-19, pp. 134–139, Jan. 1983.



Petre Stoica received the M.Sc. and Ph.D. degrees, both in automatic control, in 1972 and 1979, respectively.

Since 1972 he has been with the Department of Automatic Control, the Polytechnic Institute of Bucharest, Romania. His research interests include various aspects of system identification, time series analysis, and signal processing. For papers on these topics he received three national prizes. He is coauthor of five books, the most recent being *System Identification* (Englewood Cliffs, NJ: Prentice-Hall, 1989).

Dr. Stoica is a member of the Board of Directors of the Time Series Analysis and Forecasting (TSA&F) Society, and an Associate Editor for the *Journal of Forecasting*. He was given the Member of TSA&F (MTSA&F) honors award. He was corecipient with A. Nehorai of the IEEE Signal Processing Society's Senior Award for the paper, "Music, Maximum Likelihood, and the Cramér-Rao Bound," published in 1989.

Arye Nehorai (S'80–M'83–SM'90) received the B.Sc. and M.Sc. degrees in electrical engineering from the Technion—Israel Institute of Technology, in 1976 and 1979, respectively, and the Ph.D. degree in electrical engineering from Stanford University, Stanford, CA, in 1983.

From 1983 to 1984, he was a Research Associate at Stanford University. From 1984 to 1985, he was a Research Engineer at Systems Control Technology Inc., in Palo Alto, CA. Since 1985 he has been with the Department of Electrical Engineering and the Applied Mathematics Program at Yale University, New Haven, CT, where he is an Associate Professor. During parts of 1989 and 1990 he held Visiting Professorships at Uppsala University, Sweden, and the Technion, Israel. His current areas of interest are adaptive filtering, sensor array processing, system identification, and biomedical engineering.

Dr. Nehorai is an Associate Editor of *Circuits, Systems and Signal Processing* and was an Associate Editor of the IEEE TRANSACTIONS ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING. He is a member of the Technical Committee on Spectrum Estimation and Modeling and the Education Committee in the IEEE Signal Processing Society, and is the Chairman of the Connecticut IEEE Signal Processing Chapter. He was corecipient with P. Stoica of the IEEE Signal Processing Society's Senior Award for the paper "MUSIC, Maximum Likelihood, and Cramér-Rao Bound," published in 1989. He is a member of Sigma Xi.