June 1987 LIDS-P-1672

NONUNIFORM ARRAY PROCESSING VIA THE POLYNOMIAL APPROACH

by

Anthony J. Weiss*
Alan S. Willsky **
Bernard C. Levy**

Laboratory for Information and Decision Systems
Massachusetts Institute of Technology
Cambridge, MA 02139

ABSTRACT

A polynomial approach for maximum likelihood 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 EM algorithm in order to apply the polynomial approach to sublattice arrays and to missing samples in time series problems.

*The work of the first author was supported by a Rothschild fellowship.

**The work of these authors was supported in part by the National Science Foundation under Grant ECS-8312921 and in part by the Army Research Office under Grants DAAG-84-K-0005 and DAAL03-86-K-1071.

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 criterion. For example, minimum bearing variance is obtained by placing half of the sensors (with a spacing of half of 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. In this paper 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 EM (Expectation-Maximization) 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 which 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 described in Section II. In Section III it is shown how the EM algorithm can be used to adapt the polynomial approach to the case of sublattice arrays. Several examples of our procedure are presented in Section IV, and Section V 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 to 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

$$x_m(t) = \sum_{n=1}^{N} s_n(t-(m-1)\tau_n) + v_m(t);\ m = 1,2,\ldots,M;\ -T/2 \leq t \leq T/2, \quad (1)$$

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 γ_n the source bearing with respect to the array perpendicular, the parameter τ_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 coefficients, 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 ω_0,

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

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

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

$$\underline{X} = A\underline{S} + \underline{V} \tag{3}$$

where

$$\underline{X} = [X_1, X_2,\ldots,X_M]^T,$$

$$\underline{S} = [S_1, S_2,\ldots,S_N]^T,$$

$$\underline{V} = [V_1, V_2,\ldots,V_M]^T,$$

$$A = [\underline{a}_1, \underline{a}_2,\ldots,\underline{a}_N],$$

$$\underline{a}_n = [1, \lambda_n, \lambda_n^2,\ldots,\lambda_n^{M-1}]^T, \qquad n = 1,2,\ldots,N;$$

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

In general, the estimation procedure relies on more than one

realization of equation (3), corresponding for example to several time samples or observation intervals. In that case we use the index j to denote the different realizations:

$$\underline{X}_j = A\underline{S}_j + \underline{V}_j \qquad j = 1,2,\ldots,J. \tag{4}$$

Instead of estimating $\{\tau_n\}$ directly we concentrate on estimating $\{\lambda_n\}_{n=1}^N$. Under the assumption that the vectors $\{\underline{V}_j\}_{j=1}^J$ are i.i.d. zero mean and Gaussian with covariance $\sigma^2 I$, the maximum likelihood estimates are given by

$$\{\hat{\lambda}_n\}_{n=1}^N = \arg\min_{\lambda_n \varepsilon UC} \{R\}; \quad R = \sum_{j=1}^{J} ||\underline{X}_j - A\underline{S}_j||^2, \tag{5}$$

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 $\{\underline{S}_j\}$ and the matrix A are not known to the observer. However, whenever A is known, R is minimized by choosing

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

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

$$R = \sum_{j=1}^{J} ||\underline{X}_j - A(A^H A)^{-1} A^H \underline{X}_j||^2 = \sum_{j=1}^{J} \underline{X}_j^H P_B \underline{X}_j, \tag{7}$$

where

$$P_B = I - A(A^H A)^{-1} A^H.$$

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

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

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

$$\underline{b} = \arg \min_{\underline{b} \in \Theta_b} \sum_{j=1}^{J} \underline{X}_j^H B(B^H B)^{-1} B^H \underline{X}_j \; , \tag{8}$$

where $\underline{b} = [b_N, b_{N-1}, \ldots, b_0]^T$, and Θ_b is the space of all the vectors whose associated polynomials have zeros only on the unit circle. It can be shown that since b(z) has its roots on the unit circle, its coefficient vector is

α-conjugate-symmetric; i.e. $\underline{b} = \alpha[b_0, b_1, \ldots, b_N]^H$ where α is a constant of unit modulus.

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

$$B^H \underline{X}_j = \tilde{X}_j \underline{b} \ , \tag{9}$$

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

$$\tilde{X}_j = [\underline{X}_j(N+1:M), \ \underline{X}_j(N:M-1), \ldots, \underline{X}_j(1:M-N)],$$

and $\underline{X}_j(k:r)$ describes a subvector of $\underline{X}_j$ consisting of all of the components from the k-th component to the r-th component. Substituting (9) in (8) we obtain:

$$\underline{b} = \arg \min_{\underline{b} \varepsilon \Theta_b} \underline{b}^H C \, \underline{b}; \qquad C = \sum_{j=1}^{J} \tilde{X}_j^H (B^H B)^{-1} \tilde{X}_j. \tag{10}$$

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

(a) Initialization k=0, $\underline{b} = \underline{b}^{(0)}$

(b) Compute $C^{(k)}$ according to (10) using $\underline{b}^{(k)}$ to construct the matrix $B^{(k)}$.

(c) Find $\underline{b}^{(k+1)} = \arg \min_{\underline{b} \in \Theta_b} \underline{b}^H C(k) \underline{b}$

(d) Find the roots of the polynomial b(z) whose coefficients are given by $\underline{b}^{(k+1)}$.

In [1] the relation $\underline{b} = \alpha[b_0, b_1, \ldots, b_N]^H$ was incorporated in step (c) to yield a simple quadratic minimization problem. We now turn to the more practical situation of nonuniform arrays.

III. SUBLATTICE ARRAYS AND THE EM ALGORITHM

In this paper we are primarily interested 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, $\underline{l}$, of length M. The m-th component of $\underline{l}$ 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' x M matrix G is constructed by eliminating all the zero rows in diag($\underline{l}$). For example an array of three elements in positions 1,2,5 is described by $\underline{l}^T = (1,1,0,0,1)$ and

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

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

$$\underline{Y}_j = G\underline{X}_j = G(A\underline{S}_j + \underline{V}_j), \qquad j = 1,2,\ldots,J. \tag{11}$$

We refer to $\{\underline{X}_j\}$ as the (unavailable) "complete data" and to $\{\underline{Y}_j\}$ as the observed data.

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

$$\underline{Y} = F\underline{X} \tag{12}$$

where

$$F = \begin{bmatrix} G & & \\ & G & \\ & & \ddots \\ & & & G \end{bmatrix}$$

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

$$\underline{\theta} = [\underline{\theta}_1^T, \theta_2^T, \ldots, \underline{\theta}_J^T]^T,$$

where

$$\underline{\theta}_j = A\underline{S}_j \ .$$

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

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

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

$$||\underline{X}-\underline{\theta}||^2 = \sum_{j=1}^{J} ||\underline{X}_j - A\underline{S}_j||^2, \tag{14}$$

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 $\underline{Y}$ corresponding to an incomplete data set, if $f_{\underline{y}}(\underline{Y}|\underline{\theta})$ denotes the density of $\underline{y}$ given $\underline{\theta}$, the maximum likelihood estimate of $\underline{\theta}$ given $\underline{Y}$ is

$$\hat{\underline{\theta}} = \arg \max_{\underline{\theta}\varepsilon\Theta} f_{\underline{y}}(\underline{Y}|\underline{\theta}) = \arg \max_{\underline{\theta}\varepsilon\Theta} \ln\{f_{\underline{y}}(\underline{Y}|\underline{\theta})\} \tag{15}$$

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

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

$$\ln\{f_{\underline{y}}(\underline{Y}|\underline{\theta})\} = \ln\{f_{\underline{x}}(\underline{X}|\underline{\theta})\} - \ln\{f_{\underline{x}|\underline{y}}(\underline{X}|\underline{Y},\underline{\theta})\} \tag{16}$$

Taking the expectation of (14) over $\underline{x}$ given $\underline{Y}$ and under the assumption that

the parameter vector is equal to $\underline{\theta}'$, we obtain

$$L(\underline{\theta}) \triangleq \ln\{f_{\underline{Y}}(\underline{Y}|\underline{\theta})\} = Q(\underline{\theta}|\underline{\theta}') - H(\underline{\theta}|\underline{\theta}'), \tag{17}$$

where

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

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

Using Jensen's inequality it is easy to verify that

$$H(\underline{\theta}|\underline{\theta}') \leq H(\underline{\theta}'|\underline{\theta}'). \tag{18}$$

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

(a) Initialization: set p=0, and $\underline{\theta}^{(p)} = \underline{\theta}_0$.

(b) E-step: Determine $Q(\underline{\theta}|\underline{\theta}^{(p)})$.

(c) M-step: Choose $\underline{\theta}^{(p+1)}$ to be the value of $\underline{\theta}\varepsilon\Theta$ that maximizes $Q(\underline{\theta}|\underline{\theta}^{(p)})$.

(d) Check the convergence of $\underline{\theta}$. If no: p=p+1; go to (b); If yes: stop.

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

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

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

where the inequality holds due to (16) and due to the M-step.

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

$$\overset{\bullet}{\underline{X}} = E\{\underline{X}|\underline{Y}, \underline{\theta}'\} = \underline{\theta}' + F^H(FF^H)^{-1}(\underline{Y}-F\underline{\theta}') \tag{19}$$

for the conditional mean of $\underline{x}$, we find that

$$Q(\underline{\theta}|\underline{\theta}') = K - ||\overset{\bullet}{\underline{X}} - \underline{\theta}||^2/\sigma^2 \tag{20}$$

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

$$R_1 = ||\overset{\bullet}{\underline{X}} - \underline{\theta}||^2 = \sum_{j=1}^{J} ||\overset{\bullet}{\underline{X}}_j - A\underline{S}_j||^2, \tag{21}$$

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 $G^HG = \text{diag}(\underline{1})$ to rewrite (19) as

$$\underset{\sim}{\overset{\bullet}{X}}_j = \text{diag}(\bar{\underset{\sim}{1}})\underset{\sim}{\theta}'_j + G^H \underset{\sim}{Y}_j \tag{22}$$

where $\bar{\underset{\sim}{1}}$ is the complement of $\underset{\sim}{1}$ (zeros and ones are interchanged). The parameter vector $\underset{\sim}{\theta}'_j$ is simply the estimate of $A\underset{\sim}{S}_j$ obtained in the previous cycle and therefore (20) may be written also as:

$$\underset{\sim}{\overset{\bullet}{X}}^{(p+1)} = \text{diag}(\bar{\underset{\sim}{1}})\{A(A^HA)^{-1}A^H\underset{\sim}{\overset{\bullet}{X}}_j\}^{(p)} + G^H\underset{\sim}{Y}_j =$$

$$= \text{diag}(\bar{\underset{\sim}{1}})\{(I-B(B^HB)^{-1}B^H)\underset{\sim}{\overset{\bullet}{X}}_j\}^{(p)} + G^H\underset{\sim}{Y}_j,$$

using the notation of the polynomial approach. As one would expect equation (20) states that the components of $\underset{\sim}{\overset{\bullet}{X}}_j$ that correspond to existing sensors are always equal to the observed data, i.e., the corresponding components of $\underset{\sim}{Y}_j$.

The proposed EM algorithm maybe summarized as follows:

(a) Initialization: Select initial values for $\{\lambda_n\}_{n=1}^N$; find the corresponding $\underset{\sim}{b}^{(0)}$.

Compute: $A_1 = GA$; $\underset{\sim}{S}_j = (A_1^HA_1)^{-1}A_1^H\underset{\sim}{Y}_j$;

$\underset{\sim}{\overset{\bullet}{X}}_j^{(0)} = \text{diag}(\bar{\underset{\sim}{1}})A\underset{\sim}{S}_j + G^H\underset{\sim}{Y}_j$ (see (22))

Set: $p = 0$

(b) Use the minimization algorithm for uniform arrays:

(b.1) Construct $\tilde{X}_j = [\underset{\sim}{\overset{\bullet}{X}}_j^{(p)}(N+1:M),\ldots,\underset{\sim}{\overset{\bullet}{X}}_j^{(p)}(1:M-N)]$;

set $k=0$, $\underline{b}_1^{(0)} = \underline{b}^{(p)}$

(b.2) Construct B using $\underline{b}_1^{(k)}$.

Compute $C = \sum_{j=1}^{J} \tilde{X}_j^H (B^H B)^{-1} \tilde{X}_j$

(b.3) Compute $\underline{b}_1^{(k+1)} = \arg \min_{\underline{b}_1 \varepsilon \Theta_b} \underline{b}_1^H C \underline{b}_1$.

(b.4) Check convergence of $\underline{b}_1$. If no: $k=k+1$; go to (b.2).

If yes: $\underline{b}^{(p)} = \underline{b}_1^{(k+1)}$, continue.

(c) Construct B using $\underline{b}^{(p)}$

Compute:

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

(d) Check the convergence of $\underline{\hat{X}}_j$. NO: $p = p+1$, go to (b)

YES: continue.

(e) Find the roots of the polynomial $\underline{b}^{(p)}(z)$ whose coefficients are given by $\underline{b}^{(p)}$.

IV. EXAMPLES

To illustrate the behavior of the algorithm, let use consider two examples:

Example 1: Consider a uniform linear array of 6 sensors separated by half a wavelength of the actual narrowband source signals. Now, assume that the two middle sensors are missing (i.e., $\underline{1}^T = (1\ 1\ 0\ 0\ 1\ 1)$); this is the optimal configuration for bearing estimation when the given aperture is 2.5 wavelengths and the number of sensors is limited to 4.

The sources are two narrowband emitters located in the far field of the array. One source is located at a bearing of 10 degree, and the second source is located at a bearing of 25 degrees. We generated only 10 independent samples with a SNR of 30 dB. The initial guess was $\gamma_1^{(0)} = 3^o$, $\gamma_2^{(0)} = 17^o$. The algorithm converged to within one degree of the right result in 8 iterations, as shown in Table 1.

Example 2: Consider Example 1 where the array is reconfigured so that $\underline{1}^T = (1\ 0\ 1\ 0\ 0\ 1)$. Note that only 3 sensors are used and they are separated by one wavelength and 1.5 wavelengths. Nevertheless, the algorithm converged to within one degree of the right result in only 7 iterations as shown in Table 2. The initial guess was $\gamma_1^{(0)} = 3^o$, $\gamma_2^{(0)} = 35^o$.

V. SUMMARY

We have proposed a novel EM algorithm for the estimation of superimposed signals observed by nonuniform arrays. The algorithm is efficient and provides accurate results even when the number of samples is samll and the sensors are separated by more than half a wavelength.

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. 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 the MLM, MEM or the MUSIC techniques (see [7] for a review of these methods).

References

[1] Y. Bresler and A. Macovski, "Exact Maximum Likelihood Parameter Estimation of Superimposed Exponential Signals in Noise," IEEE Trans. on Acoustics, Speech and Signal Processing", Vol. 34, No. 5, pp. 1081-1089. October, 1986.

[2] R. Kumaresan and A.K. Shaw, "High Resolution Bearing Estimation without Eigen Decomposition," International Conference on Acoustics Speech and Signal Processing (ICASSP), Tampa, Florida 1985, pp. 576-579.

[3] R. Kumaresan, L.L. Scharf, and A.K. Shaw, "An Algorithm for Pole-Zero Modeling and Spectral Analysis," IEEE Trans. on Acoustics, Speech, and Signal Processing, Vol. ASSP-34, No. 3, pp. 637-640, June 1986.

[4] G.C. Carter, "Coherence and Time Delay Estimation," Proceeding of the IEEE, Vol. 75, No. 2, pp. 236-255, February 1987.

[5] M. Feder. and E. Weinstein, "Parameter Estimation of Superimposed Signals Using the EM Algorithm", preprint.

[6] A.P. Dempster, N.M. Laird, D.B. Rubin, "Maximum Likelihood from Incomplete Data via the EM Algorithm", Journal of the Royal Statistical Society, ser. B39, pp. 1-38, 1977.

[7] M. Wax, "Detection and Estimation of Superimposed Signals," Ph.D. Dissertation, Stanford University, Stanford, CA, 1985.

Iterations No.	$\hat{\gamma}_1$ degrees	$\hat{\gamma}_2$ degrees
0	3.00	17.0
1	6.15	19.38
2	7.29	20.49
3	8.16	21.47
4	8.78	22.30
5	9.23	22.95
6	9.55	23.46
7	9.77	23.83
8	9.93	24.12
9	10.04	24.31
10	10.12	24.47

Table 1: Evolution of the algorithm for $\underline{1}^T = (1\ 1\ 0\ 0\ 1\ 1)$.

Iterations No.	$\hat{\gamma}_1$ degrees	$\hat{\gamma}_2$ degrees
0	3.00	35.00
1	-0.01	18.13
2	3.46	18.27
3	7.18	20.16
4	8.74	21.90
5	9.39	23.01
6	9.68	23.69
7	9.84	24.10
8	9.92	24.35
9	9.96	24.51
10	9.99	24.61

Table 2: Evolution of the algorithm for $\underline{1}^T = (1\ 0\ 1\ 0\ 0\ 1)$