

Computationally Efficient Angle Estimation for Signals with Known Waveforms

Jian Li, *Member, IEEE*, Bijit Halder, Petre Stoica, *Fellow, IEEE*, and Mats Viberg, *Member, IEEE*

Abstract—This paper presents a large sample decoupled maximum likelihood (DEML) angle estimator for uncorrelated narrowband plane waves with known waveforms and unknown amplitudes arriving at a sensor array in the presence of unknown and arbitrary spatially colored noise. The DEML estimator decouples the multidimensional problem of the exact ML estimator to a set of 1-D problems and, hence, is computationally efficient. We shall derive the asymptotic statistical performance of the DEML estimator and compare the performance with its Cramér–Rao bound (CRB), i.e., the best possible performance for the class of asymptotically unbiased estimators. We will show that the DEML estimator is asymptotically statistically efficient for uncorrelated signals with known waveforms. We will also show that for moderately correlated signals with known waveforms, the DEML estimator is no longer a large sample maximum likelihood (ML) estimator, but the DEML estimator may still be used for angle estimation, and the performance degradation relative to the CRB is small. We shall show that the DEML estimator can also be used to estimate the arrival angles of desired signals with known waveforms in the presence of interfering or jamming signals by modeling the interfering or jamming signals as random processes with an unknown spatial covariance matrix. Finally, several numerical examples showing the performance of the DEML estimator are presented in this paper.

I. INTRODUCTION

MANY high-resolution algorithms for estimating the angles of arrival (AOA's) of signals incident on an array of sensors have been devised in the past few decades. These research activities are mainly motivated by military applications including the source localization applications in radar and sonar. In these conventional applications, the incident signals are usually unknown. The subspace-based algorithms, such as MUSIC [1] and ESPRIT [2], are developed without considering any knowledge of the incident signals, except for some general statistical properties such as the second-order ergodicity. Conditional or unconditional estimators [3] are also devised for such applications. The unconditional estimators,

such as the unconditional maximum likelihood (ML) estimators [3]–[5], model the unknown signals as random processes. The conditional estimators, such as the conditional ML estimators [6]–[10], model the unknown signals as unknown deterministic parameters.

As the political environment changes, however, more attention is now being paid to applying array signal processing techniques to civilian applications including friendly communications. For example, an antenna array may be used as a receiver in a communication system to enhance its communications capability. A distinguishing feature of friendly communications is that certain *a priori* knowledge on the impinging signals is available to its receiver. This *a priori* information may or may not be explicit. For example, in a digital communication system, the modulation format of the transmitted signals is known to the receiver, although the actual transmitted symbol stream is unknown. In a packet radio communication system or a mobile communication system, a known preamble may be added to the message for training purposes [11]–[13]. Such extra information may be exploited to enhance the accuracy of the AOA estimates and may be used to simplify the computational complexity of the estimation algorithms. Several algorithms have been developed to exploit such extra information including, for example, the special features of cyclostationary signals [14]–[16] and constant modulus signals [17]. Exact ML estimators for signals with known waveforms in the presence of spatially white noise are also considered [18].

The purpose of this paper is to present a large sample decoupled maximum likelihood (DEML) angle estimator for uncorrelated narrowband plane waves with known waveforms and unknown amplitudes arriving at a sensor array in the presence of unknown and arbitrary spatially colored noise. The DEML estimator decouples the multidimensional problem of the exact ML estimator to a set of 1-D problems and, hence, is computationally efficient. We shall derive the asymptotic statistical performance of the DEML estimator and compare the performance with its Cramér–Rao bound (CRB), i.e., the best possible performance for the class of asymptotically unbiased estimators. We shall show that the DEML estimator is asymptotically statistically efficient for uncorrelated signals with known waveforms, which occur often in communication systems. We shall show that for moderately correlated signals with known waveforms, the DEML estimator is no longer a large sample maximum likelihood (ML) estimator, but the DEML estimator may still be used for angle estimation, and the performance degradation relative to the CRB is small. We shall

Manuscript received February 24, 1994; revised February 1, 1995. This work was supported by NSF Grant MIP-9308302, DARPA Contract F49620-93-1-0085 monitored by the AFOSR, and the Swedish Research Council of Engineering Sciences under Contract 93-669. The associate editor coordinating the review of this paper and approving it for publication was Prof. Fu Li. J. Li is with the Department of Electrical Engineering, University of Florida, Gainesville, FL 32611 USA.

B. Halder is with the Information Systems Laboratory, Stanford University, Stanford, CA 94305 USA.

P. Stoica is with the Systems and Control Group, Department of Technology, Uppsala University, Uppsala, Sweden, on leave from the Department of Automatic Control, Polytechnic Institute of Bucharest, Bucharest, Romania.

M. Viberg is with the Information Systems Laboratory, Stanford University, Stanford, CA 94305 USA. He is now with Department of Applied Electronics, Chalmers University of Technology, Gothenburg, Sweden.

IEEE Log Number 9413305.

1053-587X/95\$04.00 © 1995 IEEE

show that the DEML estimator can also be used to estimate the arrival angles of desired signals with known waveforms in the presence of interfering or jamming signals by modeling the interfering or jamming signals as random processes with an unknown spatial covariance matrix. Finally, several numerical examples showing the performance of the DEML estimator are presented in this paper.

The rest of this paper is organized as follows. In Section II, we formulate the problem of interest. In Section III, we derive the computationally efficient large sample DEML estimator. In Section IV, we present the asymptotic statistical performance of the DEML estimator and compare the performance with its CRB. In Section V, we provide several numerical examples showing the performance of the estimator. Finally, Section VI contains our conclusions.

II. PROBLEM FORMULATION

Consider the estimation of the AOA's of K narrowband plane waves impinging on an array of M sensors. The received data vector may be modeled as

$$\mathbf{x}(t) = \mathbf{A}(\boldsymbol{\theta})\mathbf{s}(t) + \mathbf{n}(t) \quad (1)$$

where $\mathbf{x}(t) \in \mathcal{C}^{M \times 1}$ is the array output vector, $\mathbf{s}(t) \in \mathcal{C}^{K \times 1}$ is the incident signal vector, and $\mathbf{n}(t) \in \mathcal{C}^{M \times 1}$ is the additive noise vector. The $M \times K$ matrix $\mathbf{A}(\boldsymbol{\theta})$ is the direction matrix corresponding to the parameter vector $\boldsymbol{\theta} = [\theta_1, \theta_2, \dots, \theta_K]^T \in \mathcal{R}^{K \times 1}$, where $(\cdot)^T$ denotes the transpose. The columns of the matrix $\mathbf{A}(\boldsymbol{\theta})$ are the direction vectors $\mathbf{a}(\theta_k)$, $k = 1, 2, \dots, K$. The waveforms of $\mathbf{s}(t)$ are assumed to be known. The amplitudes of $\mathbf{s}(t)$ are assumed to be unknown. (We have considered both cases of known and unknown amplitudes of $\mathbf{s}(t)$ in [19].)

The array is assumed to be free of rank-1 ambiguity, i.e.

$$\mathbf{a}(\theta_i) = \alpha \mathbf{a}(\theta_j) \iff \theta_i = \theta_j \quad (2)$$

where α is a nonzero scalar. Note that this unambiguous manifold assumption does not require $K < M$. We shall show herein that for signals with known waveforms, we can avoid the assumption that $K < M$, which is necessary for signals with unknown waveforms.

The noise vector $\mathbf{n}(t)$ is assumed to be a circularly symmetric complex Gaussian random vector with zero-mean and arbitrary covariance matrix \mathbf{Q} and is sampled to be temporally white, i.e.

$$\mathbb{E}[\mathbf{n}(t_i)\mathbf{n}^*(t_j)] = \mathbf{Q}\delta_{i,j} \quad (3)$$

where $(\cdot)^*$ denotes the complex conjugate transpose, and $\delta_{i,j}$ is the Kronecker delta. The unknown covariance matrix \mathbf{Q} models both thermal noise caused by the sensor output receivers and all other outside radio interference and jamming. Note that when the incident signals have unknown waveforms and are assumed to be either random processes or unknown deterministic signals, the problem of AOA estimation is ill defined if \mathbf{Q} is unknown. The AOA estimators devised for signals with unknown waveforms may assume, for example, that \mathbf{Q} is known, such as in MUSIC and ESPRIT [1], [2], or

that $\mathbf{Q} = \sigma^2 \mathbf{I}$ with σ^2 unknown, such as in MODE [20], or that \mathbf{Q} is block diagonal [21], [22].

The k th incident signal $s_k(t)$, i.e., the k th component of the signal vector $\mathbf{s}(t)$, may be written as

$$s_k(t) = \gamma_k y_k(t), \quad k = 1, 2, \dots, K \quad (4)$$

where $y_k(t)$ denotes the known signal waveform, and γ_k denotes the complex amplitude or gain of the signal and is unknown. In matrix form, the signal model may be written as

$$\mathbf{s}(t) = \boldsymbol{\Gamma} \mathbf{y}(t) \quad (5)$$

where $\boldsymbol{\Gamma} = \text{diag}[\gamma_1, \gamma_2, \dots, \gamma_K]$ and $\mathbf{y}(t) = [y_1(t), y_2(t), \dots, y_K(t)]^T$. The incident signals are assumed to be quasistationary [23], and the "covariance matrix" of the incident signals $\mathbf{s}(t)$ is defined as

$$\mathbf{R}_{ss} = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \mathbf{s}(t_n) \mathbf{s}^*(t_n). \quad (6)$$

The "covariance matrix" \mathbf{R}_{yy} of the known waveforms $\mathbf{y}(t)$ is similarly defined. For uncorrelated incident signals, both \mathbf{R}_{ss} and \mathbf{R}_{yy} will become diagonal. When the incident signals are not coherent, i.e., not completely correlated with each other, the number of signals is known since we even know their waveforms. We consider herein the case where the incident signals are not completely correlated. Finally, we assume that the signal and noise vectors are uncorrelated, i.e.

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \mathbf{s}(t_n) \mathbf{n}^*(t_n) = 0, \quad \text{with probability 1 (w.p. 1.).} \quad (7)$$

The problem of interest herein is to determine the AOA's θ_k and the unknown amplitudes γ_k (if of interest) $k = 1, 2, \dots, K$ from N independent data samples $\mathbf{x}(t_1), \mathbf{x}(t_2), \dots, \mathbf{x}(t_N)$.

III. DECOUPLED MAXIMUM LIKELIHOOD (DEML) ESTIMATOR

We consider below a large sample ML estimator for $\{\theta_k\}$ and $\{\gamma_k\}$. It is easy to show that an exact ML estimator requires a multidimensional search over the parameter space and is computationally burdensome. We shall show below that when the incident signals are uncorrelated, a large sample DEML estimator exists and is asymptotically statistically efficient. The DEML estimator determines the AOA and the amplitude (if of interest) of each incident signal separately. The DEML estimator thus reduces the K -dimensional search problem to K 1-D search problems for an arbitrary sensor array. For the special case of uniform linear arrays, we will show that we can also avoid the 1-D search problem by devising a large sample ML algorithm to compute the parameter estimates. For correlated incident signals, the DEML estimator is no longer asymptotically statistically efficient. We shall show in Section V with numerical examples that the DEML estimator may still be used to estimate the angles and amplitudes of moderately correlated incident signals without much performance degradation.

A. Arbitrary Arrays

The log-likelihood function of the array output vectors $\mathbf{x}(t_n)$, $n = 1, 2, \dots, N$ is proportional to (within an additive constant)

$$-\ln |\mathbf{Q}| - \text{tr} \left\{ \mathbf{Q}^{-1} \frac{1}{N} \sum_{n=1}^N [\mathbf{x}(t_n) - \mathbf{B}\mathbf{y}(t_n)] \cdot [\mathbf{x}(t_n) - \mathbf{B}\mathbf{y}(t_n)]^* \right\}$$

where $|\cdot|$ denotes the determinant of a matrix, and

$$\mathbf{B} = \mathbf{A}\mathbf{\Gamma} \quad (8)$$

where we have dropped the argument θ of $\mathbf{A}(\theta)$ for convenience.

Consider first the estimate of \mathbf{Q} and the unstructured estimate of \mathbf{B} . It is easy to show that

$$\hat{\mathbf{Q}} = \frac{1}{N} \sum_{n=1}^N [\mathbf{x}(t_n) - \hat{\mathbf{B}}\mathbf{y}(t_n)] [\mathbf{x}(t_n) - \hat{\mathbf{B}}\mathbf{y}(t_n)]^*, \quad (9)$$

and $\hat{\mathbf{B}}$ may be obtained by minimizing the following cost function

$$F = \left| \frac{1}{N} \sum_{n=1}^N [\mathbf{x}(t_n) - \mathbf{B}\mathbf{y}(t_n)] [\mathbf{x}(t_n) - \mathbf{B}\mathbf{y}(t_n)]^* \right|. \quad (10)$$

Let

$$\hat{\mathbf{R}}_{yx} = \frac{1}{N} \sum_{n=1}^N \mathbf{y}(t_n) \mathbf{x}^*(t_n) \quad (11)$$

and

$$\hat{\mathbf{R}}_{yy} = \frac{1}{N} \sum_{n=1}^N \mathbf{y}(t_n) \mathbf{y}^*(t_n). \quad (12)$$

Let $\hat{\mathbf{R}}_{xx}$ be defined similarly as $\hat{\mathbf{R}}_{yy}$. Then, let [24]

$$\mathbf{G} = \frac{1}{N} \sum_{n=1}^N [\mathbf{x}(t_n) - \mathbf{B}\mathbf{y}(t_n)] [\mathbf{x}(t_n) - \mathbf{B}\mathbf{y}(t_n)]^* \quad (13)$$

$$= \hat{\mathbf{R}}_{xx} - \mathbf{B}\hat{\mathbf{R}}_{yx} - \hat{\mathbf{R}}_{yx}^* \mathbf{B}^* + \mathbf{B}\hat{\mathbf{R}}_{yy} \mathbf{B}^* \quad (14)$$

$$= [\mathbf{B} - \hat{\mathbf{R}}_{yx}^* \hat{\mathbf{R}}_{yy}^{-1}] \hat{\mathbf{R}}_{yy} [\mathbf{B} - \hat{\mathbf{R}}_{yx}^* \hat{\mathbf{R}}_{yy}^{-1}]^* + \hat{\mathbf{R}}_{xx} - \hat{\mathbf{R}}_{yx}^* \hat{\mathbf{R}}_{yy}^{-1} \hat{\mathbf{R}}_{yx}. \quad (15)$$

Since the matrix $\hat{\mathbf{R}}_{yy}$ is positive definite and the second and third terms in (15) do not depend on \mathbf{B} , it follows that

$$\mathbf{G} \geq \mathbf{G}|_{\mathbf{B}=\hat{\mathbf{B}}} \quad (16)$$

where

$$\hat{\mathbf{B}} = \hat{\mathbf{R}}_{yx}^* \hat{\mathbf{R}}_{yy}^{-1}. \quad (17)$$

Since the whole sample covariance matrix \mathbf{G} is minimized, the unstructured estimate $\hat{\mathbf{B}}$ of \mathbf{B} in (17) will minimize any nondecreasing function¹ of \mathbf{G} including the determinant of \mathbf{G} ,

¹ A function $h(\mathbf{G})$ is a nondecreasing function of positive definite \mathbf{G} if for any nonnegative definite $\Delta\mathbf{G}$, $h(\mathbf{G} + \Delta\mathbf{G}) \geq h(\mathbf{G})$, and the equality holds only for $\Delta\mathbf{G} = 0$ [24].

which is F in (10). It is easy to see that $\hat{\mathbf{B}}$ is a consistent estimate of \mathbf{B} .

By using (17) with (9), the $\hat{\mathbf{Q}}$ may be rewritten as

$$\hat{\mathbf{Q}} = \hat{\mathbf{R}}_{xx} - \hat{\mathbf{R}}_{yx}^* \hat{\mathbf{R}}_{yy}^{-1} \hat{\mathbf{R}}_{yx}. \quad (18)$$

It is easy to see that $\hat{\mathbf{Q}}$ is a consistent estimate of \mathbf{Q} .

Let us now consider the structure of \mathbf{B} . The cost function in (10) may be rewritten as

$$F = |\hat{\mathbf{R}}_{xx} - \mathbf{B}\hat{\mathbf{R}}_{yx} - \hat{\mathbf{R}}_{yx}^* \mathbf{B}^* + \mathbf{B}\hat{\mathbf{R}}_{yy} \mathbf{B}^*| \quad (19)$$

$$= |\hat{\mathbf{R}}_{xx} - \hat{\mathbf{R}}_{yy} \hat{\mathbf{B}}^* + (\mathbf{B} - \hat{\mathbf{B}}) \hat{\mathbf{R}}_{yy} (\mathbf{B} - \hat{\mathbf{B}})^*| \quad (20)$$

$$= |\hat{\mathbf{Q}}| |\mathbf{I} + \hat{\mathbf{Q}}^{-1} (\mathbf{B} - \hat{\mathbf{B}}) \hat{\mathbf{R}}_{yy} (\mathbf{B} - \hat{\mathbf{B}})^*|. \quad (21)$$

The ML estimates of the angles θ_k and the gains γ_k , $k = 1, 2, \dots, K$ may be obtained by minimizing F in (21) or equivalently $\ln F$.

Theorem 1: Minimizing $\ln F$, with F defined in (21), is asymptotically equivalent to minimizing

$$F_2 = \text{tr} [\mathbf{R}_{yy} (\mathbf{B} - \hat{\mathbf{B}})^* \hat{\mathbf{Q}}^{-1} (\mathbf{B} - \hat{\mathbf{B}})]. \quad (22)$$

Proof: The proof can be found in [19]. \square

Note that minimizing F_2 in (22) also requires a multidimensional search over the parameter space. However, when the incident signals are uncorrelated with each other, which is quite common in communication systems, \mathbf{R}_{yy} in (22) is a diagonal matrix. The decoupled DEML estimator is then obtained as a large sample approximation of the exact ML estimator. For this case, the minimization in (22) is decoupled as

$$F_3 = \min_{\theta_k, \gamma_k} [\hat{\mathbf{b}}_k - \gamma_k \mathbf{a}(\theta_k)]^* \hat{\mathbf{Q}}^{-1} [\hat{\mathbf{b}}_k - \gamma_k \mathbf{a}(\theta_k)], \quad k = 1, 2, \dots, K \quad (23)$$

where $\hat{\mathbf{b}}_k$ denotes the k th column of $\hat{\mathbf{B}}$, and $\mathbf{a}(\theta_k)$ denotes the k th column of $\mathbf{A}(\theta)$.

Thus, for uncorrelated incident signals, the DEML estimates of the gains γ_k and the angles θ_k , $k = 1, 2, \dots, K$ may be easily determined as

$$\hat{\gamma}_k = \frac{\mathbf{a}^*(\hat{\theta}_k) \hat{\mathbf{Q}}^{-1} \hat{\mathbf{b}}_k}{\mathbf{a}^*(\hat{\theta}_k) \hat{\mathbf{Q}}^{-1} \mathbf{a}(\hat{\theta}_k)}, \quad (24)$$

and

$$\hat{\theta}_k = \max_{\theta_k} \frac{|\mathbf{a}^*(\theta_k) \hat{\mathbf{Q}}^{-1} \hat{\mathbf{b}}_k|^2}{\mathbf{a}^*(\theta_k) \hat{\mathbf{Q}}^{-1} \mathbf{a}(\theta_k)} \quad (25)$$

$$= \min_{\theta_k} \hat{\mathbf{b}}_k^* \hat{\mathbf{Q}}^{-1/2} \mathbf{P}_{\hat{\mathbf{a}}_k}^\perp \hat{\mathbf{Q}}^{-1/2} \hat{\mathbf{b}}_k \quad (26)$$

where

$$\hat{\mathbf{a}}_k = \hat{\mathbf{Q}}^{-1/2} \mathbf{a}_k, \quad (27)$$

where, for simplicity, we have dropped the argument of $\mathbf{a}(\theta_k)$ and $\mathbf{P}_{\hat{\mathbf{a}}_k}^\perp = \mathbf{I} - \mathbf{P}_{\hat{\mathbf{a}}_k}$ with

$$\mathbf{P}_{\hat{\mathbf{a}}_k} = \hat{\mathbf{a}}_k (\hat{\mathbf{a}}_k^* \hat{\mathbf{a}}_k)^{-1} \hat{\mathbf{a}}_k^*. \quad (28)$$

Since $\hat{\mathbf{b}}_k$ and $\hat{\mathbf{Q}}$ are consistent estimates of \mathbf{b}_k and \mathbf{Q} , respectively, when N tends to infinity, $\hat{\theta}_k$ converges to the maximum of

$$\frac{|\mathbf{a}^*(\theta_k) \mathbf{Q}^{-1} \mathbf{b}_k|^2}{\mathbf{a}^*(\theta_k) \mathbf{Q}^{-1} \mathbf{a}(\theta_k)} = |\gamma_k|^2 \frac{|\hat{\mathbf{a}}^*(\theta_k) \hat{\mathbf{a}}_k(\hat{\theta}_k)|^2}{|\hat{\mathbf{a}}(\theta_k)|^2} \quad (29)$$

where we have used $\hat{\theta}_k$ to denote the true value of the k th incident angle in order to distinguish it from the indeterminate θ_k . By assumption (2) and the Cauchy–Schwarz inequality, the absolute maximum of (25) is given by $\theta_k = \hat{\theta}_k$. This result shows that the θ_k obtained from (25) or (26) is a consistent estimate of the k th incident angle. It is then easy to see that $\hat{\gamma}_k$ is also a consistent estimate of γ_k .

The DEML estimator for signals with unknown amplitudes may be summarized as follows:

Step 1: Compute $\hat{\mathbf{b}}_k$, $k = 1, 2, \dots, K$ and $\hat{\mathbf{Q}}$ with (17) and (18), respectively.

Step 2: Determine $\hat{\gamma}_k$ (if of interest) and $\hat{\theta}_k$, $k = 1, 2, \dots, K$ with (24) and (26), respectively.

We remark that when the incident signals are uncorrelated with each other, since $\{\hat{\theta}_k, \hat{\gamma}_k\}$ are consistent and large sample realizations of the ML estimates, it follows that $\{\hat{\theta}_k, \hat{\gamma}_k\}$ are asymptotically statistically efficient, according to the general properties of ML estimators [25]. This result will be confirmed again by the performance analysis in Section IV. When the incident signals are correlated with each other, the DEML estimator is no longer a large sample ML estimator, but the DEML estimator may still be used for angle estimation. In Section V, we will confirm that for moderately correlated incident signals, the DEML estimator is no longer asymptotically statistically efficient, but the performance degradation relative to the CRB is small.

B. Uniform Linear Arrays

For a linear array of uniformly spaced identical sensors, i.e., a uniform linear array (ULA), the $\mathbf{A}(\theta)$ in (1) becomes a Vandermonde matrix, and its k th column $\mathbf{a}(\theta_k)$, $k = 1, 2, \dots, K$ becomes

$$\mathbf{a}(\theta_k) = [1, q_k, \dots, q_k^{M-1}]^T \quad (30)$$

where

$$q_k = \exp\left(j \frac{2\pi\delta}{\lambda} \sin \theta_k\right), \quad (31)$$

with δ denoting the spacing between two adjacent sensors of the ULA, λ denoting the wavelength of the incident signals, and θ_k defined relative to the array normal.

For a ULA and an arbitrary \mathbf{Q} , the minimization of the right-hand side of (26) is equivalent to solving for the roots of a polynomial of order $2(M-1)$. Compared with the exact iterative ML methods in [18], this DEML method requires the computation of a single iteration of the exact ML method.

For a ULA and an arbitrary \mathbf{Q} , the minimization of the right-hand side of (26) may be further simplified with the following noniterative method that is asymptotically equivalent to the DEML method and yet avoids solving for the roots of a high-order polynomial. We note that each function in (26) may be reparameterized in terms of the coefficients $\mathbf{c}_k = [c_{0k}, c_{1k}]^T$, $k = 1, 2, \dots, K$ of a first-order polynomial defined as

$$c_{0k}z + c_{1k} = c_{0k}(z - q_k); \quad c_{0k} \neq 0. \quad (32)$$

Let \mathbf{C}_k be an $M \times (M-1)$ matrix defined as

$$\mathbf{C}_k^* = \begin{bmatrix} c_{1k} & c_{0k} & & 0 \\ & \ddots & \ddots & \\ 0 & & c_{1k} & c_{0k} \end{bmatrix}. \quad (33)$$

Since $\text{rank}(\mathbf{C}_k) = M-1$ and $\mathbf{C}_k^* \mathbf{a}_k = \mathbf{0}$, we have $\text{rank}(\hat{\mathbf{Q}}^{1/2} \mathbf{C}_k) = M-1$ and $(\hat{\mathbf{Q}}^{1/2} \mathbf{C}_k)^* \hat{\mathbf{Q}}^{-1/2} \mathbf{a}_k = \mathbf{0}$. These two observations imply that the columns of $\hat{\mathbf{Q}}^{1/2} \mathbf{C}_k$ span the null space of $\hat{\mathbf{a}}_k^* = \mathbf{a}_k^* \hat{\mathbf{Q}}^{-1/2}$. Thus

$$\mathbf{P}_{\hat{\mathbf{a}}_k}^\perp = \hat{\mathbf{Q}}^{1/2} \mathbf{C}_k (\mathbf{C}_k^* \hat{\mathbf{Q}} \mathbf{C}_k)^{-1} \mathbf{C}_k^* \hat{\mathbf{Q}}^{1/2}. \quad (34)$$

Then, the right-hand side of (26) may be reparameterized as a function of the polynomial coefficients \mathbf{c}_k and

$$\hat{\mathbf{c}}_k = \min_{\mathbf{c}_k} [\mathbf{C}_k (\mathbf{C}_k^* \hat{\mathbf{Q}} \mathbf{C}_k)^{-1} \mathbf{C}_k^* \hat{\mathbf{b}}_k \hat{\mathbf{b}}_k^*]. \quad (35)$$

From (11) and (12), we have $\hat{\mathbf{R}}_{yx} - \mathbf{R}_{yx} = O(1/\sqrt{N})$ and $\hat{\mathbf{R}}_{yy} - \mathbf{R}_{yy} = O(1/\sqrt{N})$, where $O(1/\sqrt{N})$ denotes the (asymptotic) order, in the root-mean-square sense, of a random variable. Using first-order approximations with (17), we have $\hat{\mathbf{B}} - \mathbf{B} = O(1/\sqrt{N})$ or $\hat{\mathbf{b}}_k - \mathbf{b}_k = O(1/\sqrt{N})$ [26]. Since $\mathbf{C}_k^* \mathbf{b}_k = \mathbf{0}$ and $\mathbf{b}_k - \hat{\mathbf{b}}_k = O(1/\sqrt{N})$, we have

$$\mathbf{C}_k^* \hat{\mathbf{b}}_k = O(1/\sqrt{N}). \quad (36)$$

Then, the \mathbf{C}_k in the $\mathbf{C}_k^* \hat{\mathbf{Q}} \mathbf{C}_k$ in (35) may be replaced by a consistent estimate without affecting the asymptotic statistical efficiency of the minimizer of (35). Hence, for a uniform linear array, we can first determine an initial estimate of \mathbf{c}_k , $k = 1, 2, \dots, K$, by minimizing $\text{tr}[\mathbf{C}_k^* \hat{\mathbf{b}}_k \hat{\mathbf{b}}_k^* \mathbf{C}_k]$. This estimate of \mathbf{c}_k is consistent since $\hat{\mathbf{b}}_k$ is a consistent estimate of \mathbf{b}_k . We then determine an efficient estimate of \mathbf{c}_k using (35) with \mathbf{C}_k in $(\mathbf{C}_k^* \hat{\mathbf{Q}} \mathbf{C}_k)^{-1}$ replaced by its consistent estimate formed from the initial estimate of \mathbf{c}_k . This algorithm avoids solving for the roots of a high-order polynomial and is computationally efficient for determining the asymptotically statistically efficient estimates of the angles θ_k . In order to obtain a meaningful solution to these quadratic minimizations, the conjugate symmetry condition $\hat{c}_{0k} = \hat{c}_{1k}^*$ [27], [28] and $\text{Re}^2\{c_{0k}\} + \text{Im}^2\{c_{0k}\} = 1$ can be imposed on $\hat{\mathbf{c}}_k$. These constraints are considered in detail in [19]. In the numerical examples given in Section V, we present results obtained by imposing these constraints on $\hat{\mathbf{c}}_k$.

C. Properties of the DEML Estimator

We now sum up five significant advantages of the DEML estimator for uncorrelated signals with known waveforms as compared with the case of signals with unknown waveforms.

First of all, the large sample and asymptotically statistically efficient DEML estimator is computationally much simpler than any existing large sample ML estimators for unknown signals. It has been shown in [9] and [10] that for the case of uncorrelated and unknown signals, the multidimensional ML angle estimation problem also asymptotically decouples into K 1-D problems, which may be solved with the standard MUSIC algorithm given in [1]. The MUSIC algorithm, however, requires the eigendecomposition of the array covariance matrix, which is computationally expensive. On the contrary,

the cost function associated with the DEML estimator does not require any eigendecomposition. Moreover, on a parallel computer, the DEML estimator can be naturally implemented in a parallel fashion, i.e., by minimizing the K functionals in (26) in parallel.

Second, the accuracy provided by the DEML estimator for uncorrelated signals with known waveforms is superior to that of the best estimators for unknown signals (see the comparisons given in [18] and the numerical examples in Section V of this paper). In fact, when unknown signals are modeled as unknown deterministic parameters and the number of array sensors is finite, no estimator can achieve its CRB [9], which is bound to be greater than or equal to the CRB for signals with known waveforms due to the parsimony principle [24].

Third, the DEML estimator does not suffer any accuracy degradation when the smallest angle separation $\Delta = \inf_{i \neq j} |\theta_i - \theta_j|$ tends to zero. By contrast, the accuracy of any estimator that does not exploit the knowledge of the waveforms significantly degrades (until complete failure) when Δ approaches zero. This result is also illustrated in Section V of this paper with a numerical example.

Fourth, the DEML estimator has no constraints on the number of incident signals at all, provided that the number of data samples is large enough, while estimators for unknown signals require that the number of signals be less than the number of array sensors.

Fifth, the DEML estimator can handle the case of *unknown* spatially colored noise with little additional difficulties. The estimators for unknown signals, however, fail to handle this case. This advantage of the DEML estimator is particularly useful for estimating the incident angles of signals with known waveforms in the presence of unknown interfering and jamming signals that are not completely correlated with any of the known waveforms. This is especially true when the number of interfering and jamming signals is large (for example, larger than the number of array sensors) and when some of the interfering and jamming signals are wideband. The unknown noise covariance matrix \mathbf{Q} may be used to accommodate both the presence of these interfering and jamming signals and any other noise, including the thermal noise.

IV. ESTIMATOR PERFORMANCE AND CRAMÉR-RAO BOUND

In this section, we present the asymptotic statistical performance analysis of the DEML estimator and compare the result to its CRB, i.e., the best possible performance for the class of asymptotically unbiased estimators. We shall confirm that if the incident signals are uncorrelated, the DEML estimator is asymptotically statistically efficient, i.e., the error covariance matrix of the estimates approaches the corresponding CRB asymptotically. We shall show in Section V with numerical examples that when the incident signals are moderately correlated, the relative efficiency of the DEML estimator, i.e., the square root of the ratio between the corresponding CRB and the error variance of an estimate, is close to 1.

We first present the large sample statistical performance of the DEML angle estimates and their CRB's since the angle estimates are of the most interest. We then present the results when both the angle and amplitude estimates are of interest.

Theorem 2: Let $\{\hat{\theta}_k\}$ be estimated with the DEML estimator, i.e., with (17), (18), and (26). Then, the asymptotic covariance matrix of $\hat{\theta}$ is given by

$$E[(\hat{\theta} - \theta)(\hat{\theta} - \theta)^T] = \frac{1}{2N} \text{Re}[(\Psi^* \Psi) \odot \mathbf{R}_{ss}^{-T}] \quad (37)$$

where \odot denotes the elementwise multiplication between two matrices, \mathbf{R}_{ss} is defined in (6), and

$$\Psi = \left[\frac{\mathbf{P}_{\check{\mathbf{a}}_1}^\perp \check{\mathbf{d}}_1}{\|\mathbf{P}_{\check{\mathbf{a}}_1}^\perp \check{\mathbf{d}}_1\|^2}, \frac{\mathbf{P}_{\check{\mathbf{a}}_2}^\perp \check{\mathbf{d}}_2}{\|\mathbf{P}_{\check{\mathbf{a}}_2}^\perp \check{\mathbf{d}}_2\|^2}, \dots, \frac{\mathbf{P}_{\check{\mathbf{a}}_K}^\perp \check{\mathbf{d}}_K}{\|\mathbf{P}_{\check{\mathbf{a}}_K}^\perp \check{\mathbf{d}}_K\|^2} \right] \quad (38)$$

with

$$\check{\mathbf{a}}_k = \mathbf{Q}^{-1/2} \mathbf{a}_k, \quad (39)$$

$$\check{\mathbf{d}}_k = \mathbf{Q}^{-1/2} \mathbf{d}_k \quad (40)$$

and

$$\mathbf{d}_k = \frac{\partial \mathbf{a}(\theta_k)}{\partial \theta_k}. \quad (41)$$

Proof: See the Appendix. \square

Theorem 3: The CRB of the angle estimates for signals with known waveforms, unknown amplitudes, and unknown noise covariance matrix may be written as

$$\text{CRB}(\theta) = \frac{1}{2N} \text{Re}^{-1}[(\check{\mathbf{D}}^* \check{\mathbf{D}}) \odot \hat{\mathbf{R}}_{ss}^T - \Delta^* \Lambda^{-1} \Delta] \quad (42)$$

where

$$\check{\mathbf{D}} = [\check{\mathbf{d}}_1, \check{\mathbf{d}}_2, \dots, \check{\mathbf{d}}_K] \quad (43)$$

with $\check{\mathbf{d}}_k$, $k = 1, 2, \dots, K$ as defined in (40), and

$$\Delta = (\check{\mathbf{A}}^* \check{\mathbf{D}}) \odot \hat{\mathbf{R}}_{ss}^T, \quad (44)$$

and

$$\Lambda = (\check{\mathbf{A}}^* \check{\mathbf{A}}) \odot \hat{\mathbf{R}}_{ss}^T \quad (45)$$

with $\check{\mathbf{A}} = \mathbf{Q}^{-1/2} \mathbf{A}$.

Proof: The proof is a straightforward extension of the corresponding one in [18] and is therefore omitted. The details can be found in [19]. \square

For large N and uncorrelated signals, $\hat{\mathbf{R}}_{ss}$ approaches \mathbf{R}_{ss} , which is a diagonal matrix. Thus, the CRB in (42) also becomes diagonal, and the CRB for the k th-angle estimate may be written as

$$\begin{aligned} \text{CRB}(\theta_k) &= \frac{\check{\mathbf{a}}_k^* \check{\mathbf{a}}_k}{2N[\mathbf{R}_{ss}]_{kk} \check{\mathbf{d}}_k^* (\check{\mathbf{a}}_k^* \check{\mathbf{a}}_k \mathbf{I} - \check{\mathbf{a}}_k \check{\mathbf{a}}_k^*) \check{\mathbf{d}}_k} \\ &= \frac{1}{2N[\mathbf{R}_{ss}]_{kk} \check{\mathbf{d}}_k^* \mathbf{P}_{\check{\mathbf{a}}_k}^\perp \check{\mathbf{d}}_k} \\ &= \frac{1}{2N[\mathbf{R}_{ss}]_{kk} \|\mathbf{P}_{\check{\mathbf{a}}_k}^\perp \check{\mathbf{d}}_k\|^2} \end{aligned} \quad (46)$$

where $[\mathbf{R}_{ss}]_{kk}$ is the power of the k th incident signal. As expected, this result is the same as (37) when \mathbf{R}_{ss} is a diagonal matrix, which confirms that the DEML estimator is asymptotically statistically efficient for uncorrelated signals with known waveforms and unknown amplitudes.

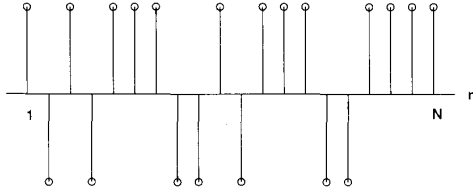


Fig. 1. Typical known waveform.

For a ULA and $\mathbf{Q} = \sigma^2 \mathbf{I}$, (46) may be further simplified to

$$\text{CRB}(\theta_k) = \frac{6\sigma^2}{N(M^2 - 1)M(2\pi\delta/\lambda)^2 [\mathbf{R}_{ss}]_{kk} \cos^2(\theta_k)} \quad (47)$$

which checks with the corresponding CRB expression in [18].

As already shown, the DEML estimator may be used to estimate not only the incident angles but the unknown amplitudes γ_k as well. The unknown amplitudes may be of interest in certain applications such as in communications and bistatic radar (where the transmitter and the receiver of the radar are at different locations). In the communications applications, the γ_k represents the channel gain and phase shift. In radar applications, γ_k represents the radar cross section of a target. Let

$$\mathbf{u} = [\mathbf{u}_1^T, \mathbf{u}_2^T, \dots, \mathbf{u}_K^T]^T \quad (48)$$

where

$$\mathbf{u}_k = [\theta_k, \text{Re}(\gamma_k), \text{Im}(\gamma_k)]^T, \quad k = 1, 2, \dots, K. \quad (49)$$

Then, the asymptotic statistic of $\hat{\mathbf{u}}$ is given by the following theorem.

Theorem 4: Let $\{\hat{\mathbf{u}}_k\}$ be estimated with the DEML estimator, i.e., with (17), (18), (26), and (24). Then, the asymptotic covariance matrix of $\hat{\mathbf{u}}$ is given by

$$E[(\hat{\mathbf{u}} - \mathbf{u})(\hat{\mathbf{u}} - \mathbf{u})^T] = \frac{1}{2N} \text{Re}^{-1}(\Xi) \text{Re}[(\mathbf{V}^* \mathbf{Q}^{-1} \mathbf{V}) \odot (\mathbf{R}_{yy}^{-T} \otimes \mathbf{E}_3)] \text{Re}^{-1}(\Xi) \quad (50)$$

where \mathbf{E}_3 denotes the 3×3 matrix with all elements equal to one, and \otimes denotes the Kronecker product

$$\mathbf{V} = [\mathbf{V}_1, \mathbf{V}_2, \dots, \mathbf{V}_K], \quad (51)$$

and

$$\Xi = \begin{bmatrix} \mathbf{V}_1^* \mathbf{Q}^{-1} \mathbf{V}_1 & & 0 \\ & \ddots & \\ 0 & & \mathbf{V}_K^* \mathbf{Q}^{-1} \mathbf{V}_K \end{bmatrix} \quad (52)$$

with

$$\mathbf{V}_k = \frac{\partial \mathbf{b}_k}{\partial \mathbf{u}_k}. \quad (53)$$

Proof: The proof can be found in [19]. \square

We can show [19] that the DEML estimator provides asymptotically statistically efficient estimates for both incident angles and unknown amplitudes when the known waveforms are uncorrelated with each other.

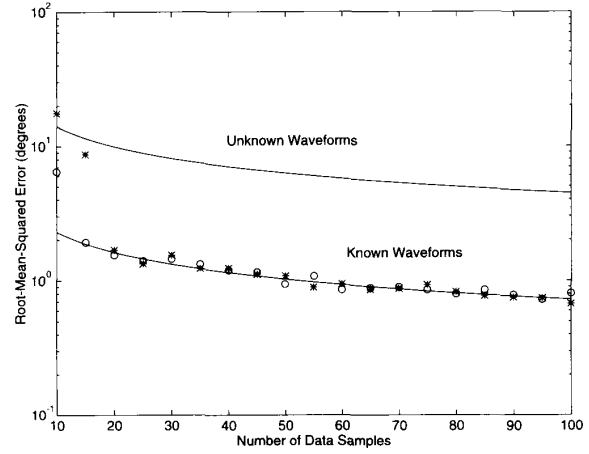


Fig. 2. Root-mean-squared error (RMSE) of the DEML and the exact ML estimators as a function of N when two uncorrelated signals arrive from $\theta_1 = 0^\circ$ and $\theta_2 = 5^\circ$, $M = 5$ and $\text{SNR} = -5$ dB. The solid lines are for the asymptotic CRB's. The statistical performance of DEML estimator coincides with its asymptotic CRB for known waveforms. The symbols "o" and "*" are for the exact ML and DEML estimators, respectively.

V. NUMERICAL RESULTS

We present in this section several examples showing the performance of the DEML estimator described in Section III. The array used in the examples consists of M identical sensors that are uniformly spaced with a spacing δ between adjacent sensors of a half wavelength. The incident signals are binary phase-shift keyed (BPSK) signals, and the gains are assumed one. A typical known waveform used in our examples is shown in Fig. 1 for $N = 20$. The performance of the estimator in each of the examples below was obtained from 100 Monte Carlo simulations. The DEML angle estimates are calculated with the computationally efficient algorithm presented in Section III-B. In Sections V-A and B, the additive noise is assumed to be the thermal noise only, which is spatially white, i.e., $\mathbf{Q} = \sigma^2 \mathbf{I}$. The examples showing the cases where the additive noise consists of both the spatially white thermal noise and unknown narrowband interfering signals can be found in [19]. The empirical results of the DEML estimator are compared with those of the exact ML estimator. The results are also compared with the asymptotic statistical performance of the DEML estimator derived in the previous section and its asymptotic CRB's.

A. Uncorrelated Signals

Consider first an example where two uncorrelated signals with known waveforms arrive at a uniform linear array of $M = 5$ sensors from $\theta_1 = 0^\circ$ and $\theta_2 = 5^\circ$. The signal-to-noise ratio (SNR) of each signal at each sensor output is assumed to be -5 dB. Fig. 2 shows the performance of the DEML estimator and the exact ML estimator as a function of the number of data samples N . This figure and the figures below show the root-mean-squared error (RMSE) for the first signal. (The results are similar for the other signal or signals.) The exact ML estimates are obtained by the iterative alternating maximization approach described in [18]. Arbitrary initial conditions are assumed for the exact ML estimator. Note that

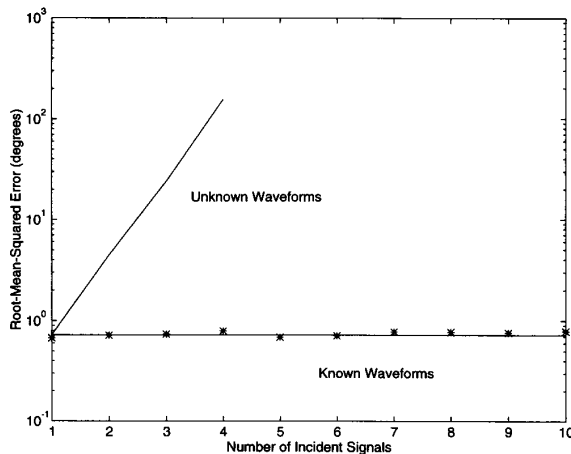


Fig. 3. Root-mean-squared error (RMSE) of the DEML and the exact ML estimators as a function of K when K uncorrelated signals arrive from $0^\circ, 5^\circ, \dots, [5(K-1)]^\circ$, $N = 100$ and $\text{SNR} = -5$ dB. The solid lines are for the asymptotic CRB's. The statistical performance of DEML estimator coincides with its asymptotic CRB for known waveforms. The symbols "+" are for the DEML estimator.

the number of iterations needed by the exact ML estimator may be reduced by using initial conditions obtained with estimators such as MODE/WSF [20], [29], [30] or IQML [27], which were developed for unknown signals. When MODE is used, the total computation needed to obtain the exact ML estimates may be reduced by 4% or less for this example. The accuracies of the exact ML estimates obtained with and without using MODE are similar. Fig. 2 also shows the asymptotic statistical performance of the DEML estimator, which coincides with the asymptotic CRB's. For comparison, the CRB's for signals with unknown waveforms (see [18]) are shown as well. Note that although the DEML estimator is a large sample ML estimator, its performance is very close to its CRB even for N as small as $N = 20$ for this example. Note that as the number of data samples becomes even smaller, the performance of both the DEML estimator and the exact ML estimator degrades as compared with the CRB. The performance of the exact ML estimator is better than that of the DEML estimator for very small N (i.e., $N < 20$ for this example). The DEML estimator, however, requires much less computation than the exact ML estimator. The computation, which is measured by the count of floating-point operations required by the DEML estimator is approximately 54 (for small N) and, respectively, 35 (for large N) times less than that required by the exact ML estimator.

Our numerical examples also show that whether a given N is large or not depends on what the M and the SNR are. We have found that the larger the M , the more data samples are needed for the DEML estimator to achieve its CRB. We have also found that the higher the SNR, the closer the DEML estimator to its asymptotic CRB even when the number of data samples is very small. More detailed discussions can be found in [19].

Fig. 3 shows the performance of the DEML estimator and its CRB as a function of the number of incident signals K when the K uncorrelated signals arrive from $0^\circ, 5^\circ, \dots, [5(K-1)]^\circ$, $M = 5$, $N = 100$, and $\text{SNR} = -5$ dB. We note that the performance of the DEML estimator and its CRB for uncorrelated signals with known waveforms are independent of K , even when K is larger than the number of array sensors M . For signals with unknown waveforms, however, K must be less than M , and the CRB increases as K increases.

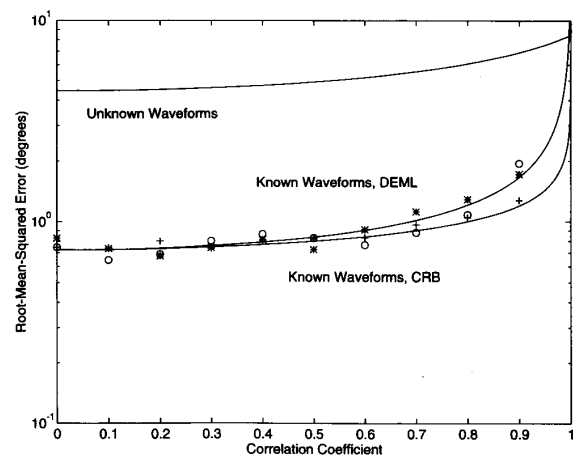


Fig. 4. Root-mean-squared error (RMSE) of the DEML and the exact ML estimators as a function of correlation coefficient when two correlated signals arrive from $\theta_1 = 0^\circ$ and $\theta_2 = 5^\circ$, $M = 5$, $N = 100$ and $\text{SNR} = -5$ dB. The solid lines are for the asymptotic CRB's and the asymptotic statistical performance of DEML estimator. The symbols "+" and "o", and "*" are for the exact ML estimator (with and without MODE), and the DEML estimator, respectively.

We remark that although we considered only the case of 5° angle separation between incident signals, the performance of the DEML and exact ML estimators for signals with known waveforms and unknown amplitudes is similar for all angle separations. In fact, the CRB's for uncorrelated signals with known waveforms are independent of angle separations as shown in [18] and in Sections III and IV of this paper.

B. Correlated Signals

In this section, we shall show how the performance of the DEML estimator degrades relative to its CRB as the incident signals become correlated. Consider an example where two correlated signals with known waveforms arrive at an array of $M = 5$ sensors from $\theta_1 = 0^\circ$ and $\theta_2 = 5^\circ$. The SNR of each signal at each sensor output is assumed to be -5 dB, and the number of data samples $N = 100$. Fig. 4 shows the performance of the DEML estimator, the exact ML estimator with arbitrary initial conditions, and the exact ML estimator with initial conditions calculated with MODE as a function of the correlation coefficient between the two incident signals. Note that when the incident signals are highly correlated, the exact ML estimator may encounter a severe local minima problem. When arbitrary initial conditions are used for the exact ML estimator, the performance of the exact ML estimator may be poorer than that of the DEML estimator. The performance of the exact ML estimator may be improved by first using MODE to calculate the initial conditions. When MODE is used, the total computation needed to obtain the exact ML estimates may be reduced by 15%. In this example, the computation required

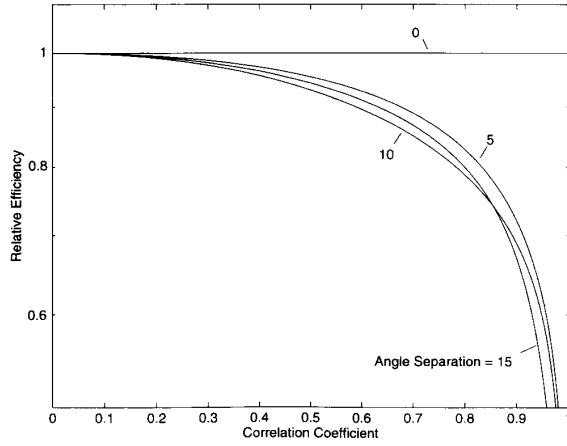


Fig. 5. Relative efficiency of the DEML estimator as a function of the correlation coefficient for various Δ when two correlated signals arrive from $\theta_1 = 0^\circ$ and $\theta_2 = \Delta$, $M = 5$, SNR = -5 dB, and $N = 100$.

by the DEML estimator is approximately 35 (for uncorrelated signals) and, respectively, 92 (for highly correlated signals) times less than that required by the exact ML estimator with arbitrary initial conditions.

Fig. 4 also shows the asymptotic statistical performance of the DEML estimator and the asymptotic CRB's. For comparison, the CRB's for signals with unknown waveforms (see [18]) are shown as well. Note that the Monte-Carlo simulation results are close to the corresponding theoretical performance analysis results for both the DEML and the exact ML estimators. We also note that although the DEML estimator is an ML estimator for uncorrelated signals only, the performance degradation for moderately correlated signals is small, as compared with its CRB's.

We consider next the relative efficiency of the DEML estimator as a function of the correlation coefficient between incident signals for various parameters. The relative efficiency of the DEML estimator is defined as the square root of the ratio between the corresponding CRB and the MSE. Consider first an example where two correlated signals with known waveforms arrive at an array of $M = 5$ sensors from $\theta_1 = 0^\circ$ and $\theta_2 = \Delta$. We also assume that SNR = -5 dB and $N = 100$. Fig. 5 shows the relative efficiency of the DEML estimator, which is obtained with the asymptotic statistical analysis results, as a function of the correlation coefficient for various Δ . We note that when $\Delta = 0^\circ$, the DEML estimator is efficient for both correlated and uncorrelated signals. For nonzero Δ , the relative efficiency decreases as the correlation coefficient increases. Yet, the relative efficiency for all Δ remains close to 1 for moderately correlated signals.

Our numerical examples also show that for correlated signals, the relative efficiency of the DEML estimator decreases as M increases. The decrease, however, is slow for moderately correlated signals. We have also found that for moderately correlated signals, the relative efficiency of the DEML estimator remains close to 1 when K is increased. This result holds even when $K > M$. More detailed discussions can be found in [19].

Finally, we remark that the DEML estimator can also be used to estimate the incident angles of desired signals

with known waveforms in the presence of interfering or jamming signals that are uncorrelated with the desired signals [18]. In particular, the interfering or jamming signals can be modeled as random processes with an unknown arbitrary spatial covariance matrix. This covariance matrix and that of the additive noise together are simply modeled with an unknown spatial covariance matrix \mathbf{Q} . This modeling approach does not add any extra difficulties to the DEML estimator since the DEML estimator is derived based on an unknown noise covariance matrix \mathbf{Q} . Moreover, when this model is used, the number of interfering or jamming signals may be greater than or equal to the number of array sensors, and the interfering or jamming signals may be wideband. Detailed examples showing the performance of this modeling approach can be found in [19].

VI. CONCLUSIONS

We have presented a large sample decoupled maximum likelihood (DEML) angle estimator for narrowband plane waves with known waveforms and unknown amplitudes arriving at a sensor array in the presence of unknown and arbitrary spatially colored noise. The DEML estimator decouples the multidimensional problem of the exact ML estimator to a set of 1-D problems and hence is computationally efficient. We have derived the asymptotic statistical performance of the DEML estimator and compared its performance with its Cramér-Rao bound (CRB), i.e., the best possible performance for the class of asymptotically unbiased estimators. We have shown that the DEML estimator is asymptotically statistically efficient for uncorrelated signals with known waveforms. We have shown that for moderately correlated signals with known waveforms, the DEML estimator is no longer asymptotically statistically efficient, but the performance degradation relative to the CRB is small. We have shown that the DEML estimator can also be used to estimate the arrival angles of desired signals with known waveforms in the presence of interfering or jamming signals by modeling the interfering or jamming signals as random processes with an unknown spatial covariance matrix. Finally, several numerical examples showing the performance of the DEML estimator have been presented in this paper.

APPENDIX PROOF OF THEOREM 2

We first derive an expression for the covariance matrix of $\text{vec}(\hat{\mathbf{B}} - \mathbf{B})$, where $\text{vec}(\cdot)$ denotes stacking all columns of a matrix into a single column vector. For sufficiently large N , we have

$$\begin{aligned} \text{vec}(\hat{\mathbf{B}} - \mathbf{B}) &= \text{vec}(\hat{\mathbf{R}}_{yx}^* \hat{\mathbf{R}}_{yy}^{-1} - \mathbf{B}) \\ &= \text{vec} \left\{ \frac{1}{N} \sum_{n=1}^N [\mathbf{x}(t_n) - \mathbf{B}\mathbf{y}(t_n)] \mathbf{y}^*(t_n) \hat{\mathbf{R}}_{yy}^{-1} \right\} \\ &= \text{vec} \left[\frac{1}{N} \sum_{n=1}^N \mathbf{n}(t_n) \mathbf{y}^*(t_n) \hat{\mathbf{R}}_{yy}^{-1} \right] \end{aligned} \quad (54)$$

$$\simeq (\mathbf{R}_{yy}^{-T} \otimes \mathbf{I}) \frac{1}{N} \sum_{n=1}^N [\mathbf{y}^{*T}(t_n) \otimes \mathbf{I}] \mathbf{n}(t_n) \quad (55)$$

$$\begin{aligned}
& E[(\hat{\theta}_k - \theta_k)(\hat{\theta}_j - \theta_j)] \\
&= \frac{E\{f'_1(\theta_k)f'_1(\theta_j)\}}{f''_1(\theta_k)f''_1(\theta_j)} \\
&= \text{Re} \left\{ \frac{2\gamma_k \gamma_j^* \tilde{\mathbf{d}}_k^* \mathbf{P}_{\tilde{\mathbf{a}}_k}^\perp \mathbf{Q}^{-1/2} E[(\hat{\mathbf{b}}_k - \mathbf{b}_k)(\hat{\mathbf{b}}_j - \mathbf{b}_j)^*] \mathbf{Q}^{-1/2} \mathbf{P}_{\tilde{\mathbf{a}}_j}^\perp \tilde{\mathbf{d}}_j}{4|\gamma_k|^2 |\gamma_j|^2 (\tilde{\mathbf{d}}_k^* \mathbf{P}_{\tilde{\mathbf{a}}_k}^\perp \tilde{\mathbf{d}}_k)(\tilde{\mathbf{d}}_j^* \mathbf{P}_{\tilde{\mathbf{a}}_j}^\perp \tilde{\mathbf{d}}_j)} \right\} = \text{Re} \left\{ \frac{[\mathbf{R}_{ss}^{-T}]_{kj} \tilde{\mathbf{d}}_k^* \mathbf{P}_{\tilde{\mathbf{a}}_k}^\perp \mathbf{P}_{\tilde{\mathbf{a}}_j}^\perp \tilde{\mathbf{d}}_j}{2N (\tilde{\mathbf{d}}_k^* \mathbf{P}_{\tilde{\mathbf{a}}_k}^\perp \tilde{\mathbf{d}}_k)(\tilde{\mathbf{d}}_j^* \mathbf{P}_{\tilde{\mathbf{a}}_j}^\perp \tilde{\mathbf{d}}_j)} \right\} \\
&= \text{Re} \left\{ \frac{[\mathbf{R}_{ss}^{-T}]_{kj}}{2N} \left(\frac{\mathbf{P}_{\tilde{\mathbf{a}}_k}^\perp \tilde{\mathbf{d}}_k}{\|\mathbf{P}_{\tilde{\mathbf{a}}_k}^\perp \tilde{\mathbf{d}}_k\|^2} \right)^* \left(\frac{\mathbf{P}_{\tilde{\mathbf{a}}_j}^\perp \tilde{\mathbf{d}}_j}{\|\mathbf{P}_{\tilde{\mathbf{a}}_j}^\perp \tilde{\mathbf{d}}_j\|^2} \right) \right\} \quad (67)
\end{aligned}$$

where we have used [31]

$$\text{vec}(\mathbf{ABC}) = (\mathbf{C}^T \otimes \mathbf{A}) \text{vec}(\mathbf{B}). \quad (56)$$

Since the noise vector $\mathbf{n}(t_n)$ is a circularly symmetric Gaussian random vector, we have

$$\lim_{N \rightarrow \infty} NE[\text{vec}(\hat{\mathbf{B}} - \mathbf{B}) \text{vec}^T(\hat{\mathbf{B}} - \mathbf{B})] = 0, \quad (57)$$

and

$$\begin{aligned}
& \lim_{N \rightarrow \infty} NE[\text{vec}(\hat{\mathbf{B}} - \mathbf{B}) \text{vec}^*(\hat{\mathbf{B}} - \mathbf{B})] \\
&= (\mathbf{R}_{yy}^{-T} \otimes \mathbf{I}) \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^N [\mathbf{y}^{*T}(t_i) \otimes \mathbf{I}] E[\mathbf{n}(t_i) \mathbf{n}^*(t_j)] \\
&\quad \cdot [\mathbf{y}^T(t_j) \otimes \mathbf{I}] (\mathbf{R}_{yy}^{-T} \otimes \mathbf{I}) \\
&= \mathbf{R}_{yy}^{-T} \otimes \mathbf{Q}. \quad (58)
\end{aligned}$$

Next, note that the DEML estimate of θ_k is obtained by maximizing the following cost function:

$$f_1(\theta_k) = \hat{\mathbf{b}}_k^* \hat{\mathbf{Q}}^{-1/2} \mathbf{P}_{\tilde{\mathbf{a}}_k}^\perp \hat{\mathbf{Q}}^{-1/2} \hat{\mathbf{b}}_k. \quad (59)$$

Then, we have

$$\hat{\theta}_k - \theta_k = -\frac{f'_1(\theta_k)}{f''_1(\theta_k)} \quad (60)$$

where $f'_1(\theta_k) = \partial f_1(\theta_k) / \partial \theta_k$.

The $f'_1(\theta_k)$ may be written as [29]

$$f'_1(\theta_k) = 2\text{Re} \{ \hat{\mathbf{b}}_k^* \hat{\mathbf{Q}}^{-1/2} \mathbf{P}_{\tilde{\mathbf{a}}_k}^\perp \tilde{\mathbf{d}}_k \tilde{\mathbf{a}}_k^\dagger \hat{\mathbf{Q}}^{-1/2} \hat{\mathbf{b}}_k \} \quad (61)$$

where

$$\tilde{\mathbf{d}}_k = \hat{\mathbf{Q}}^{-1/2} \mathbf{d}_k, \quad (62)$$

and

$$\tilde{\mathbf{a}}_k^\dagger = (\tilde{\mathbf{a}}_k \tilde{\mathbf{a}}_k)^{-1} \tilde{\mathbf{a}}_k^*. \quad (63)$$

Since $\mathbf{b}_k^* \hat{\mathbf{Q}}^{-1/2} \mathbf{P}_{\tilde{\mathbf{a}}_k}^\perp = \gamma_k^* \tilde{\mathbf{a}}_k^* \mathbf{P}_{\tilde{\mathbf{a}}_k}^\perp = 0$, we have

$$\begin{aligned}
f'_1(\theta_k) &= 2\text{Re} \{ (\hat{\mathbf{b}}_k^* - \mathbf{b}_k^*) \hat{\mathbf{Q}}^{-1/2} \mathbf{P}_{\tilde{\mathbf{a}}_k}^\perp \tilde{\mathbf{d}}_k \tilde{\mathbf{a}}_k^\dagger \hat{\mathbf{Q}}^{-1/2} \hat{\mathbf{b}}_k \} \\
&\simeq 2\text{Re} \{ (\hat{\mathbf{b}}_k^* - \mathbf{b}_k^*) \hat{\mathbf{Q}}^{-1/2} \mathbf{P}_{\tilde{\mathbf{a}}_k}^\perp \tilde{\mathbf{d}}_k \tilde{\mathbf{a}}_k^\dagger \hat{\mathbf{Q}}^{-1/2} \mathbf{b}_k \} \\
&= 2\text{Re} \{ \gamma_k (\hat{\mathbf{b}}_k - \mathbf{b}_k)^* \hat{\mathbf{Q}}^{-1/2} \mathbf{P}_{\tilde{\mathbf{a}}_k}^\perp \tilde{\mathbf{d}}_k \} \\
&\simeq 2\text{Re} \{ \gamma_k (\hat{\mathbf{b}}_k - \mathbf{b}_k)^* \hat{\mathbf{Q}}^{-1/2} \mathbf{P}_{\tilde{\mathbf{a}}_k}^\perp \tilde{\mathbf{d}}_k \} \quad (64)
\end{aligned}$$

where we have used $\tilde{\mathbf{a}}_k^\dagger \hat{\mathbf{Q}}^{-1/2} \mathbf{b}_k = \gamma_k \tilde{\mathbf{a}}_k^\dagger \tilde{\mathbf{a}}_k = \gamma_k$ and the " \simeq " is asymptotically equivalent to " $=$ " since $\hat{\mathbf{b}}_k^* - \mathbf{b}_k^* = O(1/\sqrt{N})$, and $\hat{\mathbf{Q}}$ is a consistent estimate of \mathbf{Q} .

The $f''_1(\theta_k)$ may be obtained by differentiating the right-hand side of (61) and may be written as

$$f''_1(\theta_k) = -2\text{Re} \{ \hat{\mathbf{b}}_k^* \hat{\mathbf{Q}}^{-1/2} [\mathbf{P}_{\tilde{\mathbf{a}}_k}^\perp \tilde{\mathbf{d}}_k \tilde{\mathbf{a}}_k^\dagger \tilde{\mathbf{d}}_k \tilde{\mathbf{a}}_k^\dagger + \tilde{\mathbf{a}}_k^* \tilde{\mathbf{d}}_k^* \mathbf{P}_{\tilde{\mathbf{a}}_k}^\perp \tilde{\mathbf{d}}_k \tilde{\mathbf{a}}_k^\dagger - \mathbf{P}_{\tilde{\mathbf{a}}_k}^\perp \tilde{\mathbf{d}}_k (\tilde{\mathbf{a}}_k^\dagger)' - \mathbf{P}_{\tilde{\mathbf{a}}_k}^\perp \tilde{\mathbf{d}}_k' \tilde{\mathbf{a}}_k^\dagger] \hat{\mathbf{Q}}^{-1/2} \hat{\mathbf{b}}_k \} \quad (65)$$

where $(\cdot)' = \partial / \partial \theta_k$. Since $f'_1(\theta_k) = O(1/\sqrt{N})$, the error $\hat{\theta}_k - \theta_k$ in (60) will remain asymptotically unchanged when we rewrite $f''_1(\theta_k)$ as

$$\begin{aligned}
f''_1(\theta_k) &\simeq -2\text{Re} \{ \hat{\mathbf{b}}_k^* \hat{\mathbf{Q}}^{-1/2} [\mathbf{P}_{\tilde{\mathbf{a}}_k}^\perp \tilde{\mathbf{d}}_k \tilde{\mathbf{a}}_k^\dagger \tilde{\mathbf{d}}_k \tilde{\mathbf{a}}_k^\dagger + \tilde{\mathbf{a}}_k^* \tilde{\mathbf{d}}_k^* \mathbf{P}_{\tilde{\mathbf{a}}_k}^\perp \tilde{\mathbf{d}}_k \tilde{\mathbf{a}}_k^\dagger - \mathbf{P}_{\tilde{\mathbf{a}}_k}^\perp \tilde{\mathbf{d}}_k (\tilde{\mathbf{a}}_k^\dagger)' - \mathbf{P}_{\tilde{\mathbf{a}}_k}^\perp \tilde{\mathbf{d}}_k' \tilde{\mathbf{a}}_k^\dagger] \hat{\mathbf{Q}}^{-1/2} \hat{\mathbf{b}}_k \} \\
&= -2\text{Re} \{ \hat{\mathbf{b}}_k^* \hat{\mathbf{Q}}^{-1/2} \tilde{\mathbf{a}}_k^* \tilde{\mathbf{d}}_k^* \mathbf{P}_{\tilde{\mathbf{a}}_k}^\perp \tilde{\mathbf{d}}_k \tilde{\mathbf{a}}_k^\dagger \hat{\mathbf{Q}}^{-1/2} \hat{\mathbf{b}}_k \} \\
&= -2|\gamma_k|^2 \tilde{\mathbf{d}}_k^* \mathbf{P}_{\tilde{\mathbf{a}}_k}^\perp \tilde{\mathbf{d}}_k \\
&\simeq -2|\gamma_k|^2 \tilde{\mathbf{d}}_k^* \mathbf{P}_{\tilde{\mathbf{a}}_k}^\perp \tilde{\mathbf{d}}_k. \quad (66)
\end{aligned}$$

Using (64), (66), (57), and (58) with (60), we obtain (67), which appears at the top of the page, where we have used $\mathbf{P}_{\tilde{\mathbf{a}}_k}^\perp \mathbf{P}_{\tilde{\mathbf{a}}_k}^\perp = \mathbf{P}_{\tilde{\mathbf{a}}_k}^\perp$ and $\text{Re}(z_1) \text{Re}(z_2) = \text{Re}(z_1 z_2^* + z_1^* z_2)/2$.

Writing (67) in matrix form, we prove the theorem. \square

REFERENCES

- [1] R. O. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propagat.*, vol. AP-34, no. 3, pp. 276-280, Mar. 1986.
- [2] R. Roy and T. Kailath, "ESPRIT—Estimation of signal parameters via rotational invariance techniques," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, no. 7, pp. 984-995, July 1989.
- [3] 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.
- [4] J. F. Böhme, "Estimation of spectral parameters of correlated signals in wavefields," *Signal Processing*, vol. 10, pp. 329-337, 1986.
- [5] B. Ottersten, M. Viberg, P. Stoica, and A. Nehorai, "Exact and large sample ML techniques for parameter estimation and detection in array processing," in *Radar Array Processing*, S. Haykin, Ed. New York: Springer-Verlag, 1993, ch. 4.
- [6] J. F. Böhme, "Estimation of source parameters by maximum likelihood and nonlinear regression," in *Proc. ICASSP 84*, 1984, pp. 7.3.1-7.3.4.
- [7] M. Wax, "Detection and estimation of superimposed signals," Ph.D. dissertation, Stanford Univ., CA, Mar., 1985.
- [8] I. Ziskind and M. Wax, "Maximum likelihood localization of multiple sources by alternating projection," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 36, no. 10, pp. 1553-1560, Oct. 1988.
- [9] P. Stoica and A. Nehorai, "MUSIC, maximum likelihood, and Cramér-Rao Bound," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, no. 5, pp. 720-741, May 1989.
- [10] —, "MUSIC, maximum likelihood, and Cramér-Rao bound: Further results and comparisons," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 38, no. 12, pp. 2140-2150, Dec. 1990.
- [11] J. Ward, "High throughput packet radio networks with adaptive antenna arrays," Ph.D. thesis, Ohio State Univ., Columbus, OH, 1990.
- [12] J. Ward and R. T. Compton, Jr., "Improving the performance of a slotted ALOHA packet radio network with an adaptive array," *IEEE Trans. Commun.*, vol. 40, no. 2, pp. 292-300, Feb. 1992.

- [13] W. C. Y. Lee, *Mobile Communications Engineering*. New York: McGraw-Hill, 1982.
- [14] W. A. Gardner and C. Chen, "Signal-selective time-difference-of-arrival estimation for passive location of man-made signal sources in highly corruptive environments, Part I: Theory and method," *IEEE Trans. Signal Processing*, vol. 40, no. 5, pp. 1168–1184, May 1992.
- [15] C. Chen and W. A. Gardner, "Signal-selective time-difference-of-arrival estimation for passive location of man-made signal sources in highly corruptive environments, Part I: Algorithms and performance," *IEEE Trans. Signal Processing*, vol. 40, no. 5, pp. 1185–1197, May 1992.
- [16] G. Xu and T. Kailath, "Direction-of-arrival estimation via exploitation of cyclostationarity—A combination of temporal and spatial processing," *IEEE Trans. Signal Processing*, vol. 40, pp. 1175–1185, July 1992.
- [17] R. Gooch and J. Lundell, "The CM array: An adaptive beamformer for constant modulus signals," in *Proc. ICASSP 86*, May 1986.
- [18] J. Li and R. T. Compton, Jr., "Maximum likelihood angle estimation for signals with known waveforms," *IEEE Trans. Signal Processing*, vol. 41, no. 9, pp. 2850–2862, Sept. 1993.
- [19] J. Li, B. Halder, P. Stoica, M. Viberg, and T. Kailath, "Decoupled maximum likelihood angle estimation for signals with known waveforms," Chalmers Univ. of Technol., Gothenburg, Sweden, Tech. Rep. no. CTH-TE-8, Feb. 1994.
- [20] P. Stoica and K. C. Sharman, "Maximum likelihood methods for direction-of-arrival estimation," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 38, no. 7, pp. 1132–1143, July 1990.
- [21] W. G. Chen, J. P. Reilly, and K. M. Wong, "Detection of the number of signals in noise with unknown nonwhite covariance matrices," in *Proc. ICASSP 92*, 1992, pp. V-377–V-380.
- [22] Q. Wu, K. M. Wong, and J. P. Reilly, "Maximum likelihood estimation for array processing in unknown noise environment," in *Proc. ICASSP 92*, 1992, pp. V-241–V-244.
- [23] L. Ljung, *System Identification: Theory for the User*. Englewood Cliffs, NJ: Prentice-Hall, 1987.
- [24] T. Söderström and P. Stoica, *System Identification*. London: Prentice-Hall, 1989.
- [25] H. L. Van Trees, *Detection, Estimation, and Modulation Theory, Part I*. New York: Wiley, 1968.
- [26] G. W. Stewart, *Introduction to Matrix Computations*. New York: Academic, 1973.
- [27] 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, no. 5, pp. 1081–1089, Oct. 1986.
- [28] P. Stoica and K. C. Sharman, "Novel eigenanalysis method for direction estimation," *Proc. Inst. Elec. Eng. Pt. F*, vol. 137, no. 1, pp. 19–26, Feb. 1990.
- [29] M. Viberg and B. Ottersten, "Sensor array processing based on subspace fitting," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 39, no. 5, pp. 1110–1121, May 1991.
- [30] M. Viberg, B. Ottersten, and T. Kailath, "Detection and estimation in sensor arrays using weighted subspace fitting," *IEEE Trans. Signal Processing*, vol. 39, no. 11, pp. 2436–2449, Nov. 1991.
- [31] A. Graham, *Kronecker Products and Matrix Calculus with Applications*. Chichester, UK: Ellis Horwood, 1981.



Jian Li (S'87–M'91) received the M.Sc. and Ph.D. degrees in electrical engineering from The Ohio State University, Columbus, in 1987 and 1991, respectively.

From April 1991 to June 1991, she was an Adjunct Assistant Professor with the Department of Electrical Engineering, The Ohio State University. From July 1991 to June 1993, she was an Assistant Professor with the Department of Electrical Engineering, University of Kentucky, Lexington. Since August 1993, she has been an Assistant Professor with the Department of Electrical Engineering, University of Florida, Gainesville. Her current research interests include sensor array signal processing, synthetic aperture radar image formation and understanding, radar detection and estimation theory, image segmentation and processing, and communications.

Dr. Li is a member of Sigma Xi and Phi Kappa Phi. She received the 1994 National Science Foundation Young Investigator Award.



Bijit Halder received the B.Tech degree in electronics and electrical communication engineering from the Indian Institute of Technology, Kharagpur, India, in 1989, and the M.E. degree in electrical communication engineering from the Indian Institute of Science, Bangalore, India, in 1991. He is currently working toward the Ph.D. degree in electrical engineering at Stanford University, Stanford, CA. His areas of interest include statistical signal processing and performance analysis of linear and nonlinear adaptive algorithms.



Petre Stoica (F'94) received the M.Sc. and D.Sc. degrees, both in automatic control, from the Bucharest Polytechnic Institute, Romania, in 1972 and 1979, respectively.

Since 1972, he has been with the Department of Automatic Control and Computers at the Polytechnic Institute of Bucharest, Romania, where he holds the position of Professor of System Identification and Signal Processing. He spent the years 1992, 1993, and the first half of 1994 with the Systems and Control Group at Uppsala University in Sweden as a Guest Professor. In 1993, he was awarded an Honorary Doctorate by Uppsala University, Sweden. From July 1994 to December 1994, he held a Chalmers 150th Anniversary Visiting Professorship with the Applied Electronics Department at Chalmers University of Technology, Gothenburg, Sweden. Currently, he is a guest professor in the Systems and Control Group of Uppsala University.

Dr. Stoica is a Corresponding Member of the Romanian Academy.



Mats Viberg (S'87–M'90) was born in Linköping, Sweden, on December 21, 1961. He received the M.S. degree in applied mathematics in 1985, the Lic. Eng. degree in 1987, and the Ph.D. degree in electrical engineering in 1989, all from Linköping University, Sweden.

He joined the Division of Automatic Control at the Department of Electrical Engineering, Linköping University, in 1984, and from November 1989 until August 1993, he was a Research Associate.

During October 1988 to March 1989, he was on leave at the Informations Systems Laboratory, Stanford University, as a Visiting Scholar. From August 1992 until August 1993, he held a Fulbright-Hayes grant scholarship as a Visiting Researcher at the Department of Electrical and Computer Engineering, Brigham Young University, Provo, UT, and at the Informations Systems Laboratory, Stanford University. Since September 1993, he has been a Professor of Signal Processing at the Department of Applied Electronics, Chalmers University of Technology, Sweden. His research interests are in statistical signal processing and its application to sensor array signal processing, system identification, and communication.

Dr. Viberg received the IEEE Signal Processing Society's 1993 Paper Award in the Statistical Signal and Array Processing Area for the paper "Sensor Array Processing Based on Subspace Fitting," which he coauthored with B. Ottersten.