

# Nonuniform Array Processing Via the Polynomial Approach

ANTHONY J. WEISS, Senior Member, IEEE  
Saxpy Computer Corporation

ALAN S. WILLSKY, Fellow, IEEE  
M.I.T.

BERNARD C. LEVY  
University of California

A polynomial approach for maximum likelihood (ML) estimation of superimposed signals in time series problems and array processing was recently proposed [1]–[3]. This technique was applied successfully to linear uniform arrays and to uniformly sampled complex exponential signals. However, uniformly spaced arrays are not optimal for minimum variance estimation of bearing, range, or position, and uniform sampling of signals is not always possible in practice. In this communication we make use of the expectation-maximization (EM) algorithm in order to apply the polynomial approach to sublattice arrays and to missing samples in time series problems.

Manuscript received June 19, 1987; revised February 4, and April 25, 1988.

IEEE Log No. 25816.

The work of A. Weiss was supported by the Army Research Office under Contract DAAL03-86-C-0018.

The work of A. Willsky and B. Levy was supported in part by the National Science Foundation under Grant ECS-8312921 and in part by the Army Research Office under Grant DAAG-84-K-0005 and DAAL03-86-K-1071.

Authors' addresses: A. J. Weiss, Saxpy Computer Corp., 255 San Geronimo Way, Sunnyvale, CA 94086; A. S. Willsky, Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, MA 02139; B. C. Levy, Dep't. of Electrical Engineering and Computer Science, University of California, Davis, CA 95616.

0018-9251/89/0100-0048 \$1.00 © 1989 IEEE

## I. INTRODUCTION

The estimation of multiple superimposed exponential signals in noise is of interest in time series analysis and in array processing. Recently an effective technique for computing the maximum likelihood (ML) estimates of the signals was introduced by Bresler and Macovski [1] and Kumaresan, Scharf, and Shaw [2], [3]. We refer to this technique as the *polynomial approach* since it is based on expressing the ML criterion in terms of the prediction polynomial of the noiseless signal. The polynomial approach relies on the assumption that the array of sensors is uniformly spaced. It is well known [4] that the optimal sensor configuration is not uniform under many reasonable criteria. For example, minimum bearing variance is obtained by placing half of the sensors (with a spacing of half the design wavelength) at each end of the given aperture; minimum range variance is obtained by placing one-fourth of the elements at each end and half in the middle; and optimal position estimation is obtained by placing one-third of the sensors at each end and the middle. Furthermore, when operating long uniform arrays, often some of the sensors do not function and their outputs must be ignored, yielding in effect a sublattice array. We present a method for extending the polynomial approach to sublattice arrays. We treat the sublattice array output as an incomplete data observation. Therefore the expectation-maximization (EM) algorithm is directly applicable. This algorithm was only recently applied to array processing problems by Feder and Weinstein [5]. However, in [5] the EM algorithm is used to estimate one signal at a time, while here it is employed to enable the use of the polynomial approach that estimates all the signals simultaneously. Since both the polynomial approach and the EM algorithm are not widely known, the basic principles of each of these techniques are briefly reviewed here for clarity.

Note that although we concentrate on the array problem, all the results that we describe are equally applicable to the corresponding time series problem discussed in [1], namely, the estimation of superimposed complex exponential signals in noise.

This paper is organized as follows. The polynomial approach for processing data collected over a uniform array is reviewed in Section II. In Section III the EM algorithm is briefly described with reference to sublattice arrays. Section IV presents the proposed technique. Simulation results of our procedure are presented in Section V, and Section VI contains some conclusions.

## II. UNIFORM ARRAYS AND THE POLYNOMIAL APPROACH

Consider  $N$  narrowband radiating sources observed by a linear *uniform* array composed of  $M$  sensors.

The sources are assumed to be far enough from the array, compared with the array length so that the signal wavefronts are effectively planar over the array. The signal at the output of the  $m$ th sensor can be expressed by

$$x_m(t) = \sum_{n=1}^N s_n(t - (m-1)\tau_n) + v_m(t), \quad (1)$$

$$m = 1, 2, \dots, M, \quad -T/2 \leq t \leq T/2$$

where  $\{s_n(t)\}_{n=1}^N$  are the radiated signals,  $\{v_m(t)\}_{m=1}^M$  are additive noise processes, and  $T$  is the observation interval. The delay of the  $n$ th wavefront at the  $m$ th sensor, relative to the first sensor, is given by  $(m-1)\tau_n$ .

If  $d$  denotes the sensor spacing,  $c$  the propagation velocity, and  $\gamma_n$  the source bearing with respect to the array perpendicular, the parameter  $\tau_n$  can be expressed as

$$\tau_n = (d/c)\sin(\gamma_n).$$

A convenient separation of the parameters  $\{\tau_n\}_{n=1}^N$  to be estimated may be obtained by using Fourier coefficient, defined by

$$X_m = \frac{1}{T} \int_{-T/2}^{T/2} x_m(t) e^{-j\omega_0 t} dt.$$

Since we assume that the spectrum of the signals is concentrated around  $\omega_0$ , with a bandwidth that is small compared with  $2\pi/T$ , a single Fourier coefficient is enough to completely describe the signals. Taking the Fourier coefficients of (1) we obtain

$$X_m = \sum_{n=1}^N S_n e^{-j\omega_0(m-1)\tau_n} + V_m, \quad m = 1, 2, \dots, M \quad (2)$$

where  $S_n$  and  $V_m$  are the Fourier coefficients of  $s_n(t)$  and  $v_m(t)$ , respectively. Equation (2) may be expressed using vector notation as

$$\mathbf{X} = \mathbf{A}\mathbf{S} + \mathbf{V} \quad (3)$$

where

$$\mathbf{X} \triangleq [X_1, X_2, \dots, X_M]^T$$

$$\mathbf{S} \triangleq [S_1, S_2, \dots, S_N]^T$$

$$\mathbf{V} \triangleq [V_1, V_2, \dots, V_M]^T$$

$$\mathbf{A} \triangleq [\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_N]$$

$$\mathbf{a}_n \triangleq [1, \lambda_n, \lambda_n^2, \dots, \lambda_n^{M-1}]^T, \quad n = 1, 2, \dots, N$$

$$\lambda_n \triangleq e^{-j\omega_0 \tau_n}.$$

In general, the estimation procedure relies on more than one realization of (3), corresponding for example to several time samples or observation intervals.

In that case we use the index  $j$  to denote different realizations:

$$\mathbf{X}_j = \mathbf{A}\mathbf{S}_j + \mathbf{V}_j, \quad j = 1, 2, \dots, J. \quad (4)$$

Note that in this form the parameters of interest are embedded in the matrix  $\mathbf{A}$  and they are separated from the amplitude and phase of the signals which are given by  $\mathbf{S}_j$ . Instead of estimating  $\{\tau_n\}$  directly, we concentrate on estimating  $\{\lambda_n\}_{n=1}^N$ . Under the assumption that the vectors  $\{\mathbf{V}_j\}_{j=1}^J$  are independent identically distributed (IID), zero mean, and Gaussian with covariance  $\sigma^2 \mathbf{I}$ , the ML estimates are given by

$$\{\hat{\lambda}_n\}_{n=1}^N = \arg \min_{\lambda_n \in UC} \{R\}; \quad (5)$$

$$R \triangleq \sum_{j=1}^J \|\mathbf{X}_j - \mathbf{A}\mathbf{S}_j\|^2$$

where  $\|\cdot\|$  denotes the Euclidean norm and  $UC$  stands for the unit circle which is the parameter space, in this case.

The minimization required in (5) is not trivial since the vectors  $\{\mathbf{S}_j\}$  and the matrix  $\mathbf{A}$  are not known to the observer. However, whenever  $\mathbf{A}$  is known (i.e.,  $\{\lambda_n\}_{n=1}^N$  are known),  $R$  is minimized by choosing

$$\hat{\mathbf{S}}_j = (\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H \mathbf{X}_j \quad (6)$$

as the estimate of  $\mathbf{S}_j$ , for  $j = 1, 2, \dots, J$ , where  $(\cdot)^H$  represents the Hermitian-transpose operation. Substituting (6) in (5) we obtain

$$R = \sum_{j=1}^J \|\mathbf{X}_j - \mathbf{A}(\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H \mathbf{X}_j\|^2 = \sum_{j=1}^J \mathbf{X}_j^H \mathbf{P}_B \mathbf{X}_j \quad (7)$$

where

$$\mathbf{P}_B \triangleq \mathbf{I} - \mathbf{A}(\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H.$$

The polynomial approach relies on the introduction of the polynomial  $b(z) = b_0 z^N + b_1 z^{N-1} + \dots + b_N$ , whose zeros are the parameters of interest  $\{\lambda_n\}_{n=1}^N$ . Observe that by definition the  $M \times (M-N)$  Toeplitz matrix  $\mathbf{B}$  defined by

$$\mathbf{B}^H \triangleq \begin{bmatrix} b_N & b_{N-1} & \cdots & b_0 & 0 \\ & \ddots & \ddots & & \ddots \\ 0 & & b_N & b_{N-1} & \cdots & b_0 \end{bmatrix}$$

is orthogonal to  $\mathbf{A}$ , i.e.,  $\mathbf{B}^H \mathbf{A} = 0$  and hence  $\mathbf{P}_B = \mathbf{B}(\mathbf{B}^H \mathbf{B})^{-1} \mathbf{B}^H$ . Now the minimization in (5) can be expressed in terms of the coefficients  $\{b_i\}_{i=0}^N$  as

$$\mathbf{b} = \arg \min_{\mathbf{b} \in \theta_b} \sum_{j=1}^J \mathbf{X}_j^H \mathbf{B}(\mathbf{B}^H \mathbf{B})^{-1} \mathbf{B}^H \mathbf{X}_j, \quad (8)$$

where  $\mathbf{b} = [b_N, b_{N-1}, \dots, b_0]^T$ , and  $\theta_b$  is the space of all the vectors whose associated polynomials have zeros only on the unit circle.

The algorithm for the minimization required in (8) is based on the relation

$$B^H X_j = \tilde{X}_j \mathbf{b} \quad (9)$$

where  $\tilde{X}_j$  is the  $(M - N) \times (N + 1)$  matrix defined by

$$\tilde{X}_j \triangleq [X_j(N + 1 : M), X_j(N : M - 1), \dots, X_j(1 : M - N)]$$

and  $X_j(k : r)$  describes a subvector of  $X_j$  consisting of all the components from the  $k$ th component to the  $r$ th component. Substituting (9) in (8) we obtain

$$\mathbf{b} = \arg \min_{\mathbf{b} \in \theta_b} \mathbf{b}^H C \mathbf{b}; \quad C \triangleq \sum_{j=1}^J \tilde{X}_j^H (B^H B)^{-1} \tilde{X}_j. \quad (10)$$

This relation is used in the minimization algorithm [1–3]. The algorithm starts with any initial estimate  $\mathbf{b}^{(0)}$  of  $\mathbf{b}$  and proceeds as follows.

- 1) Initialization.  $k = 0$ ,  $\mathbf{b}^{(k)} = \mathbf{b}^{(0)}$ .
- 2) Compute  $C^{(k)}$  according to (10) using  $\mathbf{b}^{(k)}$  to construct the matrix  $B^{(k)}$ .
- 3) Find  $\mathbf{b}^{(k+1)} = \arg \min_{\mathbf{b} \in \theta_b} \mathbf{b}^H C^{(k)} \mathbf{b}$ .
- 4) Check convergence of  $\mathbf{b}$ . If no:  $k = k + 1$ , go back to 2). If yes: continue.
- 5) Find the roots of the polynomial  $b(z)$  whose coefficients are given by  $\mathbf{b}^{(k+1)}$ .

Note that this algorithm can be employed only if the number of signals is known. In practice one must first estimate  $N$ , the number of signals present, and only then use the above algorithm. A method for estimating  $N$  is described in [7].

### III. SUBLATTICE ARRAYS AND THE EM ALGORITHM

We are primarily interested here in the problem where the measurements are taken along a sublattice array of  $M'$  sensors. The sublattice array may be described by a binary vector  $l$  of length  $M$ . The  $m$ th component of  $l$  is 1 if the  $m$ th sensor of the full array is part of the subarray and it is zero if the sensor is missing. Equation (4) may be converted to describe a sublattice array through a left-multiplication by a transformation matrix  $G$ . The  $M' \times M$  matrix  $G$  is constructed by eliminating all the zero rows in  $\text{diag}(l)$ . For example an array of three elements in positions 1, 2, 5 is described by  $l^T = (1, 1, 0, 0, 1)$  and

$$G = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

Multiplying (4) by  $G$  we obtain, for a given sublattice array, the equation

$$\mathbf{Y}_j = G \mathbf{X}_j = G(AS_j + \mathbf{V}_j), \quad j = 1, 2, \dots, J. \quad (11)$$

We refer to  $\{X_j\}$  as the (unavailable) *complete data* and to  $\{Y_j\}$  as the *observed data*.

Let  $\mathbf{Y} = [\mathbf{Y}_1^T, \mathbf{Y}_2^T, \dots, \mathbf{Y}_J^T]^T$  and  $\mathbf{X} = [\mathbf{X}_1^T, \mathbf{X}_2^T, \dots, \mathbf{X}_J^T]^T$  denote, respectively, the observation vector, and the complete data vector. From (11) they are related by

$$\mathbf{Y} = F \mathbf{X} \quad (12)$$

where

$$F \triangleq \text{diag}\{G\}$$

is a block diagonal matrix with  $J$  blocks of  $G$ . The complete data vector  $\mathbf{X}$  is Gaussian with given covariance  $\sigma^2 I$  and unknown mean  $\theta$ . The parameter vector  $\theta$  is defined by

$$\theta \triangleq [\theta_1^T, \theta_2^T, \dots, \theta_J^T]^T$$

where

$$\theta_j \triangleq AS_j.$$

If  $f_{\mathbf{x}}(\mathbf{X} | \theta)$  is the density of  $\mathbf{x}$  given  $\theta$ , we have therefore

$$\ln\{f_{\mathbf{x}}(\mathbf{X} | \theta)\} = -MJ \ln(\pi \sigma^2) - \|\mathbf{X} - \theta\|^2 / \sigma^2 \quad (13)$$

and the ML estimate of  $\theta$  given  $\mathbf{X}$  is then easy to compute. In fact, it requires the minimization of

$$\|\mathbf{X} - \theta\|^2 = \sum_{j=1}^J \|\mathbf{X}_j - AS_j\|^2 \quad (14)$$

and it was shown in Section II how the polynomial approach could be used to perform this minimization.

When we are only given the observation vector  $\mathbf{Y}$  corresponding to an incomplete data set, if  $f_{\mathbf{y}}(\mathbf{Y} | \theta)$  denotes the density of  $\mathbf{y}$  given  $\theta$ , the ML estimate of  $\theta$  given  $\mathbf{Y}$  is

$$\theta = \arg \max_{\theta \in \Theta} f_{\mathbf{y}}(\mathbf{Y} | \theta) = \arg \max_{\theta \in \Theta} \ln\{f_{\mathbf{y}}(\mathbf{Y} | \theta)\} \quad (15)$$

where  $\Theta$  is the parameter space. However  $\ln\{f_{\mathbf{y}}(\mathbf{Y} | \theta)\}$  cannot be expressed as simply as in (13), (14), and the maximization of  $\ln\{f_{\mathbf{y}}(\mathbf{Y} | \theta)\}$  is therefore more difficult to achieve.

The EM approach [6] to the ML estimation problem consists of estimating the complete data vector  $\mathbf{X}$  from the given observation vector  $\mathbf{y}$  and then substituting the estimate  $\mathbf{X}$  in (14) to perform the minimization over the parameter space  $\Theta$ . However, since  $\mathbf{X}$  depends, in general, on  $\theta$  as well as  $\mathbf{Y}$ , several iterations of the above procedure are necessary in order for the parameter  $\theta$  to converge. A rigorous justification of the EM algorithm is as follows. First from Bayes' rule

$$\ln\{f_{\mathbf{y}}(\mathbf{Y} | \theta)\} = \ln\{f_{\mathbf{x}}(\mathbf{X} | \theta)\} - \ln\{f_{\mathbf{x}|\mathbf{Y}}(\mathbf{X} | \mathbf{Y}, \theta)\}. \quad (16)$$

Taking the expectation of (16) over  $\mathbf{x}$  given  $\mathbf{Y}$  and under the assumption that the parameter vector is

equal to  $\theta'$ , we obtain

$$L(\theta) \triangleq \ln\{f_Y(Y|\theta)\} = Q(\theta|\theta') - H(\theta|\theta') \quad (17)$$

where

$$Q(\theta|\theta') \triangleq E\{\ln\{f_X(X|\theta)\} | Y, \theta'\}$$

$$H(\theta|\theta') \triangleq E\{\ln\{f_{X|Y}(X|Y, \theta)\} | Y, \theta'\}.$$

Using Jensen's inequality it is easy to verify that

$$H(\theta|\theta') \leq H(\theta'|\theta'). \quad (18)$$

The EM algorithm may be described by the following sequence [6].

- 1) Initialization. Set  $p = 0$ , and  $\theta^{(p)} = \theta_0$ .
- 2) E-step. Determine  $Q(\theta|\theta^{(p)})$ .
- 3) M-step. Choose  $\theta^{(p+1)}$  to be the value of  $\theta \in \Theta$  that maximizes  $Q(\theta|\theta^{(p)})$ .
- 4) Check the convergence of  $\theta$ . If no,  $p = p + 1$ , go back to 2). If yes, stop.

In every cycle of the algorithm the likelihood function  $L(\theta)$  is increased, since

$$L(\theta^{(p+1)}) = Q(\theta^{(p+1)}|\theta^{(p)}) - H(\theta^{(p+1)}|\theta^{(p)})$$

$$\geq Q(\theta^{(p)}|\theta^{(p)}) - H(\theta^{(p)}|\theta^{(p)}) = L(\theta^{(p)})$$

where the inequality holds due to (18) and due to the M-step.

In Section IV we apply this algorithm to the sublattice array problem. Step 3) is solved using the polynomial approach described in Section II.

#### IV. PROPOSED ALGORITHM

The application of this rather general algorithm to the problem at hand requires only the determination of  $Q(\theta|\theta')$ . From (13), and using the expression

$$\hat{X} = E\{X|Y, \theta'\} = \theta' + F^H(F F^H)^{-1}(Y - F\theta') \quad (19)$$

for the conditional mean of  $x$ , we find that

$$Q(\theta|\theta') = K - \|\hat{X} - \theta\|^2/\sigma^2 \quad (20)$$

where  $K$  consists of terms independent of  $\theta$ . Thus, as was claimed above, the maximization of  $Q(\theta|\theta')$  reduces to the minimization of

$$R_1 = \|\hat{X} - \theta\|^2 = \sum_{j=1}^J \|\hat{X}_j - AS_j\|^2 \quad (21)$$

and the M-step of the EM algorithm may be performed by using the polynomial approach to minimize (21).

The estimation step (19) of the EM algorithm can also be simplified further by using the block diagonal

structure of  $F$  and the relations  $GG^H = I$  and  $G^H G = \text{diag}(I)$  to rewrite (19) as

$$\hat{X}_j = \text{diag}(\bar{I})\theta'_j + G^H Y_j \quad (22)$$

where  $\bar{I}$  is the complement of  $I$  (zeros and ones are interchanged). The parameter vector  $\theta'_j$  is simply the estimate of  $AS_j$  obtained in the previous cycle and therefore (22) may be written also as

$$\hat{X}_j^{(p+1)} = \text{diag}(\bar{I})\{A(A^H A)^{-1}A^H \hat{X}_j\}^{(p)} + G^H Y_j$$

$$= \text{diag}(\bar{I})\{(I - B(B^H B)^{-1}B^H)\hat{X}_j\}^{(p)} + G^H Y_j$$

using the notation of the polynomial approach. As one would expect (22) states that the components of  $\hat{X}_j$  that correspond to existing sensors are always equal to the observed data, *i.e.*, the corresponding components of  $Y_j$ .

The proposed EM algorithm may be summarized as follows.

- 1) Initialization. Select initial values for  $\{\lambda_n\}_{n=1}^N$ ; find the corresponding  $\mathbf{b}^{(0)}$ . Compute:  $A_1 = GA$ ;  $S_j = (A_1^H A_1)^{-1}A_1^H Y_j$ ;  $\hat{X}_j^{(0)} = \text{diag}(\bar{I})AS_j + G^H Y_j$  (see (22)). Set:  $p = 0$ .
- 2) Use the minimization algorithm for uniform arrays.
  - a) Construct  $\tilde{X}_j = [\hat{X}_j^{(p)}(N+1:M), \dots, \hat{X}_j^{(p)}(1:M-N)]$ . Set  $k = 0$ ,  $\mathbf{b}_1^{(0)} = \mathbf{b}^{(p)}$ .
  - b) Construct  $B$  using  $\mathbf{b}_1^{(k)}$ . Compute  $C = \sum_{j=1}^J \tilde{X}_j^H (B^H B)^{-1} \tilde{X}_j$ .
  - c) Compute  $\mathbf{b}_1^{(k+1)} = \arg \min_{\mathbf{b}_1 \in \Theta_b} \mathbf{b}_1^H C \mathbf{b}_1$ .
  - d) Check convergence of  $\mathbf{b}_1$ . If no:  $k = k + 1$ , go back to 2)b). If yes:  $\mathbf{b}^{(p)} = \mathbf{b}_1^{(k+1)}$ , continue.
- 3) Construct  $B$  using  $\mathbf{b}^{(p)}$ . Compute  $\hat{X}_j^{(p+1)} = \text{diag}(\bar{I})(I - B(B^H B)^{-1}B^H)\hat{X}_j^{(p)} + G^H Y_j$ .
- 4) Check the convergence of  $\hat{X}_j$ . If no:  $p = p + 1$ , go back to 2). If yes: continue.
- 5) Find the roots of the polynomial  $\mathbf{b}^{(p)}(z)$  whose coefficients are given by  $\mathbf{b}^{(p)}$ . These roots are our estimates of the parameters of interest  $\{\lambda_n\}_{n=1}^N$ .

Once we have the estimates of  $\{\lambda_n\}_{n=1}^N$ , it is easy to obtain the estimates of the direction of arrival of the signals.

#### V. EXPERIMENTS

Two sets of experiments are described here. One uses signals  $S_j$  drawn from a normal distribution, and the other uses deterministic signals.

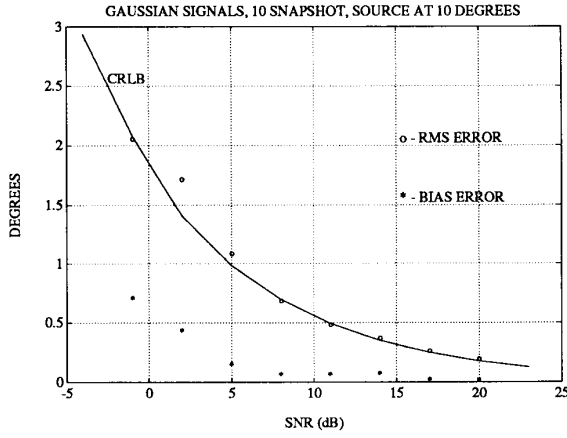


Fig. 1. Bias and rms error for Gaussian source at  $10^\circ$ , based on 200 experiments for each SNR value.

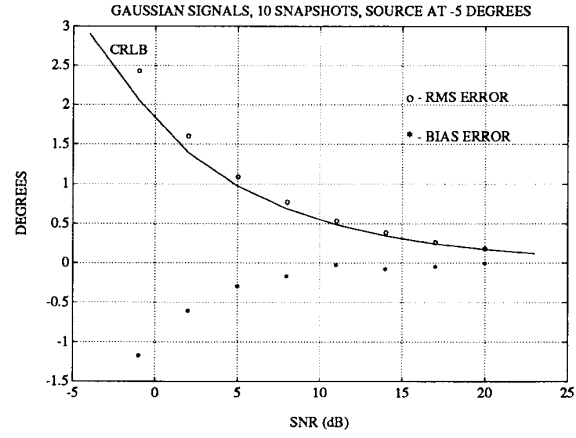


Fig. 2. Bias and rms error for Gaussian source at  $-5^\circ$ , based on 200 experiments for each SNR value.

#### A. Gaussian Signals

To illustrate the behavior of the algorithm, consider a linear uniform array of 6 sensors separated by half a wavelength of the actual narrowband source signals. Now assume that the second, fourth, and fifth sensors are removed, yielding in effect an array described by  $l^T = (1 \ 0 \ 1 \ 0 \ 0 \ 1)$ . Note that only 3 sensors are used, which are separated by one wavelength and 1.5 wavelengths. The Rayleigh resolution criterion for this array is  $23^\circ$ . To show the “super-resolution” properties of the proposed procedure, we place one source at a bearing of  $-5^\circ$  and a second source at a bearing of  $10^\circ$  relative to the perpendicular to the array baseline. The signals  $S_j$  and noise  $V_j$  are random complex Gaussian vectors, with covariance matrices  $\sigma_s^2 \cdot I$  and  $\sigma_n^2 \cdot I$ , respectively. In each experiment  $J = 10$  samples (snapshots) are used. The initial guess in each case is generated by using simple beamforming [7] over the selected field of view of  $[-20^\circ, 20^\circ]$ . Since the sources cannot be resolved by the beamformer we used  $\hat{\gamma}_1^{(0)} = \gamma_b - 1^\circ$ ,  $\hat{\gamma}_2^{(0)} = \gamma_b + 1^\circ$  where  $\hat{\gamma}_i^{(0)}$  is the initial guess of the bearing to the  $i$ th source and  $\gamma_b$  is the bearing for which the beamformer obtains its maximum. For each signal-to-noise ratio (SNR) =  $-1, 2, 5, \dots, 20$  (dB) we performed 200 experiments. The SNR is defined by  $10 \log_{10}(\sigma_s^2/\sigma_n^2)$ . The bias and the root mean square (rms) error are displayed in Fig. 1 for the first source, while the results for the second source are displayed in Fig. 2. The Cramer-Rao lower bound (CRLB) was calculated according to the formulas in [8].

The 99.74 percent confidence intervals for the rms error are between 84 and 114 percent of the rms values shown in the figures. The 99.74 percent confidence intervals for the bias are approximately between  $(\hat{b} - 0.21) \cdot \text{rms}$  and  $(\hat{b} + 0.21) \cdot \text{rms}$ , where  $\hat{b}$  and  $\text{rms}$  are the bias and rms error estimates shown in the figures.

We observe that the algorithm is efficient in the sense that it converges to the CRLB even for relatively low SNR. It is interesting to note that an algorithm that is based on modeling the signals as deterministic vectors is efficient, even though the signals are random. Usually, algorithms that are based on modeling the signals as random vectors use additional information, such as the mean and the covariance matrix of the signals [5]. We demonstrated that efficient estimation of direction of arrival, of random signals, can be obtained without a-priori knowledge of the statistics of the signal.

Finally, note that eigenstructure techniques, such as the MUSIC algorithm [7], can perform just as well when many snapshots are observed. However, the proposed technique performs well even for the case of a single snapshot and of coherent signals as demonstrated in the following section.

#### B. Deterministic Signals

Consider the same setting as in the previous section, where we used a linear array of sensors described by  $l^T = (1 \ 0 \ 1 \ 0 \ 0 \ 1)$ , and 2 sources at  $-5^\circ$  and  $10^\circ$  relative to the perpendicular to the array baseline. However, now, the signal vector is fixed at  $S = (1 \ 1)^T$  and the noise  $V$ , is a random complex Gaussian vector with covariance matrix  $\sigma_n^2 \cdot I$ . Only a single snapshot is used and therefore the index  $j$  is suppressed. Again, the initial guess in each experiment is obtained by using simple beamforming over the interval  $[-20^\circ, 20^\circ]$ . The SNR definition is now modified to be  $10 \log_{10}(1/\sigma_n^2)$ . For each SNR =  $0, 3, 6, \dots, 30$  (dB) we performed 200 experiments. The bias and rms errors are displayed in Figs. 3 and 4 for the source at  $10^\circ$  and the source at  $-5^\circ$ , respectively. The CRLB was calculated analytically. Note that the Fisher information matrix contains 36 elements (a  $6 \times 6$  matrix), since we assume that the phase, amplitude,

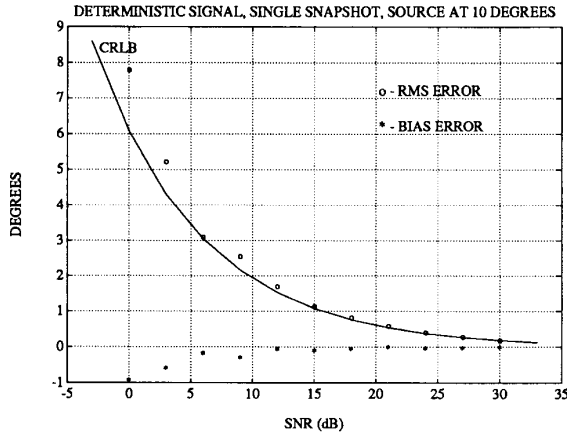


Fig. 3. Bias and rms error for deterministic source at  $10^\circ$ , based on 200 experiments for each SNR value.

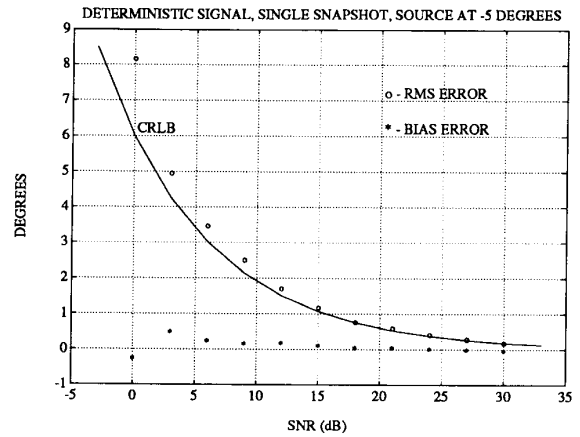


Fig. 4. Bias and rms error for deterministic source at  $-5^\circ$ , based on 200 experiments for each SNR value.

and direction-of-arrival of the signals are unknown. The 99.74 percent confidence intervals are given by the same expressions as in the random signals case.

We observe that the algorithm is efficient even in the case of a single snapshot and perfectly coherent signals. This performance cannot be matched by any of the eigenstructure methods. However, other exact ML techniques are expected to perform just as well. These exact ML techniques are described in [5, 9, and 10]. While the methods in [5, 9, and 10] are not limited to narrowband signals in the farfield of a linear sublattice array, they seem to require a higher computational load. A complete analysis of the computational load of each of these algorithms is not yet available.

## VI. SUMMARY

We have proposed a novel EM algorithm for the estimation of superimposed signals observed by nonuniform arrays. The algorithm is efficient and requires a relatively low computational load even when the number of samples is small and the signals

are perfectly coherent.

Note that convergence theorems exist for the EM method. However, convergence theorems for the polynomial approach are not yet available and therefore further investigation is required to prove the convergence of the proposed technique. The experiments reported here did not reveal any convergence problems.

Finally, we would like to emphasize that the EM algorithm is guaranteed to converge to a local maximum of the likelihood function. Thus we would expect that the algorithm described here will converge to the globally optimum result only if the initial estimates are good enough. Fast initial estimates can be obtained by using simpler methods such as beamforming or the MUSIC techniques (see [7] for a review of these methods).

## ACKNOWLEDGMENT

The authors would like to thank Prof. Y. Bresler for helpful suggestions and discussions.

## REFERENCES

- [1] Bresler, Y., and Macovski, A. (1986)  
Exact maximum likelihood parameter estimation of superimposed exponential signals in noise.  
*IEEE Transactions on Acoustics, Speech and Signal Processing*, ASSP-34, 5 (Oct. 1986), 1081-1089.
- [2] Kumaresan, R., and Shaw, A. K. (1985)  
High resolution bearing estimation without eigen decomposition.  
In *Proceedings of the International Conference on Acoustics, Speech and Signal Processing*, Tampa, Fla., 1985, pp. 576-579.
- [3] Kumaresan, R., Scharf, L. L., and Shaw, A. K. (1986)  
An algorithm for pole-zero modeling and spectral analysis.  
*IEEE Transactions on Acoustics, Speech and Signal Processing*, ASSP-34, 3 (June 1986), 637-640.
- [4] Carter, G. C. (1987)  
Coherence and time delay estimation.  
*Proceedings of the IEEE*, 75, 2 (Feb. 1987), 236-255.
- [5] Feder, M., and Weinstein, E. (1988)  
Parameter estimation of superimposed signals using the EM algorithm.  
*IEEE Transactions on Acoustics, Speech and Signal Processing*, ASSP-36, 4 (Apr. 1988), 477-489.
- [6] Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977)  
Maximum likelihood from incomplete data via the EM algorithm.  
*Journal of the Royal Statistical Society*, B39, (1977), 1-38.
- [7] Wax, M. (1985)  
Detection and estimation of superimposed signals.  
Ph.D. dissertation, Stanford University, Stanford, Calif., 1985.
- [8] White, F. M. (1984)  
Performance of Bayes-optimal angle-of-arrival estimators.  
Technical report 654, Lincoln Laboratory, M.I.T., Cambridge, Aug. 1984.
- [9] Weiss, A. J., Willsky, A. S., and Levy, B. C. (1988)  
Maximum likelihood array processing for the estimation of superimposed signals.  
*Proceeding of the IEEE*, 76, 2 (Feb. 1988), 203-205.
- [10] Ziskind, I., and Wax, M. (1987)  
Maximum likelihood estimation via the alternating projection maximization algorithm.  
In *Proceedings of the International Conference on Acoustics, Speech and Signal Processing*, Dallas, Tex., Apr. 1987, pp. 2280-2283.



**Anthony J. Weiss** (S'84—M'86—SM'86) was born in London, England, in 1951. He received the B.Sc. (Cum Laude) from the Technion-Israel Institute of Technology, Haifa, Israel, in 1973, and the M.Sc. (Summa Cum Laude) and Ph.D. (Summa Cum Laude) degrees from Tel-Aviv University, Israel, in 1982 and 1985, all in electrical engineering.

From 1973 to 1983 he was involved in research and development of numerous projects in the fields of communications, tracking systems, command and control, and emitter localization. In 1985 he became a faculty member of the Department of Electronic Systems, Tel-Aviv University. He was also a consultant for the Department of Defence, the EW Department of Tadiran Inc. and for Elta Electronics, a subsidiary of Israel Aircraft Industries (IAI). During the academic year 1986/1987 he was a Visiting Scientist with the Laboratory for Information and Decision Systems, M.I.T., Cambridge, Mass. In 1987 he joined Saxpy Computer Corporation, Sunnyvale, Calif. His research activities are detection and estimation theory, signal processing, array processing, image processing, and parallel processing architectures, with applications to radar, sonar, and passive surveillance systems.

Dr. Weiss held a Rothschild Foundation Fellowship from 1986 to 1987 and a Ygaal Alon Fellowship from 1985 to 1986. He was a co-recipient of the IEEE Acoustics, Speech, and Signal Processing Society's 1983 Senior Award for the paper, "Fundamental Limitations in Passive Time Delay Estimation."

**Alan S. Willsky** (S'70—M'73—SM'82—F'86) received the S.B. and the Ph.D. degrees from M.I.T. in 1969, and 1973, respectively.

From 1969 through 1973 he held a Fannie and John Hertz Foundation Fellowship. He joined the M.I.T. faculty in 1973 and his present position is Professor of Electrical Engineering. From 1974 to 1981 Dr. Willsky served as Assistant Director of the M.I.T. Laboratory for Information and Decision Systems. He is also a founder and member of the board of directors of Alphatech, Inc. Dr. Willsky has held visiting positions at Imperial College, London, and L'Université de Paris-Sud. He is Editor of the M.I.T. Press series on signal processing, optimization, and control, was program chairman for the 17th IEEE Conference on Decision and Control, has been an associate editor of several journals including the *IEEE Transactions on Automatic Control*, is a member of the Board of Governors of the IEEE Control Systems Society, and was program chairman for the 1981 Bilateral Seminar on Control Systems held in the People's Republic of China. Also Dr. Willsky gave the opening plenary lecture at the 20th IEEE Conference on Decision and Control.



Dr. Willsky is the author of the research monograph *Digital Signal Processing and Control and Estimation Theory* and is co-author of the undergraduate text *Signals and Systems*. In 1975 he received the Donald P. Eckman Award from the American Automatic Control Council. He was awarded the 1979 Alfred Noble Prize by the ASCE and the 1980 Browder J. Thompson Memorial Prize Award by the IEEE for a paper excerpted from his monograph. Dr. Willsky's present research interests are in problems involving abrupt changes in signals and systems, multidimensional estimation, decision-directed signal processing, and the asymptotic analysis of control and estimation systems.

**Bernard C. Levy** (S'74—M'78) was born in Princeton, N.J., on July 31, 1951. He received the diploma of Ingénieur Civil des Mines from the Ecole Nationale Supérieure des Mines in Paris, France, and the Ph.D. in electrical engineering from Stanford University, Stanford, Calif.



While at Stanford University, he held an INRIA Fellowship, and worked also as Teaching Assistant, Research Assistant, and Instructor. From June 1979 to June 1987, he was Assistant, and then Associate Professor in the Department of Electrical Engineering and Computer Science at M.I.T. Since July 1987, he has been Associate Professor in the Department of Electrical Engineering and Computer Science at the University of California, Davis. During the past two years, he has also been a consultant for the Charles Stark Draper Laboratory in Cambridge, Mass. His research interests are in the areas of multidimensional and statistical signal processing, inverse problems, estimation, detection, and scientific computing.