J. Indian Inst. Sci., Sept.–Oct. 1987, 67, pp. 341–350.

Estimation of Signal Parameters *via* Rotational Invariance Techniques—*ESPRIT*†*

R. Roy, A. Paulraj and T. Kailath
Information Systems Laboratory, Stanford University, Stanford, Ca. 94305, USA.

Abstract

A new approach to the general problem of signal parameter estimation is described. Though the technique (*ESPRIT*) is discussed in the context of direction-of-arrival estimation, it can be applied to a wide variety of problems including spectral estimation. *ESPRIT* exploits an underlying *rotational invariance* among signal subspaces induced by an array of sensors with a *translational invariance* structure (*e.g.*, pairwise matched and co-directional antenna element doublets) and has several advantages over earlier techniques such as MUSIC including improved performance, reduced computational load, freedom from array characterization/calibration, and reduced sensitivity to array perturbations. Results of computer simulations carried out to evaluate the new algorithm are presented.

Key words: Signal parameter estimation, spectral estimation, rotational invariance, translational invariance, array signal processing.

1. Introduction

High-resolution parameter estimation is important in many applications including direction-finding (DF) sensor systems. Many methods have been proposed such as the maximum likelihood (ML) method of Capon, the maximum entropy (ME) method of Burg, and conventional (delay-and-sum) beamforming. These methods have been overshadowed recently by the signal subspace method (MUSIC) developed by Schmidt[1]. Among all the methods proposed to date, only MUSIC is known to yield unbiased and efficient estimates as the amount of information (*i.e.*, the amount of data or the signal-to-noise ratio (SNR) increases without bound, though practically the amount of residual bias in most algorithms becomes insignificant as the information-to-noise ratio (INR) becomes large (*cf.* ref. 2 for extensive simulation results).

The MUSIC algorithm derives its properties from exploitation of the underlying data model of finite (low) rank signals (*e.g.*, spatially coherent wavefronts) in additive noise, a

†This work was supported in part by the JSEP, managed by the U.S. Army Research Office under Contract DAAG29–83–K–0028, and the SDI/IST Program managed by the Office of Naval Research under Contract N00014–85–K–0550.

*First presented at the Platinum Jubilee Conference on Systems and Signal Processing held at the Indian Institute of Science, Bangalore, India, during December 11–13, 1986.

situation typical in many sensor-array environments. The MUSIC algorithm first determines the *signal subspace* from the array measurements. Intersections between the estimated signal subspace and the array manifold (the set of all possible array responses as functions of the parameter(s) to be estimated) are then sought. This search is typically carried out by computing a weighted norm (Hermitian form) using the direction vectors for each angle of interest and a kernel obtained from the noise eigen vectors of the data covariance matrix. Essentially the same computation also underlies the earlier methods (*cf.* ML, ME) with the only difference being in the choice of norms (kernels).

In this paper, a new approach (*ESPRIT*) to the signal parameter estimation problem is described[3,4]. *ESPRIT* is similar to MUSIC in that it correctly exploits the underlying data model, while manifesting significant advantages over MUSIC. Moreover, *ESPRIT* does not require detailed knowledge of the array geometry and element characteristics as do other techniques, eliminating the need to calibrate the array thereby eliminating the need for the associated storage of the array manifold. *ESPRIT* is also computationally much less complex because it does not employ the search procedure inherent in other algorithms, and it manifests improved performance over the MUSIC algorithm in terms of bias and resolution. *ESPRIT* is also less sensitive to errors in sensor positions (array geometry), and in sensor gains /phases than the MUSIC algorithm, and provides a simple solution to the signal copy problem, where the objective is to extract a particular signal of interest while rejecting all others. Finally, *ESPRIT* can *simultaneously* estimate the number of sources and the parameters (*e.g.*, DOAs), unlike MUSIC where an estimate of the number of sources present is required before source parameter estimates can be obtained. However, in MUSIC there are essentially no restrictions on the array manifold, other than the design requirement to eliminate ambiguities, whereas *ESPRIT* requires the array manifold to possess a *displacement invariance*. It is precisely this symmetry/invariance which leads to the simple solution provided by *ESPRIT*, though in this sense *ESPRIT* is not completely general.

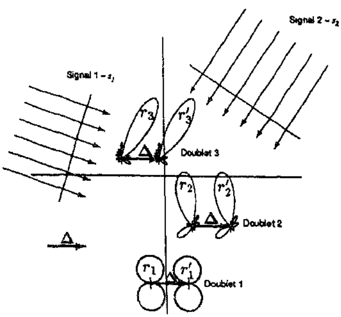

FIG. 1. Sensor-array geometry for multiple source DOA estimation using *ESPRIT*.

2. Problem formulation

The basic problem under consideration is that of estimation of parameters of finite dimensional signal processes given measurements from an array of sensors. This general problem appears in many different fields including radio astronomy, geophysics, sonar signal processing, electronic surveillance, structural (vibration) analysis, and spectral analysis. In order to simplify the description of the basic ideas behind *ESPRIT*, the ensuing discussion is couched in terms of the problem of multiple source one-dimensional DOA estimation of narrow-band emitters from data collected by an array of sensors.

Consider a planar array of *arbitrary* geometry composed of m *matched* sensor doublets whose elements are translationally separated by a *known constant displacement vector* (fig. 1). The element characteristics such as element gain and phase pattern, polarization sensitivity, etc., may be arbitrary for each doublet as long as the elements are *pairwise identical*. Assume there are $d \leq m$ narrow-band stationary zero-mean sources centered at frequency ω_0, and located sufficiently far from the array such that in homogeneous isotropic transmission media, the wavefronts impinging on the array are planar. Additive noise is present at all the $2m$ sensors and is assumed to be a stationary zero-mean random process that is uncorrelated from sensor to sensor.

To exploit the translational invariance property of the sensor array, consider the array as being comprised of two identical subarrays, X and Y, displaced from each other by a known displacement vector. The signals received at the ith doublet can then be expressed as:

$$x_i(t) = \sum_{k=1}^{d} s_k(t) a_i(\theta_k) + n_{xi}(t), \tag{1}$$
$$y_i(t) = \sum_{k=1}^{d} s_k(t) e^{j\omega_0 \Delta \sin\theta_k / c} a_i(\theta_k) + n_{yi}(t);$$

where $s_k(\cdot)$ is the kth signal (wavefront) as received at sensor 1 of the X subarray, θ_k is the DOA of the kth source relative to Δ (the displacement vector between the two arrays), $a_i(\theta_k)$ is the response of the ith sensor of either subarray relative to its response at sensor 1 of the same subarray when a single wavefront impinges at an angle θ_k. c is the speed of propagation in the transmission medium, and $n_{xi}(\cdot)$ and $n_{yi}(\cdot)$ are the additive noises at the elements in the ith doublet for subarrays X and Y respectively.

Combining the outputs of each of the sensors in the two subarrays, the received data vectors can be written as follows:

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) + \mathbf{n}_x(t), \quad \mathbf{y}(t) = \mathbf{A}\boldsymbol{\Phi}\mathbf{s}(t) + \mathbf{n}_y(t); \tag{2}$$

where $\mathbf{x}^T(t) = [x_1(t), \ldots, x_m(t)]$, and $\mathbf{y}(t)$, $\mathbf{n}_x(t)$ and $\mathbf{n}_y(t)$ are similarly defined. The vector $\mathbf{s}(t)$ is a d-vector of impinging signals (wavefronts) as observed at the reference sensor of subarray X. The $m \times d$ matrix $\mathbf{A}$ is the *direction matrix* whose columns $\{\mathbf{a}(\theta_k), k = 1, \ldots, d\}$ are the signal *direction vectors* for the d wavefronts. The matrix Φ is a diagonal $d \times d$ matrix of the phase delays, $\phi_k = \omega_0 \Delta \sin \theta_k / c$, between the doublet sensors for the d wavefronts.

Defining the data received from the entire array as $\mathbf{z}(t) = [\mathbf{x}^T(t), \mathbf{y}^T(t)]^T$ and assuming $\mathbf{z}(t)$ is zero mean, the auto-covariance of the data is $\mathbf{R}_{zz} = E[\mathbf{z}(t)\mathbf{z}^*(t)] = C_{zz} + \sigma^2 \Sigma$, where $\sigma^2 \Sigma$ is the noise covariance,

$$C_{zz} = \begin{bmatrix} C_{xx} & C_{xy} \\ C_{yx} & C_{yy} \end{bmatrix} = \begin{bmatrix} \mathbf{ASA}^* & \mathbf{AS\Phi^*A}^* \\ \mathbf{A\Phi SA}^* & \mathbf{A\Phi S\Phi^* A}^* \end{bmatrix}, \tag{3}$$

and $\mathbf{S} = E[\mathbf{s}(t)\mathbf{s}^*(t)]$, assumed to be nonsingular. Now the problem can be stated as follows: *Given measurements $\mathbf{z}(t)$ and making no assumptions about the array geometry, element characterisics, DOAs, noise powers, or the signal (wavefront) correlation, estimate the signal DOAs.*

3. Invariant subspace approach

The basic idea behind the new technique is to exploit the rotational invariance of the underlying signal subspaces induced by the translational invariance of the sensor array. The algebraic details behind this geometric interpretation of the *ESPRIT* algorithm are embodied in the following theorem.

Theorem (ESPRIT): Define Γ *as the generalized eigenvalue matrix associated with the matrix pencil* $\{C_{xx}, C_{xy}\}$. *For nonsingular* $\mathbf{S}$, *the matrices* Φ *and* Γ *are related by:*

$$\Gamma = \begin{bmatrix} \Phi & 0 \\ 0 & 0 \end{bmatrix} \tag{4}$$

to within a permutation of the elements of Φ.

Proof: Using the definitions of $C_{xx} = \mathbf{ASA}^*$ and $C_{xy} = \mathbf{AS\Phi^*A}^*$, the matrix pencil can be written as follows:

$$C_{xx} - \gamma C_{xy} = \mathbf{ASA}^* - \gamma \mathbf{AS\Phi^* A}^* = \mathbf{AS}(\mathbf{I} - \gamma \Phi^*)\mathbf{A}^*. \tag{5}$$

By inspection, the column space of both $\mathbf{ASA}^*$ and $\mathbf{AS\Phi^*A}^*$ are identical, and in general $\rho(\mathbf{ASA}^*) - \gamma \mathbf{AS\Phi^*A}^*) = d$ where $\rho(\cdot)$ denotes rank. However, when $\gamma = \gamma_i = e^{j\omega_0 \Delta \sin \theta_i / c}$, the ith row of $(\mathbf{I} - \gamma_i \Phi^*)$ is zero, and $\rho(\mathbf{I} - \gamma_i \Phi^*) = d-1$. Consequently, the pencil $C_{xx} - \gamma C_{xy}$ decreases in rank to $d-1$. By definition, these are the *generalized eigenvalues* (GEs) of the matrix pair $\{C_{xx}, C_{xy}\}$. Since both matrices span the same subspace, the common null space GEs are zero by definition. Thus, d GEs lie on the unit circle and are equal to the diagonal elements of Φ, and the remaining $m-d$ GEs are at the origin. Once Φ is known, the DOAs can be calculated using $\theta_k = \arcsin\{c\phi_k/\omega_0\Delta\}$, and the proof of the theorem is complete.

In order to obtain $\{C_{xx}, C_{xy}\}$, *i.e.*, C_{zz}, from the data covariance $\mathbf{R}_{zz}$ and knowledge of the normalized noise covariance Σ, the noise power σ^2 must be calculated. Defining $\bar{\mathbf{A}} = [\mathbf{A}^T, (\mathbf{A}\Phi)^T$, $\mathbf{R}_{zz} = \bar{\mathbf{A}}\mathbf{S}\bar{\mathbf{A}}^* + \sigma^2\Sigma$. From linear algebra, $\rho(\mathbf{ASA}^*) = \min(\rho(\bar{\mathbf{A}}), \rho(\mathbf{S}))$. Assuming there are no array ambiguities, the columns of $\bar{\mathbf{A}}$ are linearly independent; hence $\rho(\bar{\mathbf{A}}) = d$. Since $\rho(\mathbf{S}) = d$, $\rho(\bar{\mathbf{A}}\mathbf{S}\bar{\mathbf{A}}^*) = d$. Thus, $det(\bar{\mathbf{A}}\mathbf{S}\bar{\mathbf{A}}^*) = det(\mathbf{R}_{zz} - \sigma^2\Sigma) = 0$. This equation is only satisfied it σ^2 is the minimum eigenvalue $\lambda_{\min}$ of $\mathbf{R}_{zz}$ in the *metric* Σ, *i.e.*, the minimum *generalized eigenvalue* of the matrix pair $\{\mathbf{R}_{zz}, \Sigma\}$, since $\bar{\mathbf{A}}\mathbf{S}\bar{\mathbf{A}}^*$ is Hermitian and therefore non-negative definite. Consequently, $\bar{\mathbf{A}}\mathbf{S}\bar{\mathbf{A}}^* = \mathbf{R}_{zz} - \lambda_{\min}\Sigma$. Note that there will be $m-d$ minimum generalized eigenvalues, all equal to σ^2 since $\rho(\bar{\mathbf{A}}\mathbf{S}\bar{\mathbf{A}}^*) = d$.

3.1 *Signal copy*

Signal copy refers to the weighted combination of sensor measurements such that the single output contains the desired signal while completely rejecting the other $d-1$ signals. *ESPRIT* provides an elegant solution to the problem of estimating the *optimal* signal copy weight vector. Let e_i be the generalized eigenvector (GEV) corresponding to the GE γ_i. By definition:

$$\mathbf{AS}(\mathbf{I} - \gamma_i\Phi^*)\mathbf{A}^*\mathrm{e}_i = 0. \tag{6}$$

Since the column space of $\mathbf{AS}(\mathbf{I}-\gamma_i\Phi^*)\mathbf{A}^*$ is same as the subspace spanned by the vectors $\{\mathbf{a}_j, j \neq i\}$, it follows that e_i is orthogonal to all direction vectors except $\mathbf{a}_i$[†]. Thus, e_i is (proportional to) the desired weight vector for signal copy of the ith signal, rejecting signals from the remaining $d-1$ directions;

$$\mathbf{w}_i^{SC} \propto \mathrm{e}_i. \tag{7}$$

Note that this is the *optimal* copy vector in the sense defined above even when the signals are correlated.

4. Subspace rotation algorithm

The *ESPRIT* theorem is based on knowledge of $\mathbf{R}_{zz}$, a covariance matrix which in practice is not known, and which must be estimated. Due to errors in estimating $\mathbf{R}_{zz}$, from finite data as well as errors introduced during the subsequent finite precision computations, the relations in the *ESPRIT* theorem will not be satisfied exactly. A procedure which is not globally optimal, but which utilizes some well-established, stepwise-optimal techniques to deal with such issues is outlined.

The key steps of the covariance matrix formulation of *ESPRIT* are:

1. Find the $2m \times 2m$ sample covariance matrix $\hat{\mathbf{R}}_{zz}$ of the complete $2m$ sensor array, then estimate the number of sources $\hat{d}$ and the noise variance $\hat{\sigma}^2$ as the number of minimum repeated generalized eigenvalues of the matrix pair $\{\hat{\mathbf{R}}_{zz}, \Sigma\}$.

[†]For convenience in the ensuing derivations, the shorthand notation $\mathbf{a}(\theta_i) = \mathbf{a}_i$ is adopted.

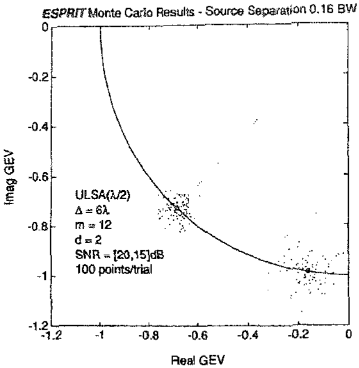

FIG. 2. *ESPRIT* GEs for $\Delta = 6\lambda$.

2. Compute a rank $\hat{d}$ approximation to $\hat{\mathbf{R}}_{zz} - \hat{\sigma}^2\Sigma$, denoting the result $\hat{\mathbf{C}}_{zz}$.
3. Use the d generalized eigenvalues (GEs) of the matrix pair $\{\hat{\mathbf{C}}_{xx}, \hat{\mathbf{C}}_{xy}\}$ and/or $\{\hat{\mathbf{C}}_{yx}, \hat{\mathbf{C}}_{yy}\}$ that lie closest to the unit circle to estimate Φ.

The rank d approximation to $\hat{\mathbf{R}}_{zz} - \sigma^2\Sigma$ is obtained using spectral decomposition, *i.e.*, $\hat{\mathbf{C}}_{zz} = \sum_{i=1}^{d} (\hat{\lambda}_i - \hat{\sigma}^2)\hat{e}_i\hat{e}_i^*$, where $\{e_1, e_2, \ldots, e_m\}$ are the ordered eigenvectors of $\hat{\mathbf{R}}_{zz} - \hat{\sigma}^2\Sigma$.

5. Simulation results

Simulations were carried out to investigate the comparative performance of the *ESPRIT* and MUSIC algorithms under similar conditions. The MUSIC algorithm was chosen as the benchmark due to its superior bias, error variance and resolution performance as compared to the more traditional methods (MLM, MEM, AAR, etc.). Two scenarios were used in this analysis in order to investigate the relative performance of *ESPRIT* and MUSIC; one in which the standard MUSIC *spectrum* fails to resolve the two sources present, and one in which it resolves the two sources with high probability.

The first scenario consisted of two planar wavefronts impinging on a 12-element array consisting of two six-element uniform ($\lambda/2$) linear subarrays which for convenience were assumed to be collinear and separated by $\Delta = 6\lambda$. Two planar uncorrelated signal wavefronts impinged on the array at angles of 26 and 27°, with SNRs of 20 and 15 dB relative to the additive noise. Covariance estimates were computed from 100 snapshots of data, and 100 trials were run using independent data sets. Figure 2 shows the *ESPRIT* results. The two sources 1° apart* are easily resolved. The sample means and sigmas of

*For $\Delta = 6\lambda$, $BW_{3dB} = 6°$, and $\delta\theta = 0.16BW$.

the *ESPRIT* estimates of $\sin(\theta)$ were $(0.4381 \pm 0.0011, 0.4540 \pm 0.0021)$ which compare favorably with the actual values (0.4384, 0.4540). Note that the *ESPRIT* algorithm did not require knowledge of the array geometry, nor did it exploit the uniform linear structure of the subarrays. Figure 3 contains MUSIC spectral estimates obtained using the sample covariances from the first 20 trials. In all cases, the number of sources was assumed known ($d = 2$), and the signal and noise subspaces estimated appropriately. The conventional MUSIC *spectrum* is given by $P(\theta) = [\mathbf{a}^*(\theta)\mathbf{E}_n\mathbf{E}_n^*\mathbf{a}(\theta)]^{-1}$, where $\mathbf{E}_n$ denotes the estimated noise subspace. In a majority of the trials, two *spectral peaks* were not resolvable in the search region [25°,28°].

To investigate the relative performance of *ESPRIT* and MUSIC in a situation where MUSIC is clearly able to resolve the sources, the scenario was changed. An eight-element uniform linear array with $\lambda/4$ spacing** was used. Two sources with SNRs of 20 dB each referenced to the additive noise were located at 24 and 28°, and 5000 Monte Carlo trials were run. A histogram of the *ESPRIT* results (using overlapping seven-element subarrays) is shown in fig. 4. The sample means and sigmas of the resulting angle estimates are (23.99° ± 0.30°, 28.01° ± 0.27°). Gaussian curves with these means and variances are also included in fig. 4. The *ESPRIT* estimates are clearly

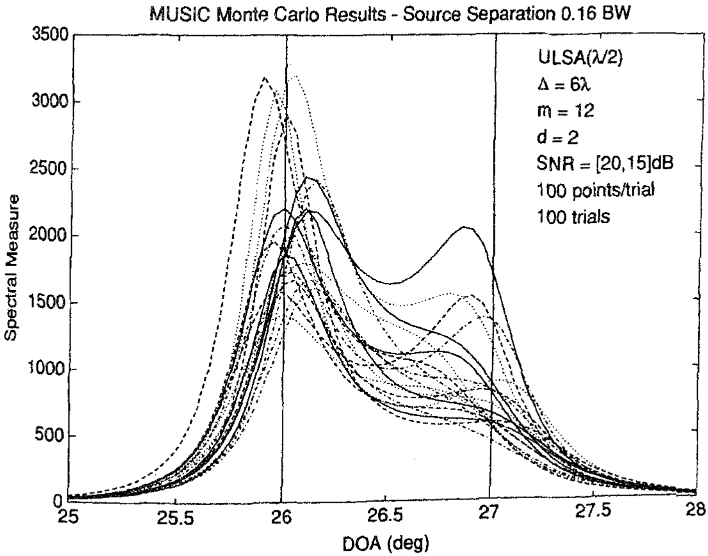

FIG. 3. MUSIC spectra for $\Delta = 6\lambda$.

**For this eight-element array steered to 26°, BW_{3dB} = 30°.

unbiased. The corresponding results for the MUSIC algorithm are given in fig. 5. The sample means and sigmas are $(24.15° \pm 0.30°, 27.86° \pm 0.27°)$. A significant bias of approximately one-half sigma is manifest in the MUSIC results. However, both *ESPRIT* and MUSIC have the same variance.

Preliminary comparisons of the sensitivity of *ESPRIT* and MUSIC to errors in sensor position, gain and phase have also been made. In these simulations, the nominal array structure had the desired *ESPRIT* displacement structure though no information other than the nominal displacement vector was used by the algorithm. On the other hand, MUSIC had a complete characterization (array manifold) of the nominal array. Sensor positions, phases and gains were perturbed randomly $N(0, \sigma^2)$ in each trial and data from the perturbed array used to obtain DOA estimates. The nominal scenario was the same as the previous case with two sources (20dB SNR) at 24 and 28°. The sigmas for the relative position, gain, and phase errors were $0.01\,\Delta$, $0.1\,\mathrm{d}B$, and 2° respectively and 5000 independent trials were run. The conventional MUSIC spectrum proved incapable of resolving the sources in 40% of the trials making a direct comparison of *ESPRIT* and MUSIC results untenable. However, the *ESPRIT* results (with no failures by definition) were unbiased with a sigma of 0.7°.

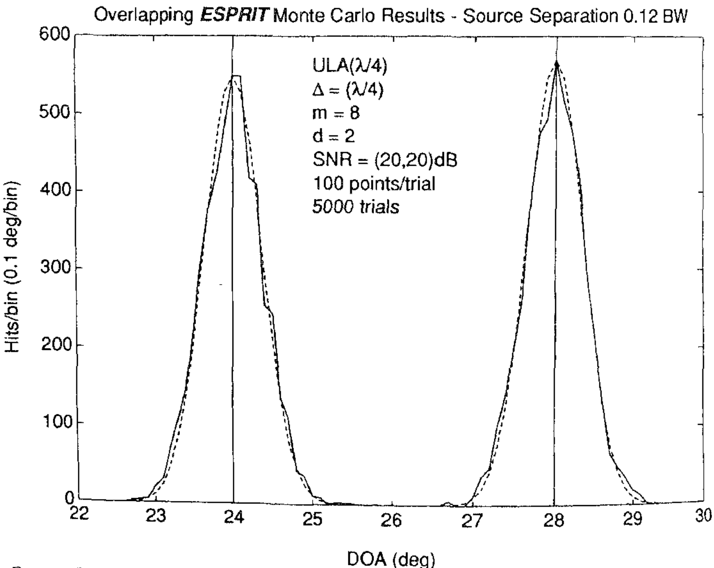

FIG. 4. Overlapping covariance *ESPRIT* results.

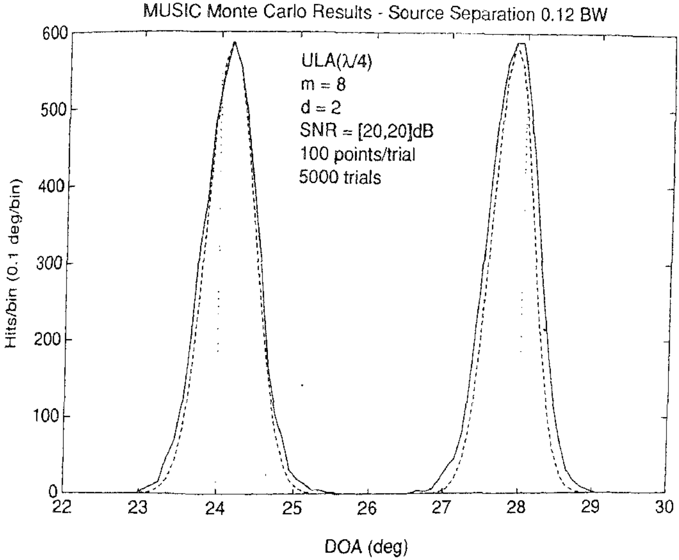

FIG. 5. Histogram of MUSIC results.

6. Concluding remarks

In this paper, a new approach to signal parameter estimation using data received by an array having a translational invariance structure has been described. The method shows considerable promise and has significant advantages over previous algorithms including improved performance, reduced computational load, indifference to array calibration (thus eliminating the associated storage) and lower sensitivity to array perturbations. For example, with a 20-element array covering an arc of 2 radians with a one milliradian resolution in both azimuth and elevation, *ESPRIT* has a computational advantage on the order of 10^5 over MUSIC. Furthermore, while MUSIC needs about 20 megabytes of storage for the array manifold (using 16-bit words), *ESPRIT* requires no storage. The fact that array calibration is not necessary is very attractive in applications such as space antennas, sonobuoys, etc., where the array geometry may not be known and may be slowly varying with time. In addition, *ESPRIT* provides a simple solution to the signal copy problem. Thus, the new technique has the potential to make high-resolution DOA estimation, signal copy, etc., feasible in the sense of making it simpler and cheaper to implement in many applications.

References

1. SCHMIDT, R.O. *A signal subspace approach to multiple emitter location and spectral estimation*, Ph.D. Thesis, Stanford University, Stanford, Ca., 1981.

2. BARABELL, A.J., CAPON, J., DELONG, D.F., JOHNSON, J.R. AND SENNE, K. *Performance comparison of superresolution array processing algorithms*, Technical Report TST-72, Lincoln Laboratory, M.I.T., 1984.

3. PAULRAJ, A., ROY, R. AND KAILATH, T. Patent application: *Methods and means for signal reception and parameter estimation*, Stanford University, Stanford, Ca., 1985.

4. ROY, R., PAULRAJ, A. AND KAILATH, T. *ESPRIT*—a subspace rotation approach to estimation of parameters of cisoids in noise, *IEEE Trans.*, 1986, **ASSP-34**, 1340–1342.

Note added in proof

Since the submission of the paper, the ESPRIT algorithm has been improved[1] and is now patented too[2].

1. ROY, R. H. *ESPRIT—Estimation of Signal Parameters via Rotational Invariance Techniques*, Ph.D. thesis, Stanford University, Stanford, Ca., 1987.

2. ROY, R., PAULRAJ, A. AND KAILATH, T. *Method for estimating signal source locations and signal parameters using an array of signal sensor pairs*, United States Patent, Patent No. 4,750,147, June 7, 1988.