

Positive-Definite Toeplitz Completion in DOA Estimation for Nonuniform Linear Antenna Arrays—Part I: Fully Augmentable Arrays

Yuri I. Abramovich, *Member, IEEE*, Douglas A. Gray, *Member, IEEE*,
Alexei Y. Gorokhov, *Member, IEEE*, and Nicholas K. Spencer

Abstract—This paper considers the problem of direction-of-arrival (DOA) estimation for multiple uncorrelated plane waves incident on so-called “fully augmentable” sparse linear arrays. In situations where a decision is made on the number of existing signal sources (m) prior to the estimation stage, we investigate the conditions under which DOA estimation accuracy is effective (in the maximum-likelihood sense).

In the case where m is less than the number of antenna sensors (M), a new approach called “MUSIC-maximum-entropy equalization” is proposed to improve DOA estimation performance in the “preasymptotic region” of finite sample size (N) and signal-to-noise ratio. A full-sized positive definite (p.d.) Toeplitz matrix is constructed from the $M \times M$ direct data covariance matrix, and then, alternating projections are applied to find a p.d. Toeplitz matrix with m -variate signal eigensubspace (“signal subspace truncation”).

When $m \geq M$, Cramér–Rao bound analysis suggests that the minimal useful sample size N is rather large, even for arbitrarily strong signals. It is demonstrated that the well-known direct augmentation approach (DAA) cannot approach the accuracy of the corresponding Cramér–Rao bound, even asymptotically (as $N \rightarrow \infty$) and, therefore, needs to be improved. We present a new estimation method whereby signal subspace truncation of the DAA augmented matrix is used for initialization and is followed by a local maximum-likelihood optimization routine. The accuracy of this method is demonstrated to be asymptotically optimal for the various superior scenarios ($m \geq M$) presented.

Index Terms—Direction of arrival estimation, nonuniformly spaced arrays, linear arrays, Toeplitz matrices.

I. INTRODUCTION

NONUNIFORMLY-spaced antenna arrays, and linear sparse arrays in particular, attract considerable attention, especially when the number of receiving channels is limited [1]–[5].

A *fully augmentable* M -element nonuniform linear array (NLA) has a full co-array (i.e., the set of $M(M-1)/2$ intersensor separations has no missing lags) and is identical

Manuscript received January 13, 1996; revised February 19, 1998. Some results of this paper were presented at the ICASSP Conference, Atlanta, GA, 1996. This work was supported in part by the INTAS SASPARC Grant. The associate editor coordinating the review of this paper and approving it for publication was Dr. Petar M. Djuric.

Y. I. Abramovich, D. A. Gray, and N. K. Spencer are with the Cooperative Research Centre for Sensor Signal and Information Processing (CSSIP), Adelaide, Australia.

A. Y. Gorokhov is with the Département Signal, Télécom Paris, Paris, France.

Publisher Item Identifier S 1053-587X(98)05958-3.

to that of the corresponding M_α -element uniform linear array (ULA). Here, $(M_\alpha - 1)$ is the array aperture, which is usually measured in half-wavelength units. (In the sequel to this paper (Part II), we investigate completion methods for *partially augmentable* arrays that have some number of missing lags.) Consequently, fully augmentable arrays allow direction-of-arrival (DOA) estimation in the so-called “superior case”

$$M \leq m < M_\alpha \quad (1)$$

where the standard method utilized is the direct augmentation approach (DAA) proposed by Pillai *et al.* [6], [7]. Of course, for the “conventional case”

$$1 \leq m < M \quad (2)$$

high-resolution signal eigensubspace techniques, such as MUSIC and its variants, can be applied directly. Moreover, for uncorrelated signals with linearly independent manifold (“steering”) vectors, these methods are known to be asymptotically optimal in the maximum-likelihood (ML) sense [8] and cannot be significantly improved upon in the asymptotic domain, where the signal-to-noise ratio (SNR) and sample volume N are sufficiently large. However, in most practical applications, we are interested in defining threshold parameters (such as the minimum SNR and N) that ensure that the desired DOA estimation accuracy is met.

Naturally, if nothing is known *a priori* about the signal environment, then the overall problem is a detection or joint detection-estimation problem, rather than the isolated DOA estimation problem. Nevertheless, there are many applications where the primary detection problem is followed by a sequence of estimation procedures for the evolving DOA’s. This primary detection stage may be performed on the basis of MDL or Akaike criteria [9] or their latest variants [10], [11] in all conventional cases. For the superior case, these algorithms, which are applied to the augmented covariance matrix, should be modified appropriately. This idea can be justified for the asymptotic domain ($N \rightarrow \infty$), where the augmented matrix is treated as Gaussian distributed. Regardless of the sample volume that is necessary for the detection stage, a decision scheme with a sequence of ongoing estimation procedures creates a strong desire to explore the limit accuracy conditions for the estimation problem alone.

In the conventional case, Cramér–Rao bound (CRB) analysis suggests that a high ML limit accuracy is attainable even

for a relatively low SNR and sample volume. Nevertheless, it has long been known that in spite of asymptotic optimality, the performance of MUSIC and its variants is far from the Cramér–Rao bound prediction in the so-called “preasymptotic domain,” that is, where the signal spatial frequency separation, SNR, and N are all comparatively small. The misidentification of peaks in the MUSIC pseudo-spectrum occurs with high probability under such preasymptotic conditions, and here, the MUSIC-type methods can no longer be considered to be the best available. Techniques that in some way involve array geometry claim to achieve much better performance in the preasymptotic domain [12]. The way in which array geometry can be used for DOA estimation is not unique and obviously depends on this geometry. Correspondingly, the fundamental property of the class of fully augmentable arrays should be used to improve the preasymptotic behavior for conventional scenarios.

For superior scenarios, on the other hand, CRB analysis indicates that a reasonable DOA estimation accuracy can only be achieved for large N , almost irrespective of the source signal powers. Even when the sample volume is very large ($N \gg 1$), the problem of estimation accuracy improvement still is important since, even in the asymptotic domain, the DAA is far from being effective in the ML sense. This fact is illustrated in this paper and stimulates a search for some new approaches to DOA estimation for a superior number of signals. Once again, all available *a priori* information should be used to achieve this goal.

Thus, the problem of DOA estimation performance improvement has significantly different features for the conventional ($1 \leq m < M$) and superior ($M \leq m < M_\alpha$) cases when dealing with fully augmentable nonuniform linear arrays. In the conventional case, we only need to improve the preasymptotic behavior by reducing in some way the probability of incorrect identification. In the superior case, we need to advance DOA estimation accuracy beyond the standard DAA limit in the asymptotic domain.

Accordingly, this paper is organized as follows. Section II summarizes background material: Cramér–Rao bounds are introduced for both conventional and superior scenarios; preasymptotic conditions for MUSIC are identified for the conventional case in order to facilitate comparison with the new techniques. Section III describes the new conventional-case estimation technique, comprising “MUSIC-maximum-entropy equalization” (MMEE) followed by “signal subspace truncation” (SST), which is designed to improve DOA estimation accuracy in the preasymptotic domain. Section IV is devoted to the superior situation, where the DAA is significantly modified to achieve asymptotically optimal results in the ML sense. We call this the ATML algorithm. A summary appears in Section V, and necessary mathematical details are attached in the Appendices.

II. BACKGROUND

Consider an M -element sparse linear array with sensors situated at positions d_j ($j = 1, \dots, M$)

$$\mathbf{d} = [d_1 \equiv 0, d_2, d_3, \dots, d_M =: M_\alpha - 1]. \quad (3)$$

Let d be the greatest common divisor of the set of interelement distances

$$\mathcal{D} = \{d_j - d_k \mid j, k = 1, \dots, M; j > k\}. \quad (4)$$

To avoid trivial ambiguity, d is usually set equal to a half wavelength of the incident radiation ($\lambda/2$). Fully augmentable arrays have the property that all intermediate distances are realized in the set of lags, i.e., given the sequence of natural numbers $\kappa = 1, \dots, M_\alpha - 1$, we have $\kappa d \in \mathcal{D}$.

We assume Gaussian processes are observed as a combination of m uncorrelated plane wave signals with DOA's $\boldsymbol{\theta} \equiv [\theta_1, \dots, \theta_m]^T$, powers $P \equiv \text{diag}[p_1, \dots, p_m]$, and white noise of power p_0

$$\mathbf{y}(t) = \mathbf{S}\mathbf{x}(t) + \boldsymbol{\eta}(t), \quad \text{for } t = 1, \dots, N \quad (5)$$

where $\mathbf{y}(t) \in \mathcal{C}^{M \times 1}$ is the vector of observed sensor outputs (the “snapshot”), $\mathbf{x}(t) \in \mathcal{C}^{m \times 1}$ is the vector of Gaussian signal amplitudes

$$\mathcal{E}\{\mathbf{x}(t_1)\mathbf{x}^H(t_2)\} = \begin{cases} P & \text{for } t_1 = t_2 \\ 0 & \text{for } t_1 \neq t_2 \end{cases} \quad (6)$$

and $\boldsymbol{\eta}(t) \in \mathcal{C}^{M \times 1}$ is additive white Gaussian noise. Here, as usual, $\mathcal{C}^{p \times q}$ is the space of $p \times q$ complex-valued matrices, and $\mathcal{E}\{\cdot\}$ is the expectation operator. The signal manifold matrix is $S \equiv [\mathbf{s}(w_1), \dots, \mathbf{s}(w_m)] \in \mathcal{C}^{M \times m}$, where

$$\begin{aligned} \mathbf{s}(w) &= \left[\exp\left(i2\pi \frac{d}{\lambda} d_j \sin \theta\right) \right]_{j=1, M} \\ &= [1, \exp(i\pi d_2 w), \dots, \exp(i\pi d_M w)]^T \end{aligned} \quad (7)$$

is the “steering vector” associated with the DOA w and where $w \equiv 2(d/\lambda) \sin \theta = \sin \theta$ in this case.

In this study, we assume that the manifold matrix S is of full rank, i.e., that all the steering vectors are linearly independent. A special investigation of degenerate scenarios with linearly dependent steering vectors has been conducted elsewhere [13], [14].

Given N independent snapshots, the sufficient statistic for DOA estimation is the direct data covariance (DDC) matrix

$$\hat{R} = \frac{1}{N} \sum_{t=1}^N \mathbf{y}(t)\mathbf{y}^H(t). \quad (8)$$

Under our assumptions, \hat{R} is characterized by a complex Wishart distribution $N\hat{R} \sim \mathcal{CW}(N, M, R)$ that is nondegenerate for $N \geq M$, and thus, the CRB can be readily calculated [15] to provide an asymptotic bound on DOA estimation accuracy.

Let $\{\hat{\boldsymbol{\lambda}}, \hat{U}\}$ be the eigendecomposition of \hat{R} , where the eigenvalues $\hat{\boldsymbol{\lambda}} = [\hat{\lambda}_1, \dots, \hat{\lambda}_M]$ are arranged in decreasing order of size, with corresponding normalized eigenvectors $\hat{U} = [\hat{\mathbf{u}}_1, \dots, \hat{\mathbf{u}}_M]$. Let us partition these into signal and noise subspace components with the notation $\hat{U}_s = [\hat{\mathbf{u}}_1, \dots, \hat{\mathbf{u}}_m]$ and $\hat{U}_n = [\hat{\mathbf{u}}_{m+1}, \dots, \hat{\mathbf{u}}_M]$ so that $\hat{U} = [\hat{U}_s, \hat{U}_n]$, and similarly, $\hat{\boldsymbol{\lambda}} = [\hat{\boldsymbol{\lambda}}_s, \hat{\boldsymbol{\lambda}}_n]$.

In the conventional case, the MUSIC algorithm applied to the DDC matrix \hat{R} forms DOA estimates $\hat{\mathbf{w}}$ by choosing the

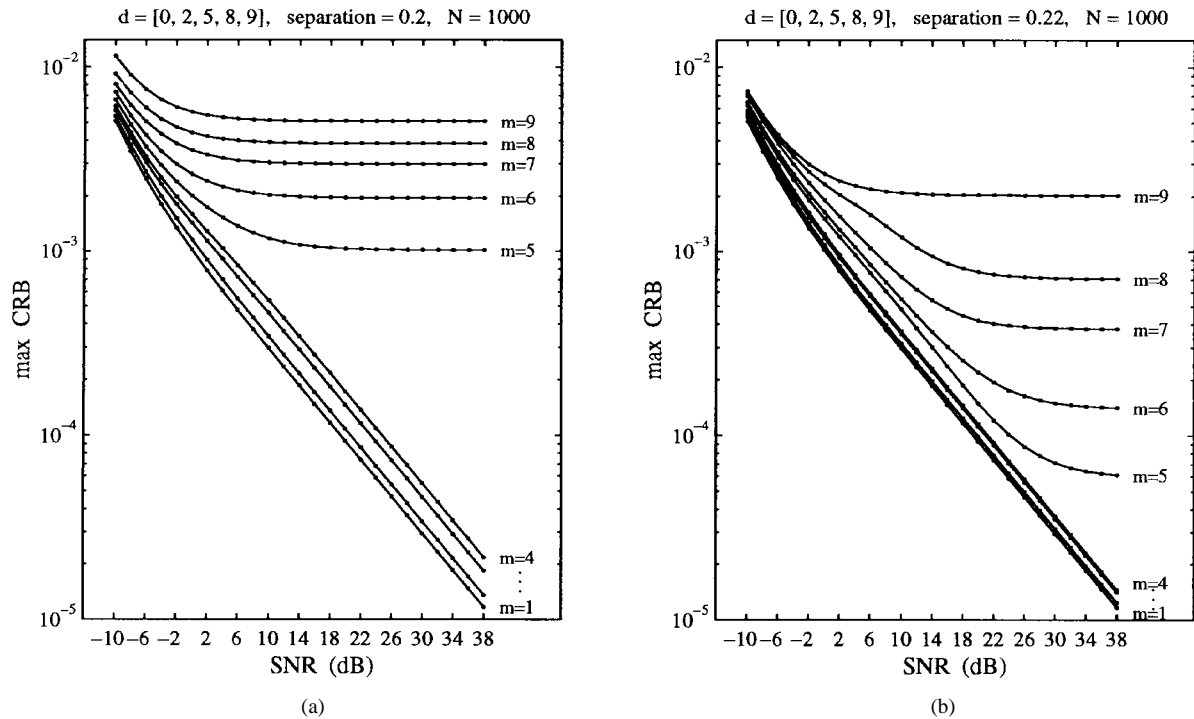


Fig. 1. Maximum CRB for a five-element NLA with antennas located at positions $\mathbf{d} = [0, 2, 5, 8, 9]$ (measured in half-wavelength units), $N = 1000$ snapshots and m sources located at DOA's. (a) $w \equiv \sin \theta = [-1.00, -0.80, -0.60, \dots]$. (b) $w = [-1.00, -0.78, -0.56, \dots]$, each with common SNR.

m values of w for which the function $f_{\hat{R}}(w)$ is minimized

$$f_{\hat{R}}(w) = \mathbf{s}^H(w) \hat{U}_n \hat{U}_n^H \mathbf{s}(w). \quad (9)$$

It has been demonstrated [8] that under this uncorrelated signal model, the DOA estimation accuracy of MUSIC asymptotically approaches the CRB limit (for $m < M$). To illustrate the typical distinction between conventional and superior cases, consider the following CRB analysis of the five-element "minimum-redundancy" array [16], where

$$\mathbf{d}_5 = [0, 2, 5, 8, 9] \quad (10)$$

with $M = 5$ and $M_\alpha = 10$ (described by Moffet [1] as a "restricted C-sequence"). The signal environment consists of $m = 1, \dots, 9$ sources with common SNR located at DOA's $w = [-1.0, -0.8, -0.6, \dots]$, and we use $N = 1000$ snapshots. Fig. 1(a) shows the maximum CRB calculated by the well-known expression [15] involving the Fisher information matrix element

$$J_{xy} = N \text{Tr} \left\{ R^{-1} \frac{\partial R}{\partial x} R^{-1} \frac{\partial R}{\partial y} \right\}. \quad (11)$$

We see a clear distinction between the limit accuracy for conventional ($1 \leq m \leq 4$) and superior ($5 \leq m \leq 9$) scenarios. As expected, the conventional regime displays a strong inverse semi-logarithmic dependence on SNR. On the contrary, the superior regime shows negligible dependence on SNR above some positive threshold value and demonstrates the very large reduction in potential accuracy compared with the conventional case for large SNR. However, when signals are below the white noise level ($\text{SNR} \ll 1$), degradation is not significant, and a large sample volume ($N \gg 1$) is necessary to achieve a reasonable accuracy in both regimes.

Fig. 1(b) shows the CRB analysis for the same scenario but with a slightly increased source separation. We see that here, the threshold SNR values are increased and are now more dependent on the actual number of superior sources. The limit DOA estimation accuracy (for $\text{SNR} \rightarrow \infty$) also depends more strongly on the actual number of sources. Nevertheless, the main distinction between conventional and superior scenarios is essentially the same for both separations. Thus, for the superior case, we always require a large sample volume, whereas for the conventional case, a lack of samples may be compensated for by a corresponding increase in SNR.

Obviously, the actual accuracy of any particular algorithm does not necessarily coincide with the CRB behavior, even if the algorithm is asymptotically efficient. It has long been known that for a relatively small signal separation, SNR, and sample volume, MUSIC fails to approach the corresponding CRB.

In order to carefully examine some relevant issues, let us consider Fig. 2, which presents the results of DOA estimation simulations using the MUSIC algorithm. These calculations are based on the minimum-redundancy array ("optimal-lag" array [16], in fact)

$$\mathbf{d}_4 = [0, 1, 4, 6] \quad (12)$$

measured in half wavelength units. This example is specifically tailored to enable comparison with another approach [12]. The covariance matrix estimate \hat{R} is obtained from $N = 100$ snapshots. The signal environment consists of two sources with a common but varying SNR located at spatial frequencies $w = [0, 0.12]$, i.e., at angles $\theta = [0^\circ, 6.9^\circ]$. DOA estimation statistics are presented as a function of SNR. The number of so-called abnormal trials is plotted in Fig. 2(b). In this paper,

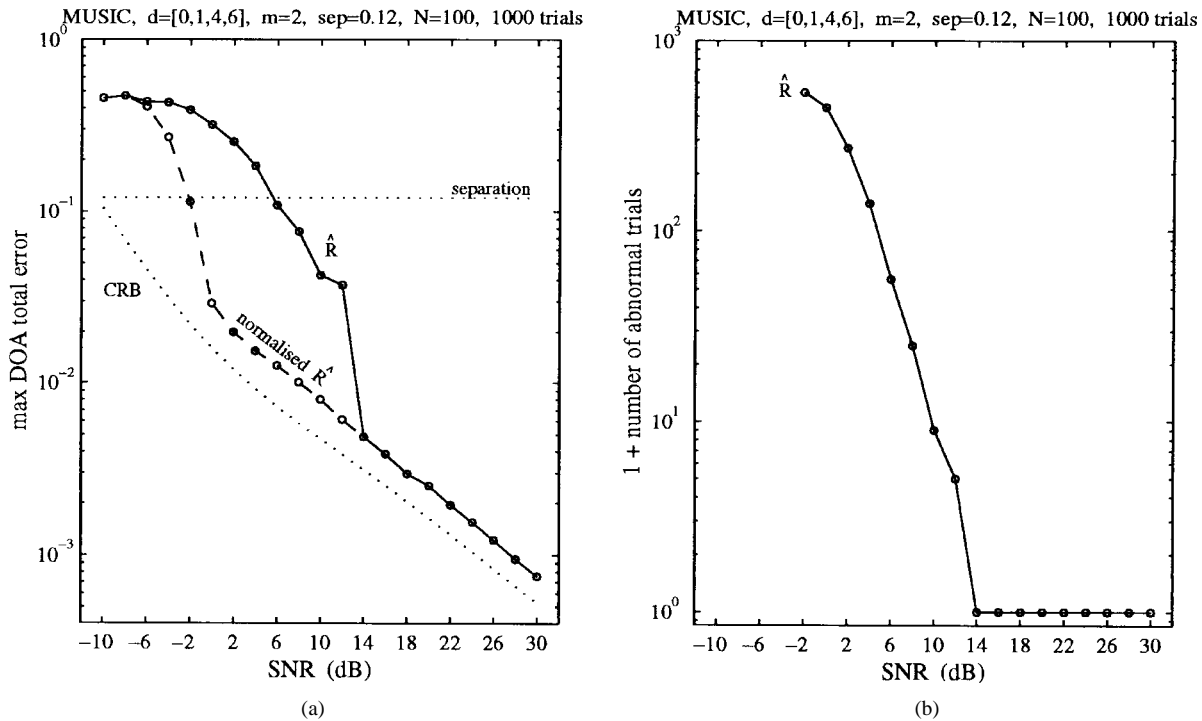


Fig. 2. Typical DOA estimation performance for MUSIC over 1000 simulated trials for the four-element NLA $\mathbf{d} = [0, 1, 4, 6]$, $m = 2$ sources located at DOA's $w = [0, 0.12]$ with common SNR, and $N = 100$ snapshots. (a) Comparison of MUSIC performance against “normalized” results and the maximum CRB. (b) Corresponding number of “abnormal” trials.

we define “abnormal” estimates to be DOA estimates \hat{w}_j lying outside the range $[w_j - \Delta, w_j + \Delta]$, where Δ is 20 times the appropriate CRB. In turn, an abnormal trial is defined to be one that includes one or more abnormal estimates.

Interpretation of these results confirm some well-known facts. There is a clear division between the preasymptotic domain and the asymptotic domain at some threshold SNR value (in this case, at 14 dB). The asymptotic domain is characterized by DOA estimates that essentially correspond to the theoretical limit accuracy for ML methods provided by the CRB with the number of abnormal estimates diminishing to zero. Clearly, in this domain, the MUSIC algorithm is almost optimally accurate, and a method for improving DOA estimates is unnecessary. In the preasymptotic domain, however, the performance of MUSIC degrades dramatically due to contamination by abnormal trials. To demonstrate this, Fig. 2(a) also shows the performance of the “normalized” estimates, where the abnormal trials are simply excluded. Normalized \hat{R} performance closely matches the theoretic bound, which confirms the assertion that the asymptotic domain may be defined by the conditions under which the probability of abnormal estimates is negligible.

Thus, for conventional scenarios, we should focus our efforts on investigating new approaches that reduce the probability of abnormal estimates in the preasymptotic domain. For superior scenarios, the problem is quite different. According to the results of CRB analysis, we should only be interested in the asymptotic behavior ($N \gg 1$). Comparison of the direct augmentation approach (DAA) [6], [7] against the CRB demonstrates that even asymptotically, DAA performance is usually far worse than the corresponding CRB, and there-

fore, considerable effort should be directed toward asymptotic efficiency improvement.

III. FINITE SIGNAL SUBSPACE POSITIVE-DEFINITE TOEPLITZ COMPLETION FOR THE CONVENTIONAL CASE

The idea behind the proposed approach is quite straightforward. When the MUSIC algorithm is applied to any arbitrary positive-definite (p.d.) Hermitian matrix (such as \hat{R}), the resulting MUSIC pseudo-spectrum is not guaranteed to contain exactly m peaks, even if it has an m -dimensional signal subspace. Since a fully augmentable M_α -element sparse array is similar to a M_α -element uniform linear array (ULA) in that all covariance lags are present, we may try to construct from the DDC matrix \hat{R} the M_α -variate p.d. Toeplitz matrix \hat{T} that is “closest” in some sense to \hat{R} , which has its $(M_\alpha - m)$ smallest eigenvalues equal in value. Such a Toeplitz matrix would have precisely m peaks, provided the largest m eigenvalues (the signal subspace eigenvalues) are different from the remaining $(M_\alpha - m)$ eigenvalues (the noise subspace eigenvalues). Naturally, the known number of sources m that is used to construct the traditional estimator $\text{MUSIC}(\hat{R})$ is once again used to construct \hat{T} .

Note that the uniform (“augmented” or “virtual”) array geometry is implied by the properties of the corresponding Toeplitz matrix. Indeed, if we try to find the best Hermitian approximation in the L_2 sense with an m -dimensional signal subspace, then we should simply equalize the $(M_\alpha - m)$ smallest eigenvalues [17]. Obviously, such “signal subspace truncation” (SST) alters neither the eigenvectors nor, correspondingly, the MUSIC pseudo-spectrum. However, for a Toeplitz matrix, this is no longer true, and the SST process

must necessarily change the eigenvectors so that the MUSIC pseudo-spectrum corresponds to a mixture of m sinusoids in white noise.

The following two-step procedure is proposed to find such a p.d. Toeplitz matrix with m -variate signal eigensubspace. First, given \hat{R} , we find the closest (in some sense) M_α -variate Toeplitz estimate \hat{T}_0 , which need not have a signal subspace of the correct dimension. Second, we utilize recently described [18] alternating projections to form a p.d. Toeplitz solution \hat{T}_m with the required m -dimensional signal subspace. This alternating projection method minimizes the Frobenius norm $\|\hat{T}_0 - \hat{T}_m\|_F$.

The construction method for \hat{T}_0 must meet the condition that in the asymptotic domain the DOA estimates obtained from MUSIC(\hat{T}_0) are statistically as accurate as those obtained from the asymptotically optimal estimator MUSIC(\hat{R}). To meet this condition, the following approach is proposed. Given the DDC matrix \hat{R} , form the standard MUSIC pseudo-spectrum $f_{\hat{R}}(w) \geq 0$, and then, find the M_α -variate p.d. Toeplitz matrix \hat{T}_0 whose maximum-entropy (ME) spectrum is strictly equal to $f_{\hat{R}}(w)$.

Let $\{\hat{\sigma}, \hat{V}\}$ be the eigendecomposition of some M_α -variate p.d. matrix \hat{T}_0 (for this paragraph, we only require that \hat{T}_0 is Hermitian and not Toeplitz)

$$\hat{T}_0 = \sum_{j=1}^{M_\alpha} \hat{\sigma}_j \hat{\mathbf{v}}_j \hat{\mathbf{v}}_j^H, \quad \hat{\mathbf{v}}_j^H \hat{\mathbf{v}}_k = \delta_{jk}. \quad (13)$$

Then, the ME spectrum of \hat{T}_0 may be written in the form

$$f_{\text{ME}}^{-1}(w) = \left| \mathbf{a}^H(w) \left(\sum_{j=1}^{M_\alpha} \hat{\sigma}_j^{-1} \hat{\mathbf{v}}_j \hat{\mathbf{v}}_j^H \right) \mathbf{e}_1 \right|^{-2} \quad (14)$$

where $A \equiv [\mathbf{a}(w_1), \dots, \mathbf{a}(w_m)]$ is the M_α -variate ULA signal manifold matrix, and $\mathbf{e}_1 = [1, 0, \dots, 0]^T$. For a sufficiently large SNR, we may assume that

$$\hat{\sigma}_j \gg \hat{\sigma}_{m+1} \quad \text{for } j = 1, \dots, m. \quad (15)$$

Hence

$$f_{\text{ME}}^{-1}(w) = \left| \mathbf{a}^H(w) \sum_{j=m+1}^{M_\alpha} \hat{\sigma}_j^{-1} (\hat{\mathbf{v}}_j^H \mathbf{e}_1) \hat{\mathbf{v}}_j \right|^{-2} \quad (16)$$

whereas the MUSIC pseudo-spectrum of \hat{T}_0 is

$$f_{\hat{T}_0}^{-1}(w) = \left| \mathbf{a}^H(w) \sum_{j=m+1}^{M_\alpha} \hat{\mathbf{v}}_j \hat{\mathbf{v}}_j^H \mathbf{a}(w) \right|^{-1}. \quad (17)$$

Comparison of the last two equations leads to the conclusion that in order to receive a spurious peak in the MUSIC pseudo-spectrum in some direction w_0 , all noise subspace eigenvectors \hat{V}_n must exhibit the orthogonal property

$$|\mathbf{a}^H(w_0) \hat{\mathbf{v}}_j| \simeq 0 \quad \text{for } j = m+1, \dots, M_\alpha \quad (18)$$

whereas for the ME spectrum, the coherent weighted sum

$$\sum_{j=m+1}^{M_\alpha} \mathbf{a}^H(w_0) \hat{\mathbf{v}}_j [\hat{\sigma}_j^{-1} \hat{\mathbf{v}}_j^H \mathbf{e}_1] \quad (19)$$

may give rise to a false peak, even when (18) does not hold. In other words, a peak in the MUSIC pseudo-spectrum necessarily implies a (corresponding) peak in the ME spectrum but not *vice versa*. Because of this correspondence, a MUSIC pseudo-spectrum cannot contain spurious peaks if the ME spectrum has none. Thus, for the asymptotic domain, where ME(\hat{T}_0) duplicates the (correct) peaks of MUSIC(\hat{R}), we see that MUSIC(\hat{T}_0) is guaranteed to duplicate those same correct peaks.

Obviously, the same known number of signal sources m is used for both MUSIC pseudo-spectra $f_{\hat{R}}(w)$ and $f_{\hat{T}_0}(w)$. This is important in order to understand the mechanism of abnormal estimate removal. As we have mentioned above, even if MUSIC(\hat{R}) has a spurious peak, this is not necessarily reproduced in MUSIC(\hat{T}_0). Moreover, when the process of signal subspace truncation takes place, only the true number of peaks m will remain. A false peak in the original MUSIC pseudo-spectrum of \hat{R} should survive if it is significantly higher than the true DOA peaks.

Thus, the first algorithm in our approach, which is called "MUSIC-ME equalization" (MMEE), is to construct the p.d. Toeplitz matrix \hat{T}_0 with a given ME spectrum. This is achieved on the basis of the following lemma.

Lemma 1 [19]: The necessary and sufficient condition for an arbitrary M_α -variate vector $\hat{\mathbf{w}} \equiv [\hat{w}_1, \dots, \hat{w}_{M_\alpha}]^T$ to be presented in the form

$$\hat{\mathbf{w}} = \hat{T}_0^{-1} \mathbf{e}_1, \quad \hat{T}_0 > 0, \quad \hat{w}_1 > 0 \quad (20)$$

is the absence of zeros inside the unit circle for the polynomial

$$\hat{W}(z) := \sum_{j=1}^{M_\alpha} \hat{w}_j z^{j-1} \quad (21)$$

i.e.,

$$\hat{W}(z) \neq 0 \quad \text{for } |z| \leq 1. \quad (22)$$

The proof of this lemma appears in Appendix A. Based on this lemma, we may now present the MMEE algorithm.

A. MUSIC-ME Equalization (MMEE)

Conventional-Case Algorithm— $\hat{R} \xrightarrow{\text{MMEE}} \hat{T}_0$

Step 1: Given the M -variate DDC matrix \hat{R} , define the MUSIC pseudo-spectrum

$$f_{\hat{R}}(w) = \mathbf{s}^H(w) \hat{U}_n \hat{U}_n^H \mathbf{s}(w). \quad (23)$$

Step 2: For $z = e^{i\pi w}$, and where $f_{\hat{R}}(z) = f_{\hat{R}}(w)$ on the unit circle $|z| = 1$, conduct the spectral factorization

$$f_{\hat{R}}(z) = \hat{W}(z) \hat{W}^H(1/z^*) \quad (24)$$

in order to define the unique M_α -variate polynomial $\hat{W}(z)$ with the property

$$\hat{W}(z) \neq 0 \quad \text{for } |z| \leq 1. \quad (25)$$

Step 3: Calculate the vector $\hat{\mathbf{w}}'$ by the coefficients of the modified polynomial $\hat{W}'(z) = e^{i\gamma}\hat{W}(z)$, where

$$\hat{W}(z) = \prod_{j=1}^{M_\alpha-1} (z - \hat{z}_j) \quad (26)$$

and

$$\gamma = \sum_{j=1}^{M_\alpha-1} \arg \hat{z}_j + \pi(M_\alpha - 1) \quad (27)$$

so that

$$\hat{w}'_1 = \hat{W}'(0) = \prod_{j=1}^{M_\alpha-1} |\hat{z}_j| > 0. \quad (28)$$

Step 4: Given the vector $\hat{\mathbf{w}}' \equiv [\hat{w}'_1, \dots, \hat{w}'_{M_\alpha}]^T$, compute the M_α -variate p.d. Toeplitz matrix \hat{T}_0 from the Gohberg–Semencul formula [20], [21]

$$\hat{T}_0 = \hat{w}'_1(AA^H - B^HB)^{-1} \quad (29)$$

where A and B are the Toeplitz matrices

$$A = \begin{bmatrix} \hat{w}'_1 & 0 & 0 & \cdots & 0 \\ \hat{w}'_2 & \hat{w}'_1 & 0 & \cdots & 0 \\ \hat{w}'_3 & \hat{w}'_2 & \hat{w}'_1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \hat{w}'_{M_\alpha} & \hat{w}'_{M_\alpha-1} & \hat{w}'_{M_\alpha-2} & \cdots & \hat{w}'_1 \end{bmatrix}$$

$$B = \begin{bmatrix} 0 & \hat{w}'_{M_\alpha} & \hat{w}'_{M_\alpha-1} & \hat{w}'_{M_\alpha-2} & \cdots & \hat{w}'_2 \\ 0 & 0 & \hat{w}'_{M_\alpha} & \hat{w}'_{M_\alpha-1} & \cdots & \hat{w}'_3 \\ 0 & 0 & 0 & \hat{w}'_{M_\alpha} & \cdots & \hat{w}'_4 \\ 0 & 0 & 0 & 0 & \cdots & \hat{w}'_5 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 0 \end{bmatrix}. \quad (30)$$

Several comments are in order. Since the MUSIC pseudo-spectrum is defined as a non-negative integrable function, namely, $f_{\hat{R}}(z) \geq 0$ for $|z| = 1$, the spectral factorization in Step 2 is guaranteed to exist by the Fejer–Riesz theorem [22]. In Step 3, we simply shift the initial phase in order to have a positive \hat{w}'_1 . Thus, after Step 3, we have the vector $\hat{\mathbf{w}}'$, which satisfies the necessary and sufficient condition of Lemma 1 and has the spectrum $|\hat{W}'(z)|^2$ for $|z| = 1$ equal to the given MUSIC pseudo-spectrum $f_{\hat{R}}(w)$. The famous Gohberg–Semencul formula in Step 4 simply reconstructs the entire p.d. Toeplitz Hermitian matrix, given the first column of its matrix inversion.

Now that we have computed the p.d. Toeplitz Hermitian matrix \hat{T}_0 with the desired ME spectrum, we may apply alternating projections to define \hat{T}_m such that

$$\hat{T}_m = \arg \inf_{\hat{T}_m} \|\hat{T}_0 - \hat{T}_m\|_F \quad (31)$$

where \hat{T}_m is in the class of matrices defined to be of the form

$$\hat{T}_m = p_0 I_{M_\alpha} + V_s \overset{\circ}{\Sigma}_s V_s^H \in \mathcal{T}^{M_\alpha} \quad (32)$$

$$\overset{\circ}{\Sigma}_s = \text{diag}[\sigma_1 - p_0, \dots, \sigma_m - p_0] > 0, \quad p_0 \geq 0 \quad (33)$$

where \mathcal{T}^p is the space of $p \times p$ Hermitian Toeplitz matrices, and V_s is the $M_\alpha \times m$ matrix composed of the normalized signal subspace eigenvectors of \hat{T}_m . For the Frobenius norm, this optimization problem is found to be a slight modification of those investigated by Grigoriadis *et al.* [18].

Based on this approach, the following SST algorithm is proposed.

B. Signal Subspace Truncation (SST)

Conventional-Case Algorithm— $\hat{T}_0 \xrightarrow{\text{SST}} \hat{T}_m$

Step 1: Given the M_α -variate p.d. Toeplitz matrix \hat{T}_0 , with

$$\hat{T}_0 = \sum_{j=1}^{M_\alpha} \hat{\sigma}_j \mathbf{v}_j \mathbf{v}_j^H, \quad \hat{\sigma}_j > 0, \quad \mathbf{v}_j^H \mathbf{v}_k = \delta_{jk} \quad (34)$$

set the initial estimate $\hat{T}^{(0)} = \hat{T}_0$, and form the corresponding p.d. Hermitian matrix $\hat{H}^{(0)} \equiv \{\hat{h}_{jk}^{(0)}\}_{j,k=1}^{M_\alpha}$ with m -dimensional signal subspace

$$\hat{H}^{(0)} = \bar{\sigma}_n^{(0)} I_{M_\alpha} + \sum_{j=1}^m (\hat{\sigma}_j - \bar{\sigma}_n^{(0)}) \mathbf{v}_j \mathbf{v}_j^H \quad (35)$$

where

$$\bar{\sigma}_n^{(0)} = \frac{1}{M_\alpha - m} \sum_{j=m+1}^{M_\alpha} \hat{\sigma}_j. \quad (36)$$

Step 2: Find the corresponding Toeplitz matrix $\hat{T}^{(1)} \equiv \{\hat{t}_{jk}^{(1)}\}_{j,k=1}^{M_\alpha}$ by “redundancy averaging”

$$\hat{t}_0^{(1)} = \frac{1}{M_\alpha} \sum_{j=1}^{M_\alpha} \hat{h}_{jj}^{(0)}, \quad \hat{t}_1^{(1)} = \frac{1}{M_\alpha - 1} \sum_{j=1}^{M_\alpha-1} \hat{h}_{j,j+1}^{(0)}$$

$$\hat{t}_2^{(1)} = \frac{1}{M_\alpha - 2} \sum_{j=1}^{M_\alpha-2} \hat{h}_{j,j+2}^{(0)}, \dots, \hat{t}_{M_\alpha-1}^{(1)} = \hat{h}_{1,M_\alpha}^{(0)}. \quad (37)$$

Step 3: Perform an eigendecomposition of $\hat{T}^{(1)}$ to obtain the eigenvalues $\hat{\sigma}^{(1)}$. If

$$\frac{\hat{\sigma}_{m+1}^{(1)} - \hat{\sigma}_{M_\alpha}^{(1)}}{\hat{\sigma}_{M_\alpha}^{(1)}} > \varepsilon \ll 1 \quad (38)$$

then replace $\hat{T}^{(0)}$ in Step 1 with $\hat{T}^{(1)}$, and repeat Step 2 to obtain the next iterate $\hat{T}^{(2)}$.

Step 4: Continue iterating the alternating projections until the stopping condition

$$\frac{\hat{\sigma}_{m+1}^{(\kappa)} - \hat{\sigma}_{M_\alpha}^{(\kappa)}}{\hat{\sigma}_{M_\alpha}^{(\kappa)}} \leq \varepsilon \quad (39)$$

is satisfied, and then, treat $\hat{T}^{(\kappa)}$ as the final solution \hat{T}_m .

The following comments are made to justify this approach.

Given any initial positive-semidefinite Hermitian matrix \hat{T}_0 , Step 1 defines an orthogonal projection on the subset of p.d. Hermitian matrices with m -dimensional signal subspace. The proof of this assertion can be easily obtained from the minimum Frobenius norm condition [17]. Step 2 is known to perform the orthogonal projection on the set of Toeplitz

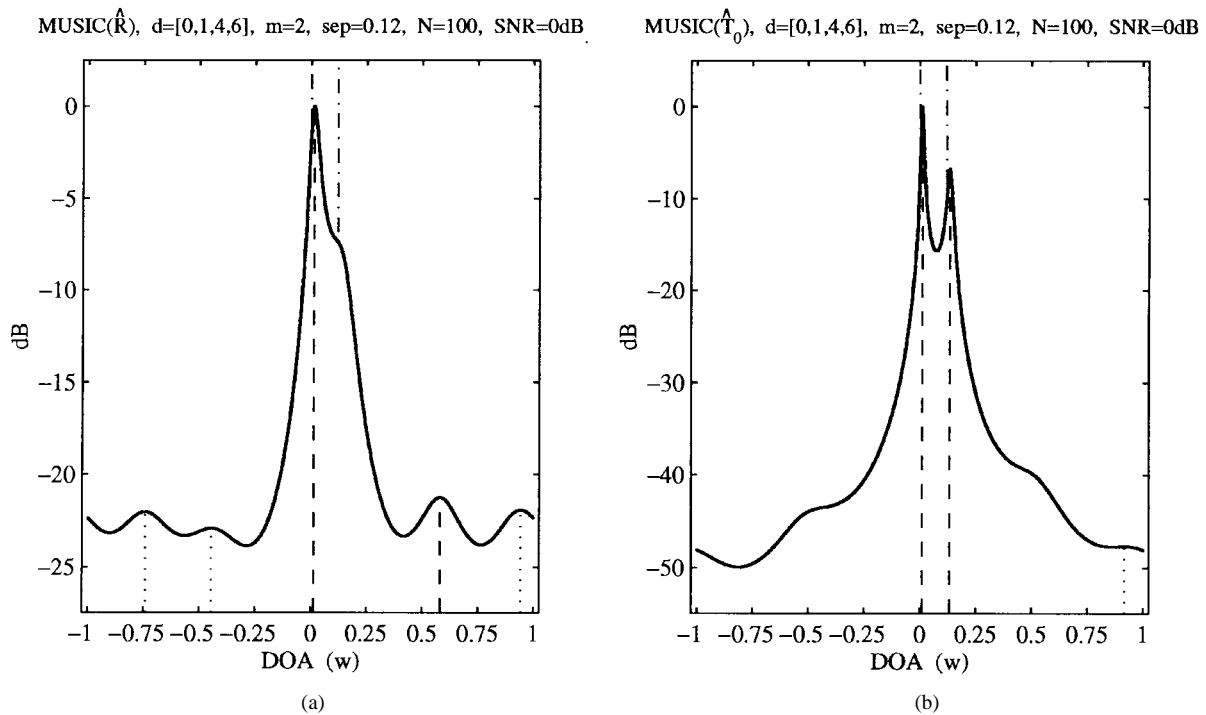


Fig. 3. Example trial from the Fig. 2 scenario, where (a) the MUSIC pseudo-spectrum of \hat{R} gives rise to an abnormal DOA estimate for the second source $w_2 = 0.12$, whereas (b) the MUSIC pseudo-spectrum of \hat{T}_0 gives rise to two normal DOA estimates (and a lower noise floor). Dot-dashed vertical lines indicate the two exact DOA's; dashed lines show the two MUSIC-selected maxima; dotted lines mark the remaining ancillary maxima.

matrices for any given matrix [18]. The algorithm is written here in the form dealing with unknown white noise power p_0 (estimated by $\hat{p}_0 = \bar{\sigma}_n^{(\kappa)}$), which is a typical result from the MMEE algorithm. In some applications, the white noise power p_0 may be known, in which case, the corresponding matrices should be corrected by the term

$$(p_0 - \bar{\sigma}_n^{(\ell)})I_{M_\alpha}, \quad \text{for } \ell = 0, \dots, \kappa. \quad (40)$$

Note that the redundancy averaging process in Step 2 generally does not guarantee the positive-definiteness of the Toeplitz matrices $\hat{T}^{(\ell)}$. Since in most practical situations at least the first m (signal) eigenvalues are assumed to be always positive, the white noise power estimate $\bar{\sigma}_n^{(\kappa)}$ should be obtained by averaging over the positive subset of the last $(M_\alpha - m)$ eigenvalues

$$\bar{\sigma}_n^{(\ell+1)} = \frac{1}{M_\alpha - m} \sum_{j=m+1}^{M_\alpha} \hat{\sigma}_j^{(\ell)+}, \quad \hat{\sigma}_j^{(\ell)+} > 0. \quad (41)$$

It is straightforward to show that whereas the set of p.d. Toeplitz matrices is convex, the subset of all \hat{T}_m of the form shown in (32) is not. Thus, the convergence of our alternating projections to the global extremum is not guaranteed, and SST may result in an unsatisfactory local minimum. Hence, the accuracy of the initial estimate \hat{T}_0 is essential for the success of the entire approach.

To illustrate the typical mechanism for the occurrence of abnormal estimates and their removal by the MMEE and SST algorithms, Fig. 3(a) shows the MUSIC pseudo-spectrum for one particular abnormal trial from the set of 1000 \mathbf{d}_4 simulation trials shown in Fig. 2. We see that the primary peak more or

less correctly identifies the first source $w_1 = 0$ at $\hat{w}_1 = 0.01$, whereas the secondary peak associated with the second source at $w_2 = 0.12$ has merged into the primary peak; hence, MUSIC erroneously identifies one of the smaller ancillary peaks as $\hat{w}_2 = 0.58$. (It is interesting that in this particular trial, root-MUSIC gave a normal estimate; the reverse situation, where root-MUSIC is abnormal but MUSIC is normal, also occurs.) Fig. 3(b) now shows the MUSIC pseudo-spectrum of \hat{T}_0 for the same trial. Due to the improved pseudo-spectral peak resolution of the estimator \hat{T}_0 , the abnormal DOA estimate has been eliminated. The noise floor has also been reduced.

This mechanism accounts for the overall reduction in the proportion of abnormal estimates when using \hat{T}_m compared with the standard estimator \hat{R} . This performance improvement is illustrated in Fig. 4 for the same set of 1000 \mathbf{d}_4 simulation trials as in Fig. 2. Since the asymptotic domain is defined by the SNR region where abnormal DOA estimates have negligible probability, our proposed MMEE and SST algorithms result in the estimator \hat{T}_m , which effectively extends the asymptotic domain by as much as 16 dB (in this case). We may also note that the MMEE algorithm alone provides a significant reduction in the number of abnormal estimates since \hat{T}_0 demonstrates much better performance than \hat{R} .

It should be noted that our approach cannot eliminate all abnormal estimates; it can only extend the area where they do not exist, compared with conventional subspace-based algorithms like MUSIC. As expected, the DOA estimation accuracy for the new technique MUSIC(\hat{T}_m) is identical to that of the conventional technique MUSIC(\hat{R}) if abnormal estimates are either not present (in the asymptotic domain) or are not taken into account (for the "normalized" datasets).

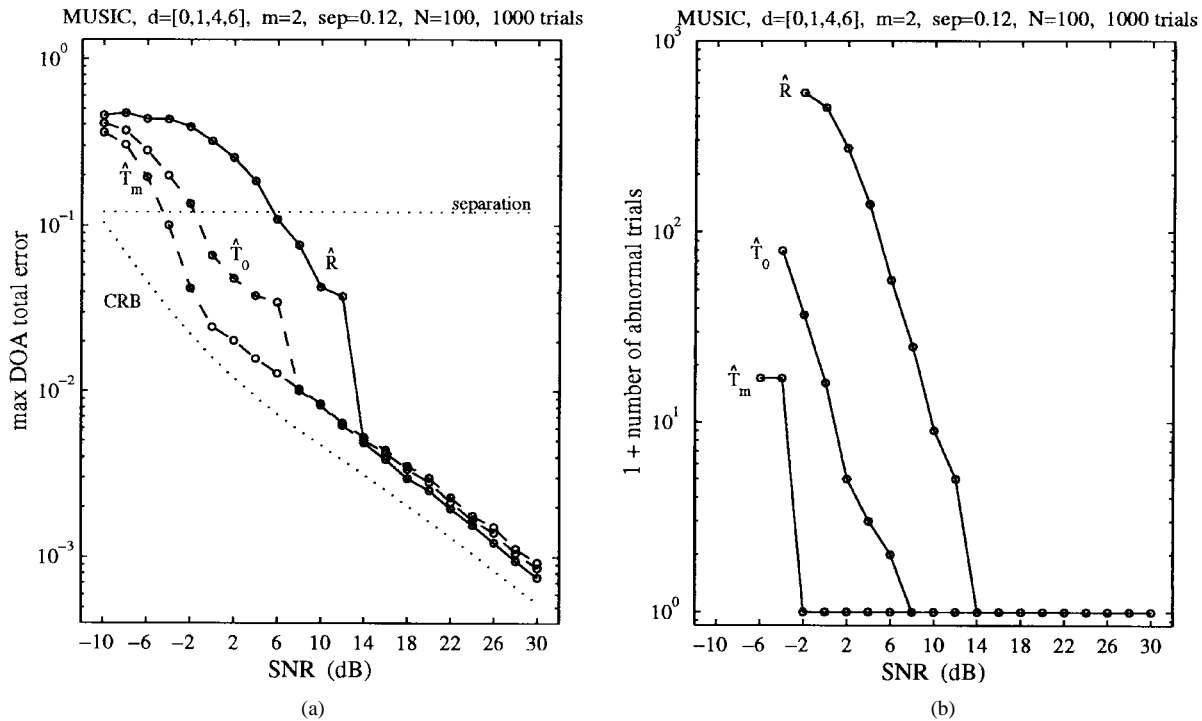


Fig. 4. Identical scenario as in Fig. 2 but showing (a) the improved performance of the new technique involving the estimators \hat{T}_0 and \hat{T}_m compared with the standard \hat{R} approach and (b) the corresponding number of “abnormal” trials.

Note that instead of using alternating projections to minimize the Frobenius norm, we can apply the slightly different norm with its associated projections proposed in [23].

Naturally, there is no need for an asymptotic analysis of the proposed approach since such an analysis would ignore abnormalities, whereas the standard first-order approximation would be essentially the same as for the conventional MUSIC(\hat{R}). It seems that a preasymptotic statistical analysis is still the challenge that has yet to be properly addressed.

IV. FINITE SIGNAL SUBSPACE POSITIVE-DEFINITE TOEPLITZ COMPLETION FOR THE SUPERIOR CASE

In the superior case ($M \leq m < M_\alpha$), the M -variate spatial covariance matrix R and its stochastic estimate \hat{R} is of full rank, even if the additive white noise power is set to zero. Obviously, none of the subspace-based techniques are applicable here. A well-known approach in this case is to utilize the DAA [6], [7], which relies on the one-to-one correspondence between the covariance lags of R and T : the Toeplitz covariance matrix of the M_α -variate “augmented” ULA

$$R_{jk} = T(d_j - d_k) \quad \text{for } j, k = 1, \dots, M. \quad (42)$$

This correspondence evidently exists for all fully augmentable arrays. The DAA philosophy is to estimate the Toeplitz matrix covariance lags

$$\bar{T}_0 \equiv \{\bar{t}_{j-k}\}_{j,k=1}^{M_\alpha} \quad (43)$$

by simply averaging over the set of corresponding covariance

lags taken from \hat{R}

$$\bar{t}_{j-k=\kappa} = \frac{\sum_{j,k=1}^M \hat{R}_{jk} \delta(\kappa d, d_j - d_k)}{\sum_{j,k=1}^M \delta(\kappa d, d_j - d_k)}, \quad j > k \quad (44)$$

where $\delta(a, b)$ is the generalized Kronecker delta function. (We have introduced the notation \bar{T}_0 instead of \hat{T}_0 to emphasize the different origin of these two Toeplitz matrices.) Clearly, the augmented covariance matrix \bar{T}_0 is even not necessarily p.d. Since the signal subspace cannot be separated from the noise subspace for the DDC matrix \hat{R} , we should expect good DOA estimation accuracy when the estimates \hat{R}_{jk} are simply sufficiently close to the true values R_{jk} (so-called “strong convergence”), i.e., when

$$\|\hat{R}_{jk} - R_{jk}\| < \varepsilon \ll 1 \quad \text{for } j, k = 1, \dots, M. \quad (45)$$

For conventional scenarios, we may instead use the weaker noise subspace separability condition

$$\|\hat{V}_s^H V_n V_n^H\| < \varepsilon. \quad (46)$$

Even m snapshots of reasonably strong signals are sufficient to form a \hat{V}_s , which meets this condition. This explains why, for the conventional case, the lack of samples may always be compensated for by increasing SNR, whereas the strong convergence condition always demands an appropriate N , regardless of SNR. This is precisely what Fig. 1 demonstrates. Thus, we should now focus on asymptotically large sample volumes that are sufficient for strong convergence.

Note that previous attempts at DAA statistical analysis [7] did not involve super-resolution techniques, such as MUSIC.

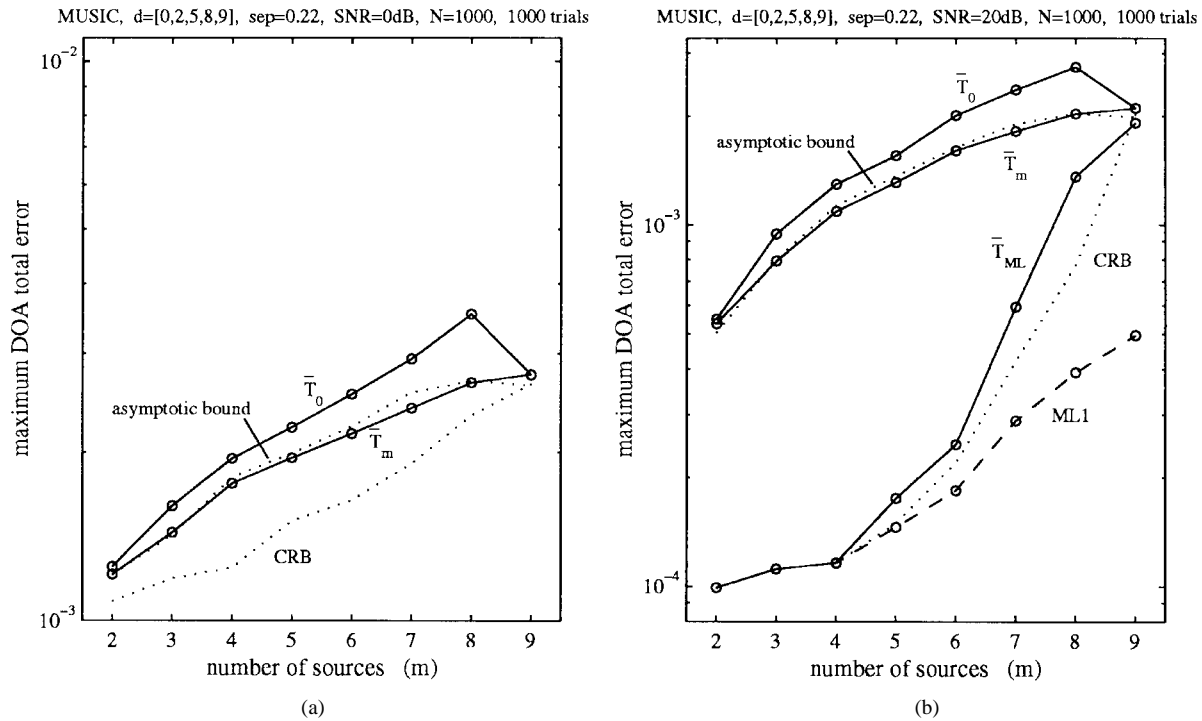


Fig. 5. Typical MUSIC DOA estimation performance for the same signal model as in Fig. 1(b) with $N = 1000$ snapshots, 1000 trials, and SNR of (a) 0 dB and (b) 20 dB, showing good performance for the new technique involving the estimators \bar{T}_0 , \bar{T}_m , and \bar{T}_{ML} compared with the CRB and two other computed bounds.

An adequate analysis has been performed only recently [24]. The main results of this asymptotic analysis are reproduced below in order to demonstrate that DAA provides reasonable accuracy (in the CRB sense) only for weak signals ($\text{SNR} \leq 0$ dB). For strong signals, DAA's accuracy is much poorer than CRB predictions. Thus, DAA should be treated only as an initial step: one that needs to be substantially improved upon.

The following theorem describes the asymptotic distribution of DOA estimates obtained from the estimator MUSIC(\bar{T}_0). The proof appears in Appendix B.

Theorem 1 [24], [25]: Let $N\hat{R} \sim CW(N, M, R)$, and let \bar{T}_0 be defined by the element-wise linear transform (DAA) of (44), i.e., $\bar{T}_0 = \mathcal{L}(\hat{R})$. As before, let $R = p_0 I_M + SPS^H = U\Lambda U^H$, where $U \equiv [\mathbf{u}_1, \dots, \mathbf{u}_M]$ is the M -variate unitary matrix of eigenvectors with corresponding eigenvalues $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_M$, so that $\Lambda = \text{diag}[\lambda_j]_{j=1, M}$, and let $T = p_0 I_{M_\alpha} + APA^H = V\Sigma V^H$, where $V \equiv [\mathbf{v}_1, \dots, \mathbf{v}_{M_\alpha}]$ is the M_α -variate unitary matrix of eigenvectors with corresponding eigenvalues $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_{M_\alpha}$, so that $\Sigma = \text{diag}[\sigma_j]_{j=1, M_\alpha}$ and $\sigma_{m+1} = \dots = \sigma_{M_\alpha} = p_0$, and where A is the array manifold matrix for the M_α -element ULA. Then, the vector of DOA estimates $\hat{\mathbf{w}} \equiv [\hat{w}_1, \dots, \hat{w}_m]^T$ has the asymptotic Gaussian distribution $\mathcal{N}(m, W, \Theta)$ with respect to N , where

$$\Theta = \{\Theta_{pq}\}_{p, q=1}^m$$

$$\Theta_{pq} = \frac{1}{2N} \sum_{j \leq k}^M \lambda_j \lambda_k (1 + \delta_{jk})^{-1} \text{Re}\{\Gamma_p^{jk} \Gamma_q^{jk*}\} + O(N^{-1}) \quad (47)$$

$$\Gamma_\ell^{jk} = \frac{1}{\gamma_\ell} [\boldsymbol{\alpha}_\ell^H \mathcal{L}(\mathbf{u}_j \mathbf{u}_k^H) \boldsymbol{\beta}_\ell - \boldsymbol{\beta}_\ell^H \mathcal{L}(\mathbf{u}_j \mathbf{u}_k^H) \boldsymbol{\alpha}_\ell] \quad \text{for } \ell = 1, \dots, m \quad (48)$$

$$\boldsymbol{\alpha}_\ell = V_s \hat{\Sigma}_s^{-1} V_s^H \mathbf{a}(w_\ell), \quad \boldsymbol{\beta}_\ell = V_n V_n^H \hat{\mathbf{a}}(w_\ell) \quad (49)$$

$$\gamma_\ell = \hat{\mathbf{a}}^H(w_\ell) \boldsymbol{\beta}_\ell \quad (49)$$

$$\hat{\mathbf{a}}(w_\ell) = i \left. \frac{\partial \mathbf{a}(w)}{\partial w} \right|_{w=w_\ell} \quad (50)$$

and

$$\hat{\Sigma}_s = \text{diag}[\sigma_j - p_0]_{j=1, m}$$

hence

$$T = p_0 I_{M_\alpha} + V_s \hat{\Sigma}_s V_s^H. \quad (51)$$

In [25], this theorem is introduced in the more general form for an arbitrary (locally) differentiable matrix-valued transformation $\mathcal{L}(\hat{R})$ and for any of the orthogonal subspace (MUSIC, MIN-NORM, etc.) or signal subspace (ESPRIT, SSR, etc.) methods.

Note that by using the same technique used to prove this theorem, we can derive the asymptotic properties (with respect to N) of the eigenvalues of \bar{T}_0 . Consequently, we can comment on the asymptotic behavior of the model selection (detection) rules used to estimate a superior number of signals.

Let us now consider the results of DOA estimation simulations for the same signal model whose CRB analysis was presented in Fig. 1(b) using the MUSIC algorithm. Two representative SNR values (0 and 20 dB) have been chosen to demonstrate significantly different DAA performance. Fig. 5 shows the maximum DOA total error over the m sources for the estimator \bar{T}_0 (DAA only), compared with the CRB and

the asymptotic bound of DAA errors

$$\max_p \Theta_{pp} \quad \text{for } p = 1, \dots, m \quad (52)$$

calculated by (47).

In Fig. 5(a) (SNR = 0 dB), the results are in agreement with our expectations, namely, that DAA performance is close to the CRB for weak signals; moreover, even for conventional scenarios ($m = 2, 3, 4$), the accuracy delivered by DAA is practically the same as for MUSIC(\hat{R}). Obviously, for weak signals, there is no need to improve upon DAA. The theoretical prediction of the DAA asymptotic bound is in good agreement with the simulation results for \bar{T}_0 .

Fig. 5(b) corresponds to the same simulation but with SNR = 20 dB. Here, the difference between the accuracy of the DAA-only estimator \bar{T}_0 and the CRB is tremendous for most m ; only for $m = 9$ sources is the performance acceptable. It is clear that for strong signals (SNR $\gg 0$ dB), there is a great need to improve on the DAA estimator MUSIC(\bar{T}_0), especially for the superior case.

Fig. 5 also shows performance results for the estimator \bar{T}_m , which is defined by applying the alternating projections described in the previous section to the augmented covariance matrix $\hat{R} \xrightarrow{\text{DAA}} \bar{T}_0 \xrightarrow{\text{SST}} \bar{T}_m$. As expected, the accuracy improvement in the estimator \bar{T}_m is more pronounced for the stronger signal scenario. However, even at 20 dB SNR, the performance is significantly worse than the CRB suggests. We therefore choose to treat \bar{T}_m as an initial covariance estimate and perform a local ML search in the vicinity of these DOA estimates. The result of this ML optimization is denoted as \bar{T}_{ML} . Thus, we may write $\bar{T}_m \xrightarrow{\text{ML}} \bar{T}_{\text{ML}}$.

One possible iterative approach for this ML search stems from the following interpretation of the optimum ML DOA estimation algorithm for fluctuating signals, which is masked by a mixture of white noise and $(m - 1)$ interfering sources with known powers p_j and DOA's w_j ($j = 1, \dots, m; j \neq k$ th source) [26]. In that case, \hat{w}_{ML} is the argument of the solution \hat{w} of the equation

$$\frac{\text{Re } \hat{s}^H(w) R_k^{-1} \mathbf{s}(w)}{\mathbf{s}^H(w) R_k^{-1} \mathbf{s}(w)} - \frac{\text{Re } \sum_{t=1}^N \mathbf{y}^H(t) R_k^{-1} \mathbf{s}(w) \hat{s}^H(w) R_k^{-1} \mathbf{y}(t)}{\sum_{t=1}^N \mathbf{y}^H(t) R_k^{-1} \mathbf{s}(w)} = 0 \quad (53)$$

where

$$\hat{s}(w) = \frac{\partial \mathbf{s}(w)}{\partial w}, \quad R_k = p_0 I_M + \sum_{\substack{j=1 \\ j \neq k}}^m p_j \mathbf{s}(w_j) \mathbf{s}^H(w_j). \quad (54)$$

As mentioned above, the initial estimates $\hat{\mathbf{w}}^{(0)}$ and $\hat{\mathbf{p}}^{(0)}$ for this local refinement method are obtained from the matrix \bar{T}_0 .

Note that the standard first-order expansion $\hat{w}_{\text{ML}} \simeq \hat{w}_k^{(0)} + \hat{\Delta}_k$ with $\sin \pi d_M \hat{\Delta}_k \simeq \pi d_M \hat{\Delta}_k$ and $\cos \pi d_M \hat{\Delta}_k \simeq 1$ may be used to derive an analytic expression for the perturbation $\hat{\Delta}_k$. Unfortunately, we have found this expansion to be

insufficiently accurate in many cases, whereas the above one-dimensional (MUSIC-like) neighborhood search has proven to be more accurate.

The method of updating the power estimates $\hat{\mathbf{p}}^{(n)}$ at iteration n is via the minimization

$$\min_p \left\| \begin{bmatrix} \text{Re } \hat{R}_{jk} \\ \text{Im } \hat{R}_{jk} \end{bmatrix} - \begin{bmatrix} \text{Re } S(\hat{w}^{(n)}) \\ \text{Im } S(\hat{w}^{(n)}) \end{bmatrix} \hat{P} \right\|^2 \quad \text{for } j, k = 1, \dots, M. \quad (55)$$

In most cases with strong signals (only where this ML refinement is required), the initial power estimates $\hat{\mathbf{p}}^{(0)}$ inferred from \bar{T}_m are sufficiently accurate and need not be updated in the iteration.

To summarize, the entire proposed approach for superior DOA estimation may be defined as follows.

A. Augmentation-Truncation-ML (ATML) Superior-Case

Algorithm— $\hat{R} \xrightarrow{\text{DAA}} \bar{T}_0 \xrightarrow{\text{SST}} \bar{T}_m \xrightarrow{\text{ML}} \bar{T}_{\text{ML}}$

Step 1: Given the M -variate DDC matrix \hat{R} , apply the DAA algorithm of (44) to obtain the M_α -variate Toeplitz matrix \bar{T}_0 .

Step 2: Apply the SST algorithm to \bar{T}_0 to obtain the M_α -variate p.d. Toeplitz matrix \bar{T}_m with equal $(M_\alpha - m)$ smallest eigenvalues.

Step 3: Apply MUSIC or root-MUSIC to \bar{T}_m to obtain the m initial DOA estimates $\hat{\mathbf{w}}^{(0)}$. Define the m initial power estimates $\hat{\mathbf{p}}^{(0)}$ by (55).

Step 4: Apply the local ML search defined by (53) iteratively, using $\hat{\mathbf{w}}^{(0)}$ and $\hat{\mathbf{p}}^{(0)}$ as initial estimates, until appropriate convergence is met

$$\sum_{j=1}^m |\hat{w}_j^{(\kappa)} - \hat{w}_j^{(\kappa-1)}| < \xi \text{ CRB} \quad \text{with } \xi < 1. \quad (56)$$

In the trials presented here, the convergence tolerance was set to $\xi = 0.5$.

A further useful bound can be calculated by setting the DOA and power parameters for the $(m - 1)$ “other” (interfering) sources in (53) to be their true values; thus, the search procedure produces the exact ML DOA estimate for each source. Calculation of this bound, which is denoted ML1, enables us to assert that the limit accuracy of the CRB can be reached in practice. Obviously, this ML1 bound should always be somewhat below the CRB, calculated using unknown parameters for *all* sources.

Fig. 5(b) also shows the results for the ML1 bound and the performance of the \bar{T}_{ML} estimator; as usual, the maximum (worst) total error over all m sources is presented. MUSIC(\bar{T}_{ML}) approaches the corresponding CRB, whereas ML1 proves that the CRB limit accuracy is attainable. These results demonstrate the efficiency (in the ML sense) of our new technique.

V. SUMMARY AND CONCLUSIONS

The main point of this paper has been to clarify the conditions under which accurate DOA estimation (in the maximum-likelihood sense) is possible for a fully augmentable

M -element nonuniform linear array (NLA), given that the actual number of plane-wave sources m is known or has been algorithmically estimated. The sources are assumed to produce uncorrelated Gaussian processes with linearly independent manifold (steering) vectors.

A presented CRB analysis demonstrates the significant difference between the conventional ($1 \leq m < M_\alpha$) and superior ($M \leq m < M_\alpha$) scenarios for fully augmentable arrays with aperture ($M_\alpha - 1$), which is usually measured in half-wavelength units. The CRB in conventional scenarios displays a strong inverse semi-logarithmic dependence on SNR, where a lack in the number of snapshots N can be compensated (for $N \geq m$) by a corresponding increase in SNR. On the contrary, for the superior case with sufficiently strong signals ($\text{SNR} \gg 0$ dB) the CRB is practically independent of SNR, and the limit DOA estimation accuracy is essentially set by the sample volume N .

Correspondingly, it has been demonstrated that the problem of algorithmically approaching the CRB limit accuracy is completely different for the conventional and superior scenarios.

In the conventional case, standard subspace-based techniques (MUSIC and its variants) are asymptotically optimal in the ML sense and do not require improvement in the asymptotic domain. However, MUSIC performance degrades from the CRB in the preasymptotic domain, where MUSIC fails to resolve merged source peaks in the pseudo-spectrum and instead misidentifies some completely erroneous spurious peak (creating an ‘‘abnormal’’ DOA estimate). This well-known misidentification phenomenon is shown to be reduced in extent by a new method that involves specific properties of fully augmentable arrays and their covariance matrices.

The introduced method consists of two stages: MUSIC-maximum-entropy equalization (MMEE) followed by signal (eigen)subspace truncation (SST). MMEE takes the standard MUSIC pseudo-spectrum of the M -variate DDC matrix \hat{R} and constructs the ‘‘closest’’ M_α -variate p.d. Toeplitz matrix \hat{T}_0 with ME spectrum identical to MUSIC(\hat{R}). SST then modifies \hat{T}_0 to produce \hat{T}_m , which is another M_α -variate p.d. Toeplitz matrix, but with exactly $(M_\alpha - m)$ smallest eigenvalues equal in size. We write this process as $\hat{R} \xrightarrow{\text{MMEE}} \hat{T}_0 \xrightarrow{\text{SST}} \hat{T}_m$. The simulation results presented demonstrate a 16-dB reduction in threshold SNR at which the asymptotic domain begins (where the probability of abnormal estimates becomes negligible).

In the superior case, it is necessary to use an asymptotically large sample volume even to have a reasonable CRB; moreover, until now, the problem of algorithmically approaching this CRB has been unanswered. It has been demonstrated, both analytically and by simulation, that the well-known direct augmentation approach (DAA) is effective (in the ML sense) only for weak signals of $\text{SNR} \lesssim 0$ dB. For strong signals, DAA and its proposed variant DAA-SST still perform much worse than the corresponding CRB, even asymptotically. Thus, we have introduced a local ML optimization to act as a refining procedure for these DOA estimates. We call this overall algorithm ATML and denote the approach by $\hat{R} \xrightarrow{\text{DAA}} \bar{T}_0 \xrightarrow{\text{SST}} \bar{T}_m \xrightarrow{\text{ML}} \bar{T}_{\text{ML}}$. In all investigated cases, where the sample size has been chosen to guarantee the correct

resolution of all m sources at the initialization stage of DAA-SST, the local ML search has converged essentially to the limit accuracy defined by the CRB.

In general, the techniques introduced in this paper significantly expand the conditions under which the practically optimal (in the ML sense) DOA estimates can be inferred from a fully augmentable nonuniform linear array for both conventional and superior scenarios. Needless to say, the performance of nonuniform arrays is much better than for their corresponding M -element uniform arrays in the conventional case, while uniform arrays cannot resolve a superior number of sources at all.

APPENDIX A

PROOF OF LEMMA 1 [18]

The fact that (22) is a necessary condition for any vector to be expressed in the form of (20) is well known (see for example [21, ch. 6]). We will prove sufficiency. Let

$$z = e^{i\mu}, \quad f(\mu) = \frac{1}{|\hat{W}(e^{i\mu})|^2}. \quad (57)$$

Then, the function $f(\mu)$ is integrable and positive since $\hat{W}(z)$ fails to vanish on the circle $|z| = 1$, and therefore, $f(\mu)$ is a spectral density function of the Toeplitz correlation matrix \hat{T}_0 with the elements

$$\hat{t}_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-ik\mu} f(\mu) d\mu \quad \text{for } k = 0, \dots, M_\alpha - 1. \quad (58)$$

We must therefore prove that the vector $\hat{\mathbf{w}}$ that satisfies (22) may be written in the form of (20) and that we consequently have the orthogonality condition

$$\sum_{j=1}^{M_\alpha} \hat{t}_{k-j} \hat{w}_j = 0, \quad \text{for } k = 2, \dots, M_\alpha. \quad (59)$$

Denoting the left-hand side of this equation by \mathcal{J}_k and substituting (58), we have

$$\mathcal{J}_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-ik\mu} f(\mu) \sum_{j=1}^{M_\alpha} \hat{w}_j e^{-ij\mu} d\mu \quad \text{for } k = 2, \dots, M_\alpha. \quad (60)$$

Since, according to (21)

$$\sum_{j=1}^{M_\alpha} \hat{w}_j e^{i(j-1)\mu} = \hat{W}(e^{i\mu}) \quad (61)$$

we obtain

$$\mathcal{J}_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{e^{-i(k-1)\mu} \hat{W}(e^{i\mu})}{\hat{W}(e^{i\mu}) \hat{W}^*(e^{i\mu})} d\mu. \quad (62)$$

It now remains show that $\mathcal{J}_k = \mathcal{J}_k^* = 0$ for $k = 2, \dots, M_\alpha$. In fact

$$\mathcal{J}_k^* = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{e^{i(k-1)\mu}}{\hat{W}(e^{i\mu})} d\mu. \quad (63)$$

Therefore, by making the change of variable $z = e^{i\mu}$, this reduces to

$$\mathcal{J}_k^* = \frac{1}{2\pi k} \oint_{|z|=1} \frac{z^{k-2}}{\hat{W}(z)} \quad \text{for } k = 2, \dots, M_\alpha. \quad (64)$$

By virtue of (22), the function $\hat{W}^{-1}(z)$ is analytic within the unit circle, and so is $z^n \hat{W}^{-1}(z)$ for $n = 0, \dots, M_\alpha - 2$. Cauchy's theorem then gives us $\mathcal{J}_k = \mathcal{J}_k^* = 0$ for $k = 2, \dots, M_\alpha$, and the lemma is proven. ■

APPENDIX B

PROOF OF THEOREM 1 [22], [23]

Let us assume that DOA estimation errors are small enough to admit an accurate first-order expansion

$$\mathbf{a}(\hat{w}_j) \equiv \mathbf{a}(w_j + \Delta \hat{w}_j) = \{\exp[i\pi k(w_j + \Delta \hat{w}_j)]\}_{k=0, M_\alpha-1} \quad (65)$$

$$\simeq \{(1 + i\pi k \Delta \hat{w}_j) \exp(i\pi k w_j)\}_{k=0, M_\alpha-1} \quad (66)$$

and

$$\mathbf{a}(\hat{w}_j) = \mathbf{a}(w_j) + \Delta \hat{w}_j \left. \frac{\partial \mathbf{a}(w)}{\partial w} \right|_{w=w_j}. \quad (67)$$

From (50)

$$\hat{\mathbf{a}}(w_j) = [\pi k e^{i\pi k w_j}]_{k=0, M_\alpha-1}$$

and

$$\left. \frac{\partial \mathbf{a}(w)}{\partial w} \right|_{w=w_j} = -i \hat{\mathbf{a}}(w_j). \quad (68)$$

Now, we can rewrite the MUSIC pseudo-spectrum function $f_{\bar{T}}(w) = \mathbf{a}^H(w) \hat{V}_n \hat{V}_n^H \mathbf{a}(w)$ as

$$f_{\bar{T}}(w_j + \Delta \hat{w}_j) \simeq [\mathbf{a}^H(w_j) - i \Delta \hat{w}_j \hat{\mathbf{a}}^H(w_j)] \hat{V}_n \hat{V}_n^H \cdot [\mathbf{a}(w_j) + i \Delta \hat{w}_j \hat{\mathbf{a}}(w_j)] \quad (69)$$

$$= \mathbf{a}^H(w_j) \hat{V}_n \hat{V}_n^H \mathbf{a}(w_j) + \Delta \hat{w}_j^2 \hat{\mathbf{a}}^H(w_j) \hat{V}_n \hat{V}_n^H \hat{\mathbf{a}}(w_j) \quad (70)$$

$$- i \Delta \hat{w}_j \{ \hat{\mathbf{a}}^H(w_j) \hat{V}_n \hat{V}_n^H \hat{\mathbf{a}}(w_j) - \mathbf{a}^H(w_j) \hat{V}_n \hat{V}_n^H \hat{\mathbf{a}}(w_j) \}. \quad (71)$$

Since \hat{w}_j is a local maximum of the pseudo-spectrum $f_{\bar{T}}(w)$

$$\frac{\partial f_{\bar{T}}(w_j + \Delta \hat{w}_j)}{\partial (\Delta \hat{w}_j)} = 0 \quad (72)$$

and hence

$$\Delta \hat{w}_j = - \frac{\text{Im}\{\mathbf{a}^H(w_j) \hat{V}_n \hat{V}_n^H\} \hat{\mathbf{a}}(w_j)}{\hat{\mathbf{a}}^H(w_j) \hat{V}_n \hat{V}_n^H \hat{\mathbf{a}}(w_j)}. \quad (73)$$

Note that the essential perturbations for the noise subspace eigenvectors \hat{V}_n belong to the (true) signal subspace, i.e.,

$$\hat{V}_n = V_n + \Delta \hat{V}_n = V_n + V_s V_s^H \hat{V}_n. \quad (74)$$

For the true DOA's, we have $V_n^H \mathbf{a}(w_j) = 0$; therefore

$$\Delta \hat{w}_j \simeq - \frac{\text{Im}\{\mathbf{a}^H(w_j) V_s V_s^H \hat{V}_n V_n^H \hat{\mathbf{a}}(w_j)\}}{\hat{\mathbf{a}}^H(w_j) V_n V_n^H \hat{\mathbf{a}}(w_j)}. \quad (75)$$

The second-order terms are neglected in this equation. Eventually, the property of the projection $\Delta \hat{V}_n = V_s V_s^H \hat{V}_n$ defines

the DOA estimation asymptotic accuracy for the augmentation technique.

Let us now define asymptotic expressions for the signal and noise eigensubspace projectors as a function of the first-order perturbation of the augmented matrix \bar{T}_0 . Let

$$\Delta_T = \bar{T}_0 - T$$

$$\Delta_\Sigma = \hat{\Sigma} - \Sigma$$

and

$$\hat{\Sigma}_n = \hat{\Sigma}_n - p_0 I_{M_\alpha - m}. \quad (76)$$

Then, we have

$$\bar{T}_0(V_n + \Delta \hat{V}_n) = (V_n + \Delta \hat{V}_n)(p_0 I_{M_\alpha} + \hat{\Sigma}_n) \quad (77)$$

and neglecting second-order terms, we obtain

$$(T - p_0 I_{M_\alpha}) \Delta \hat{V}_n = -\Delta_T V_n + V_n \hat{\Sigma}_n. \quad (78)$$

Since

$$T - p_0 I_{M_\alpha} = V_s \hat{\Sigma}_s V_s^H \quad \text{and} \quad \Delta \hat{V}_n = V_s V_s^H \hat{V}_n \quad (79)$$

we can rewrite this as

$$V_s \hat{\Sigma}_s V_s^H V_s V_s^H \hat{V}_n = -\Delta_T V_n + V_n \hat{\Sigma}_n \quad (80)$$

and hence

$$V_s^H \hat{V}_n = -\hat{\Sigma}_s^{-1} V_s^H \Delta_T V_n \quad (81)$$

while

$$\Delta \hat{V}_n \equiv V_s V_s^H \hat{V}_n = -V_s \hat{\Sigma}_s^{-1} V_s^H \Delta_T V_n. \quad (82)$$

We may also introduce the asymptotic expression

$$\hat{\Sigma}_n = V_n^H \Delta_T V_n \quad (83)$$

which is important for the asymptotic efficiency analysis of any eigenvalue-based detection scheme [9]. Equation (82) substituted into (75) proves that the DOA errors are

$$\Delta \hat{w}_j = \frac{1}{\gamma_j} \text{Im}\{\boldsymbol{\alpha}_j^H \Delta_T \boldsymbol{\beta}_j\} \quad (84)$$

in the notation of (49).

To establish the statistical (asymptotic) properties of these errors

$$\Theta_{pq} = \mathcal{E}\{\Delta \hat{w}_p \Delta \hat{w}_q\} \quad \text{for } p, q = 1, \dots, m \quad (85)$$

recall the element-wise linear dependence based on the DAA

$$\bar{T}_0 = \mathcal{L}(\hat{R}). \quad (86)$$

For the matrix \hat{R} with complex Wishart distribution $N \hat{R} \sim \mathcal{CW}(N, M, R)$ and eigendecomposition $\hat{R} = \hat{U} \hat{\Lambda} \hat{U}^H$, we find

$$\hat{\Lambda} \sim \mathcal{CW}(N, M, \Lambda) \quad (87)$$

accordingly, $\Delta_T \equiv \mathcal{L}(\Delta_R)$, where $\Delta_R = \hat{U}\Delta_{\hat{\Lambda}}\hat{U}^H$, i.e.,

$$\Delta_R = \sum_{j,k=1}^M U c_j c_k^H U^H (\Delta_{\hat{\Lambda}})_{jk} \quad (88)$$

and

$$\Delta_T = \sum_{j,k=1}^M (\Delta_{\hat{\Lambda}})_{jk} \mathcal{L}(\mathbf{u}_j \mathbf{u}_k^H) \quad (89)$$

whence

$$\Delta \hat{w}_p = \frac{1}{\gamma_j} \text{Im} \left\{ \sum_{j,k=1}^M (\Delta_{\hat{\Lambda}})_{jk} \alpha_p^H \mathcal{L}(\mathbf{u}_j \mathbf{u}_k^H) \beta_p \right\}. \quad (90)$$

Denoting $\varepsilon_{jk}^p = \alpha_p^H \mathcal{L}(\mathbf{u}_j \mathbf{u}_k^H) \beta_p$, we have

$$\begin{aligned} \gamma_j \Delta \hat{w}_p &= \sum_{j=1}^M (\Delta_{\hat{\Lambda}})_{jj} \text{Im} \varepsilon_{jj}^p + \sum_{j < k} \\ &\quad \{ \text{Re}(\Delta_{\hat{\Lambda}})_{jk} \text{Im}(\varepsilon_{jk}^p + \varepsilon_{kj}^p) \\ &\quad + \text{Im}(\Delta_{\hat{\Lambda}})_{jk} \text{Re}(\varepsilon_{jk}^p - \varepsilon_{kj}^p) \} \end{aligned} \quad (91)$$

since $(\Delta_{\hat{\Lambda}})_{jk} = (\Delta_{\hat{\Lambda}})_{kj}^*$. The covariance values of Wishart-distributed matrix elements are defined via the fourth moment of the initial complex Gaussian samples [27]

$$\mathcal{E}\{(\Delta_{\hat{\Lambda}})_{jk} (\Delta_{\hat{\Lambda}})_{pq}^* \} = \Lambda_{jp} \Lambda_{kq}^*. \quad (92)$$

Since Λ is here a diagonal matrix

$$\mathcal{E}\{(\Delta_{\hat{\Lambda}})_{jk} (\Delta_{\hat{\Lambda}})_{pq}^* \} = \frac{1}{N} \lambda_j \lambda_k \delta_{jp} \delta_{kq} \quad (93)$$

and hence

$$\text{var}\{\text{Re}(\Delta_{\hat{\Lambda}})_{jk}\} = \text{var}\{\text{Im}(\Delta_{\hat{\Lambda}})_{jk}\} = \frac{\lambda_j \lambda_k}{2N}$$

and

$$\text{var}\{(\Delta_{\hat{\Lambda}})_{jk}\} = \frac{\lambda_j^2}{N} \quad (94)$$

and

$$\begin{aligned} \text{cov}\{\text{Re}(\Delta_{\hat{\Lambda}})_{jk} \text{Re}(\Delta_{\hat{\Lambda}})_{pq}\} &= 0 \\ &\quad \text{for } j \neq p, k \neq q \end{aligned} \quad (95)$$

$$\begin{aligned} \text{cov}\{\text{Re}(\Delta_{\hat{\Lambda}})_{jk} \text{Im}(\Delta_{\hat{\Lambda}})_{pq}\} &= 0 \\ &\quad \text{for } j, k, p, q = 1, \dots, M. \end{aligned} \quad (96)$$

From (94), we have

$$\begin{aligned} \gamma_p \gamma_q \text{cov}\{\hat{w}_p, \hat{w}_q\} &= \sum_{j=1}^M \text{Im} \varepsilon_{jj}^p \text{Im} \varepsilon_{jj}^q \frac{\lambda_j^2}{N} \\ &\quad + \frac{1}{2N} \sum_{j < k} \lambda_j \lambda_k \\ &\quad \{ [\text{Im}(\varepsilon_{jk}^p + \varepsilon_{kj}^p) \text{Im}(\varepsilon_{jk}^q + \varepsilon_{kj}^q) \\ &\quad + \text{Re}(\varepsilon_{jk}^p - \varepsilon_{kj}^p) \text{Re}(\varepsilon_{jk}^q - \varepsilon_{kj}^q)] \}. \end{aligned} \quad (98)$$

Note that the second component on the right-hand side can be written as

$$\frac{1}{2N} \sum_{j < k} \lambda_j \lambda_k \text{Re}\{(\varepsilon_{jk}^p - \varepsilon_{kj}^{p*})(\varepsilon_{jk}^q - \varepsilon_{kj}^{q*})\} \quad (99)$$

and

$$\varepsilon_{kj}^{p*} = \beta_p^H \mathcal{L}(\mathbf{u}_j \mathbf{u}_k^H) \alpha_p. \quad (100)$$

Since

$$\text{Im}(\varepsilon_{jj}^p) = \frac{1}{2i} (\varepsilon_{jj}^p - \varepsilon_{jj}^{p*}) \quad (101)$$

the first component can be written as

$$\frac{1}{2N} \sum_{j=1}^M \frac{\lambda_j^2}{2} (\varepsilon_{jj}^p - \varepsilon_{jj}^{p*})(\varepsilon_{jj}^q - \varepsilon_{jj}^{q*})^* \quad (102)$$

thus giving us (47). Furthermore, assuming that the asymptotic normality of the real and imaginary parts of the elements of the matrix $\Delta_{\hat{\Lambda}}$ obey the Central Limit Theorem, by (74) and the fact that $\lim_{N \rightarrow \infty} \mathcal{E}\{\Delta_{\hat{\Lambda}}\} = 0$, we finally have $\hat{w} \sim \mathcal{N}(m, W, \Theta)$ with Θ defined by (47). ■

REFERENCES

- [1] A. T. Moffet, "Minimum-redundancy linear arrays," *IEEE Trans. Antennas Propagat.*, vol. AP-16, pp. 172-175, 1968.
- [2] S. Haykin, J. P. Reilly, V. Kezys, and E. Vertatschitsch, "Some aspects of array signal processing," *Proc. Inst. Elect. Eng.*, pt. F, vol. 139, no. 1, pp. 1-26, 1992.
- [3] S. W. Lang, G. L. Duckworth, and J. H. McClellan, "Array design for MEM and MLM array processing," in *Proc. ICASSP*, 1981, pp. 145-148.
- [4] D. Pearson, S. U. Pillai, and Y. Lee, "An algorithm for near-optimal placement of sensor elements," *IEEE Trans. Inform. Theory*, vol. 36, pp. 1280-1284, 1990.
- [5] C. Chambers, T. C. Tozer, K. C. Sharman, and T. S. Durrani, "Temporal and spatial sampling influence on the estimates of superimposed narrowband signals: When less can mean more," *IEEE Trans. Signal Processing*, vol. 44, pp. 3085-3098, Dec. 1996.
- [6] S. Pillai, Y. Bar-Ness, and F. Haber, "A new approach to array geometry for improved spatial spectrum estimation," *Proc. IEEE*, vol. 73, pp. 1522-1524, Oct. 1985.
- [7] S. Pillai and F. Haber, "Statistical analysis of a high resolution spectrum estimator utilizing an augmented covariance matrix," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 35, pp. 1517-1523, Nov. 1987.
- [8] P. Stoica and A. Nehorai, "MUSIC, maximum likelihood and Cramér-Rao bound," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, pp. 720-741, May 1989.
- [9] M. Wax and T. Kailath, "Detection of signals by information theoretic criteria," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-33, pp. 387-392, 1985.
- [10] J.-J. Fuchs, "Estimation of the number of signals in the presence of unknown correlated sensor noise," *IEEE Trans. Signal Processing*, vol. 40, pp. 1053-1061, May 1992.
- [11] P. M. Djurić, "A model selection rule for sinusoids in white Gaussian noise," *IEEE Trans. Signal Processing*, vol. 44, pp. 1744-1757, July 1996.
- [12] J.-J. Fuchs, "Extension of the Pisarenko method to sparse linear arrays," in *Proc. ICASSP*, Detroit, MI, 1995, pp. 2100-2103.
- [13] Y. I. Abramovich, D. A. Gray, N. K. Spencer, and A. Y. Gorokhov, "Ambiguities in direction-of-arrival estimation for nonuniform linear antenna arrays," in *Proc. ISSPA*, Gold Coast, Australia, 1996, pp. 631-634.
- [14] Y. I. Abramovich, N. K. Spencer, and A. Y. Gorokhov, "Resolving manifold ambiguities in direction-of-arrival estimation for nonuniform linear antenna arrays," submitted for publication.
- [15] P. Stoica and A. Nehorai, "Performance study of conditional and unconditional direction-of-arrival estimation," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 38, pp. 1783-1795, Oct. 1990.
- [16] Y. I. Abramovich, D. A. Gray, A. Y. Gorokhov, and N. K. Spencer, "Comparison of DOA estimation performance for various types of sparse antenna array geometries," in *Proc. EUSIPCO*, Trieste, Italy, 1996, vol. 2, pp. 915-918.
- [17] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge U.K.: Cambridge Univ. Press, 1990.
- [18] K. M. Grigoriadis, A. E. Frazho, and R. E. Skelton, "Application of alternating convex projection methods for computation of positive definite Toeplitz matrices," *IEEE Trans. Signal Processing*, vol. 42, pp. 1873-1875, July 1994.
- [19] Y. I. Abramovich, D. Z. Arov, and V. G. Kachur, "Adaptive cancellation filters for stationary interference with a Toeplitz correlation matrix," *Soviet J. Commun. Technol. Electron.*, vol. 33, no. 4, pp. 54-61, 1988 (English translation from *Radiotekhnika i Elektronika*).

- [20] I. Gohberg and A. A. Semencul, "On the inversion of finite Toeplitz matrices and their continuous analogues," *Mat. Issled.*, vol. 7, no. 2, pp. 201–223, 1972 (in Russian).
- [21] S. Haykin, *Adaptive Filter Theory*, 3rd ed. Englewood Cliffs, NJ: Prentice-Hall, 1996.
- [22] U. Grenander and G. Szegő, *Toeplitz Forms and Their Applications*. Berkeley, CA: Univ. of Calif. Press, 1958.
- [23] P. Forster and G. Vezzosi, "Optical Toeplitzification with a given rank constraint," in *Proc. ICASSP*, 1989, pp. 2783–2786.
- [24] Y. I. Abramovich and A. Y. Gorokhov, "Improved analysis of high resolution spatial spectrum estimators in minimum redundancy linear arrays," in *Proc. RADAR*, Paris, France, 1994, pp. 127–132.
- [25] A. Y. Gorokhov, Y. I. Abramovich, and J. F. Böhme, "Unified analysis of DOA estimation algorithms for covariance matrix transforms," *Signal Process.*, vol. 55, pp. 107–115, 1996.
- [26] R. C. Davis, L. E. Brennan, and I. S. Reed, "Angle estimation with adaptive arrays in external noise fields," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-12, pp. 179–186, 1976.
- [27] M. Siotani, T. Hayakawa, and Y. Fujikoshi, *Modern Multivariate Statistical Analysis*. Columbus, OH: Amer. Sci., 1985.



Yuri I. Abramovich (M'96) received the Dip. Eng. (Hons.) degree in radio electronics in 1967 and the Candidate of Science degree (Ph.D. equivalent) in theoretic radio techniques in 1971, both from the Odessa Polytechnic University, Odessa, U.S.S.R. He received the D.Sc. degree in radar and navigation from the Leningrad Institute for Avionics, Leningrad, U.S.S.R. in 1981.

From 1968 to 1994, he was with the Odessa State Polytechnic University, Odessa, Ukraine, as a Research Fellow, Professor, and finally as Vice

Chancellor of Science and Research. Since 1994, he has been at the Cooperative Research Centre for Sensor Signal and Information Processing, Adelaide, Australia, as a Visiting Professor. His research interests are in the application of signal processing to radar, particularly over-the-horizon radar, electronic warfare and communication, particularly spatio-temporal adaptive processing, beamforming, signal detection, and estimation.



Douglas A. Gray (M'82) received the B.Sc. and Ph.D. degrees in mathematical physics from the University of Adelaide, Adelaide, Australia, in 1969 and 1974, respectively.

From 1973 to 1993, he was with the Australian Defence Science and Technology Organization, Adelaide, applying signal processing to sonar and electronic warfare. From 1977 to 1979, he was a Visiting Scientist at the Royal Aircraft Establishment, and in 1985, he was a Visiting Fellow at the Australian National University,

Canberra. He is currently Professor of Sensor Signal Processing at the University of Adelaide and Deputy Director of the Cooperative Research Centre for Sensor Signal and Information Processing. His research interests are in the application of signal processing to sonar, radar, and electronic warfare, particularly in adaptive processes, beamforming, signal sorting, and classification techniques.



Alexei Y. Gorokhov (M'93) received the Dip. Eng. (Hons.) degree in radio electronics in 1993 from the Odessa State Polytechnic University, Odessa, Ukraine. From 1993 to 1994, he was with the Scientific Research Department of the same university. From 1994 to 1997, he was preparing the Ph.D. degree in signal processing at Télécom Paris, Paris, France.

He is currently a Research Fellow at Centre Nationale de la Recherche Scientifique (CNRS), Laboratoire des Signaux et Systèmes (L2S), Paris, France. His research interests are in adaptive signal processing, blind signal separation, and equalization for radar, electronic warfare, and communications applications.



Nicholas K. Spencer received the B.Sc. (Hons.) degree in applied mathematics in 1985 and the M.Sc. degree in computational mathematics in 1992, both from the Australian National University, Canberra.

His past employment includes periods at the Department of Defence, the Flinders University of South Australia, Adelaide, the University of Adelaide, and the Australian Centre for Remote Sensing, in areas concerning the computing and mathematical sciences. He is currently a Research Fellow at the Cooperative Research Centre for Sensor Signal and Information Processing (CSSIP), Adelaide. His research interests include array signal processing, parallel and supercomputing, numerical solution of partial differential equations, theoretical astrophysics, and cellular automata.