Abstract—In this paper, the effect of using a spatially smoothed forward–backward covariance matrix on the performance of weighted eigen-based state space methods/ESPRIT, and weighted MUSIC for the direction-of-arrival (DOA) estimation is analyzed. Expressions for the mean-squared error in the estimates of the signal zeros and the DOA estimates, along with some general properties of the estimates and optimal weighting matrices are derived. A key result of the analysis is that optimally weighted MUSIC and weighted state space methods/ESPRIT have identical asymptotic performance. It is also shown that by properly choosing the number of subarrays, the performance of unweighted state space methods can be significantly improved. Then it is shown that the mean-squared error in the DOA estimates obtained using subspace based methods is independent of the exact distribution of the source amplitudes. This results in an unified framework for dealing with DOA estimation using a uniformly spaced linear sensor array (ULA), and the time series frequency estimation problem. The resulting analysis of the time series case is shown to be more accurate than previous results.

I. INTRODUCTION

In this paper, weighted subspace based methods for estimating the direction of arrival (DOA) of plane waves in noise using a uniformly spaced linear array (ULA) are considered. In particular, the effect of using a spatially smoothed forward-backward covariance matrix on the statistical performance of weighted eigen-based state space based methods [1] or ESPRIT [2] and weighted MUSIC is analyzed. An examination of unweighted MUSIC method under these conditions has recently been studied in [3]–[6], and hence in this paper there is greater emphasis on the weighted state space methods/ESPRIT. ESPRIT was developed for general arrays which possess a displacement invariance [2] and reduces to a state space method for the ULA case [1], [7], [8]. However, the state space methods were developed first and deal extensively with the ULA case, the problem of interest here [1], [8]–[10]. In particular, the use of spatial smoothing and the forward–backward approach in the manner developed here were first discussed in the context of state space methods.

Hence, the use of the term state space methods appears more appropriate and will be used in this paper.

In recent years, a statistical evaluation of subspace based methods has been conducted by a number of researchers [3], [4], [8], [11]–[19]. Particular attention has been paid to the case where the estimate of the covariance matrix was obtained by a straight forward averaging of the outer product of the snapshots. In [15], [17], [3], [6], it was shown that MUSIC performs better than eigen-based state space methods with the difference becoming significant as the length of the array increases. However, state space methods are computationally more attractive and there is incentive to seek ways to improve the statistical properties. Weighted eigen-based state space methods and weighted MUSIC are analyzed, and it is shown that this deficiency can be overcome by using a spatially smoothed forward–backward covariance matrix and proper weighting. These results are consistent with the past successful uses of the forward–backward approach, first for autoregressive parameter estimation [20], and later for the problem of sinusoid frequency estimation [21], [9], and for DOA estimation of coherent sources [22]–[24].

Another important attribute of the forward–backward smoothing approach is that it enables a natural link between the problem of narrow-band DOA estimation using an ULA, and the time series frequency estimation problem, i.e., sinusoids in noise problem. The natural commonality between these problems is well known, however, unfortunately the analysis of these problems has been treated separately. This paper capitalizes on the analysis framework developed in [3], [5], and establishes an unified framework for these problems.

The outline of the paper is as follows. Section II provides some background information. The various covariance matrix estimators are discussed and the framework for analysis is outlined. Amongst the covariance estimators, it is shown that combining spatial smoothing with the forward–backward approach is more effective than using forward spatial smoothing alone. The statistics relevant to the analysis are derived under the assumption that the source amplitudes and the noise are independent and are circularly Gaussian random vectors.

In Section III, an expression for the mean-squared error (MSE) in the DOA estimates obtained using eigen-based state space methods is derived. Some interesting proper-
ties of the estimator along with the optimal weighting issue are discussed.

In Section IV, an expression for the mean-squared error in the DOA estimates obtained using weighted MUSIC are presented. Its properties along with the optimal weighting issue are discussed. A key result of the analysis is that optimally weighted MUSIC and weighted state space methods/ESPRIT have identical asymptotic performance.

In Section V, the one and two source problem is examined in detail for the weighted state space method. Along with the issue of optimal choice of subarrays in the unweighted case, the effect of optimal weighting on the performance is examined, and shown to significantly improve the performance of state-space methods.

In Section VI, it is shown that the relevant statistics used in the analysis are independent of the distribution of the source amplitudes. This provides a mechanism to extend the results to the time series frequency estimation problem, i.e., to deal with finite number of snapshots and deterministic source amplitudes, and results in a unified framework for dealing with DOA estimation using ULA's, and the time series frequency estimation problem. The ensuing results in the time series case are shown to be more general than existing results.

In Section VII, some computer simulations are provided to support the theoretical observations. Some of these results were first presented in [25]-[27].

II. Problem Formulation
A. Data Model
For the nth observation period (snapshot), the spatial samples of the signal plus noise when M, possibly coherent, plane waves are incident on a ULA of L1 sensors is given by the output vector

\[ Y(n) = \sum_{i=1}^{M} p_i(n)e^{i\omega_i} + N(n) \]

where \( \omega_i = (2\pi d/\lambda) \sin \theta_i \), \( d \) being the separation between sensors, \( \lambda \) the wavelength of the incident signal, and \( \theta_i \) the DOA. Due to the simple mapping between \( \omega_i \) and \( \theta_i \), for simplicity, \( \omega_i \) will also be often referred to as the DOA.

B. Different Covariance Matrix Estimators
Exploiting the ULA structure of the array, different covariance matrix estimates are possible from the data. Some popular estimates are discussed next.

1) Forward Only Smoothing (FS) Approach: In this technique, the array is divided into \( K \) smaller subarrays containing \( L_K \) sensors each, where \( K = L_1 - L_K + 1 \). The covariance matrices of the subarrays are determined and then averaged to obtain the smoothed estimate of \( R_{fs} \) [22]-[24], i.e.,

\[ \hat{R}_{fs} = \frac{1}{K} \sum_{p=1}^{K} \hat{R}_{fs}^p \]

and \( Y_p(n) \) is the output vector of the \( p \)th subarray. The special case where \( K = 1 \) has been extensively studied [11]-[14], [17], [16], and is referred to here as the forward only (FO) approach.

2) Forward-Backward Smoothing (FBS) Approach: Combining the smoothing approach with the forward-backward approach, the general FBS estimate is given by [22]-[24]

\[ \hat{R}_{fbs} = \frac{\hat{R}_{fs} + J\hat{R}_{fbs}J}{2} \]

where \( J = \begin{bmatrix} 0 & 0 & 1 \\ \vdots & \vdots & \vdots \\ 1 & 0 & 0 \end{bmatrix} \)

The exchange matrix \( J \) satisfies \( J = J^T \) and \( J^2 = I \). The FBS approach for the special case of \( K = 1 \) is called the forward-backward (FB) approach. An important property of \( R_{fbs} \) and \( \hat{R}_{fbs} \) is that they are centrohermitian, i.e.,

\[ R_{fbs} = JR_{fbs}^J \text{ and } \hat{R}_{fbs} = J\hat{R}_{fbs}^*J. \]

Compared to the FS approach, the FBS approach has been found to result in superior estimates. A potential explanation for this observation is provided in the following theorem whose proof is presented in Appendix A.

Theorem 1: For a given number of subarrays \( K \), in the noise free case the FBS covariance matrix is better conditioned than the FS covariance matrix, i.e.,

\[ (\lambda_1)_{fbs} \leq (\lambda_1)_{fs} \text{ and } (\lambda_M)_{fbs} \geq (\lambda_M)_{fs} \]

resulting in

\[ \kappa(R_{fbs}) \leq \kappa(R_{fs}) \]

where \( \kappa(R) = ||R||_2 ||R^+||_2 \) and hence the result is still true. \( \lambda_M \) plays an important role in low SNR and high resolution cases [17], [29]. A larger \( \lambda_M \) implies that the quality of the estimated signal subspace is better [28], thereby indicating that the FBS approach is preferable to the FS approach.

C. Subspace Decomposition
The subspace based methods estimate the DOA using the eigenspaces of the covariance matrix \( R \). Given any
one of the exact covariance matrices, its eigendecomposition is given by

\[ R = E \Lambda E^H + \sigma^2 I, \]

where \( E = \{E_1, E_2, \ldots, E_n\}, \) and \( \Lambda = \{ \lambda_1, \lambda_2, \ldots, \lambda_n \}. \)

Also

\[ \lambda_1 = \lambda_1^2 + \sigma^2_1 \geq \lambda_2 = \lambda_2^2 + \sigma^2_2 \geq \ldots \]

\[ \lambda_n = \lambda_n^2 + \sigma^2_n \geq \lambda_{n-1} = \lambda_{n-1}^2 + \sigma^2_{n-1} = \ldots = \lambda_2 = \lambda_2^2. \]

**Remark:** \( \lambda_i, \) the eigenvectors, and eigenvalues depend on the covariance approach used, and are also a function of the number of subarrays \( K. \) However, for simplicity of notation this is not explicitly indicated in the notation and should be clear from the context. Also \( L \) is used instead of \( L_K \) for simplicity.

**D. Performance Analysis of Subspace Methods**

For analysis purposes it is useful to divide the subspace methods into three steps as shown in Fig. 1. Step 3 is dependent on the subspace approach used to estimate the DOA, and the weighted subspace methods are the focus of this paper. Steps 1 and 2 are common to subspace methods. The issues involved in steps 1 and 2 have been discussed when MUSIC and the Minimum-Norm method were analyzed in [5], [6], and so only the main results are summarized here. For the proofs, the reader is referred to [5], [6].

Step 1 involves the mapping from the data to the covariance matrix estimate, and the statistics of the covariance matrix estimate are necessary for the analysis. This is summarized in Theorem 2. Step 2 involves the computing of the eigenvectors from the covariance matrix estimate. Errors in the covariance matrix affect the subspaces, and this effect needs to be quantified in order to carry out the analysis. This is done in Theorem 3. It is useful to note that previous analyses of the subspace methods were based on the statistics of the eigenvectors [11]–[14], [17]. In this paper, though the statistics of the eigenvectors are not computed explicitly, Theorems 2 and 3 combined provide the necessary statistical characterization of the error in the subspaces. The use of Theorems 2 and 3 greatly enhances the tractability of the analysis, and the results of this paper should support this.

![Fig. 1. The various steps in the subspace methods are shown.](image)

1) **Statistics of the Covariance Estimates:** The statistics in connection with the covariance matrix estimate \( \hat{R} \) which is necessary for the analysis is \( \Gamma_{\theta \gamma} \), is defined as

\[ \Gamma_{\theta \gamma} = \text{cov} [(\alpha^H \hat{R} \delta), (\beta^H \hat{R} \gamma)] \]

\[ = \alpha^H B_{\theta \gamma} \delta = \alpha^H B_{\theta \gamma} \beta \]

2) **Perturbation in the Subspaces:** The error in the estimated signal subspace eigenvectors projected onto the noise subspace is given by

\[ P_n \Delta \lambda_{E_3} = P_n \Delta R R_n^H E_3 = P_n \hat{R} R_n^H E_3 \]

where the projection matrix on the noise subspace is denoted by \( P_n \), with \( P_n = E_n E_n^H \). Also \( \hat{R} = R + \Delta R \), and \( \hat{E}_3 = E_3 + \Delta E_3 \).

\[ \text{cov} (X, Y) = (X - \bar{X})(Y - \bar{Y})^H. \]

\[ \text{In this paper, the overbar '·'·'·' will be used to denote the expectation operator.} \]

\[ \text{For simplicity of notation, the same } \Gamma \text{ is used to denote the statistics both for the FS and FBS case. From the context it should be clear which one to use.} \]
A similar result can be obtained in connection with the noise subspace eigenvectors [5].

III. ANALYSIS OF EIGEN-BASED STATE SPACE METHODS

In this section we analyze the eigen-based state space method. These results are an extension and improvement of the results in [8], [15] where the forward only case is examined. Therefore, the results in [8], [15] will be used as a starting point for the analysis.

A. Background

The array manifold is spanned by the columns of $A$, where

$$A = [V(\omega_1), V(\omega_2), \ldots, V(\omega_M)]$$

$$= \begin{bmatrix}
a_1 \\
a_2 \\
\vdots \\
a_L \\
a_M
\end{bmatrix} = A_1 = \begin{bmatrix}
a_1 \\
a_1 \psi \\
\vdots \\
a_1 \psi^{L-1}
\end{bmatrix}$$

(9)

with $V(\omega) = (1/\sqrt{L})[1, e^{j\omega_1}, e^{j\omega_2}, \ldots, e^{j(L-1)\omega}]^T$ and $\psi$ is a diagonal matrix with the diagonal entries equal to $z_i = e^{j\omega_i}$, i.e., $\psi = \text{diag}(z_i)$. Since the signal eigenvectors span the same space as the columns of $A$, there exists a unique, nonsingular matrix $Q$ such that

$$E_s = AQ \quad \text{or} \quad A = E_sQ^{-1}.$$  

(10)

This implies that $E_s$ also has a shift structure like $A$ and this is exploited by the eigen-based state space method and ESPRIT, i.e.,

$$E_s = \begin{bmatrix}
e_1 \\
e_2 \\
\vdots \\
e_L
\end{bmatrix} = \begin{bmatrix}
E_1 \\
E_2 \\
\vdots \\
E_L
\end{bmatrix} = \begin{bmatrix}
h_x \\
h_F \psi \\
\vdots \\
h_F \psi^{L-1}
\end{bmatrix}$$

(11)

where $e_i$ is the $i$th row of $E_s$ and $E_1, E_2$ are appropriate $(L - 1) \times M$ matrices. From (9)-(11), $h_x = a_1Q$ and $F$ is an $M \times M$ matrix which is equal to

$$F_s = Q^{-1}\psi Q.$$  

Hence, the eigenvalues of $F_s$ are equal to $z_i = e^{j\omega_i}$. Therefore, the DOA can be determined by obtaining $F_s$ and then its eigenvalues. From the structure in $E_s$, it is clear that many procedures to estimate $F_s$ can be devised. Here we only consider the procedure which exploits the following linear relationship:

$$E_1F_s = E_2.$$  

(12)

In the presence of errors, there is no exact solution to (12) and so a least squares or total least squares procedure can be used to estimate $F_s$. In [8], it was shown that the DOA estimates obtained from both these methods were asymptotically equivalent. Hence, we only consider the generalized version of the least squares estimate, the weighted least squares estimate, which is given by [15]

$$\hat{F}_s = (\hat{E}_s^H\hat{W}E_s)^{-1}\hat{E}_s^H\hat{W}E_2$$

where $W$ is a Hermitian symmetric, positive semidefinite matrix. Note that simple scalar scaling of $W$ does not change the estimate, and hence we will not differentiate weighting matrices that are related by such scaling.

B. Mean Squared Error (MSE) in Signal Zero and DOA

The analysis begins with an expression for $\Delta z_i$, the first-order perturbation in the eigenvalues of $F_s$, which can be shown to be given by [15], [8]

$$\Delta z_i = z_i e^{H} \Delta E_s x_i$$

(13)

where

$$e_i^H = -i e^{H}(A^H W A)_i^{-1} A^H W T$$

(14)

$$T = [I_{L-1 \times L-1}; 0_{L-1 \times L-1}] - z_i [0_{L-1 \times L-1}; I_{L-1 \times L-1}]$$

(15)

$l_i$ is the $i$th column of $L \times M$, and $x_i$ is the $i$th column of $Q^{-1}$ given in (10). Note that

$$Q^{-1} = [x_1, x_2, \ldots, x_M] = E_s^H A$$

or

$$x_i = Q^{-1} l_i = E_s^H A l_i = E_s^H V(\omega).$$

(16)

Equation (16) follows from (10). An interesting and useful property of $e_i$ is that it is in the noise subspace [15], i.e.,

$$e_i^H V(\omega_m) = 0, \quad m = 1, \ldots, M, \quad \text{and} \quad e_i^H P_s = e_i^H.$$

(17)

Using (17) and (16) in (13), we have

$$\Delta z_i = z_i e_i^H P_s \Delta E_s e_i^H V(\omega).$$

(18)

Intuitively, the above expression is appealing as it indicates that the error in the estimate depends only on $P_s \Delta E_s$, the component of the error projected onto the noise subspace. This is the counterpart of the observation in the context of noise subspace methods that the component of the errors in the noise subspace eigenvectors projected onto the signal subspace determine the error in the DOA estimates [3], [14]. Using (8) in (18),

$$\Delta z_i = z_i e_i^H R_i \tilde{R}_i \gamma(\omega) = z_i e_i^H \tilde{R}_i,$$

(19)

where $\tilde{R}_i = R_i^H V(\omega_i)$ and is a vector in the signal subspace. Taking expectations and using Theorem 2, we have

$$|\Delta z_i|^2 = (e_i^H \tilde{R}_i) (e_i^H \tilde{R}_i)^* = \Gamma_{en\beta i\beta}.$$  

Since (29)

$$\langle \Delta \omega \rangle^2 = \frac{1}{L} \left\{ |\Delta z_i|^2 - \text{Re} (z_i^H \tilde{R}_i) (z_i^H \tilde{R}_i)^* \right\}$$

(20)
one needs to determine \((\Delta z_i)^2\). Using (19),

\[
(\Delta z_i)^2 = z_i^*(\epsilon_i^H R_{\beta_i})(\epsilon_i^H R_{\beta_i})^* \quad z_i^* R_{\beta_i} = z_i^2 \Gamma_{\beta_i,\beta_i}.
\]

Substituting in (20), the MSE in the DOA estimate is given by

\[
(\Delta \omega_i)^2 = \frac{1}{2} \text{Re} \left( \Gamma_{\omega,\omega_i} - \Gamma_{\epsilon,\epsilon_i} \right). \tag{21}
\]

Similarly it can be shown that \(\Delta \omega_i \Delta \omega_j = \frac{1}{2} \text{Re} \left( \Gamma_{\epsilon,\epsilon_i} - \Gamma_{\epsilon,\epsilon_j} \right)\). The MSE in the estimate of \(\theta_i\) can be computed as

\[
(\Delta \theta_i)^2 = \left( \frac{\lambda}{2 \pi \cos \theta_i} \right) (\Delta \omega_i)^2. \tag{22}
\]

### C. Properties of the DOA Estimates

Due to the simple relationship between \((\Delta \theta_i)^2\) and \((\Delta \omega_i)^2\), only \((\Delta \omega_i)^2\) is examined, and it is shown to have some interesting properties \[27\], \[26\].

**Property 1:** For the forward only (FO) approach,

\[
(\Delta \omega_i)^2 = \frac{|\Delta z_i|^2}{2}.
\]

**Proof:** From (7) setting \(K = 1\), for the forward only (FO) approach

\[
\Gamma_{\epsilon,\epsilon_i} = \frac{1}{N} (\epsilon_i^H R_{\beta_i})(\beta_i^H R_{\epsilon_i}) = \frac{1}{N} |\epsilon_i^H R_{\beta_i}|^2 = \frac{1}{N} |\sigma_i^2 \epsilon_i^H \beta_i|^2 = 0.
\]

The last equality is a result of the fact that \(\epsilon_i\) is in the noise subspace, and hence a noise eigenvector orthogonal to \(\beta_i\) which lies in the signal subspace. Thus the result follows from (21) and (20).

**Property 2:** For the FBS approach using a centrohermitian weighting matrix, i.e., \(W = J W^* J\),

\[
(\Delta \omega_i)^2_{\text{FBS}} = |\Delta z_i|^2_{\text{FBS}}
\]

and therefore

\[
(\Delta \omega_i)^2_{\text{FBS}} = |\Delta z_i|^2_{\text{FBS}} = \Gamma_{\epsilon,\epsilon_i}.
\]

with

\[
\Gamma_{\epsilon,\epsilon_i} = \frac{1}{2NK^2} \left( \sum_{p=1}^{K} \sum_{q=1}^{K} (\beta_i^H R_{p\beta_i})(\epsilon_i^H N_{q\epsilon_i}) ight) \]

\[
+ \sum_{p=1}^{K} \sum_{q=1}^{K} 2 |\beta_i^H N_{pq\epsilon_i}|^2 \right) \quad \tag{26}
\]

Where

\[
B_{\beta_i} = \frac{1}{2NK^2} \left( \sum_{p=1}^{K} \sum_{q=1}^{K} (\beta_i^H R_{p\beta_i}) N_{pq} ight) + \sum_{p=1}^{K} \sum_{q=1}^{K} 2 |N_{pq\beta_i}|^2). \tag{25}
\]

and \(R_{pq} = Y_p(n) Y_q^H(n)\) and \(N_{pq} = P_p R_{pq} P_q = N_p(n) N_q^H(n)\).

**Remark:** The proof of this property is presented in Appendix C. Note the unweighted least squares corresponds to the case where \(W = I\), and hence satisfies the above property. Equation (23) results in a MSE expression considerably simpler than (21). Also from (23) one can infer that the estimated signal zeros will lie on the tangent to the line joining the origin and the true zero location in the \(z\) plane. The second term in (24) depends on the square of the noise-to-signal ratio while the first term depends on the noise-to-signal ratio and its square. In high SNR conditions, the second term can be neglected.

**Property 3:** For the FS approach,

\[
(\Delta \omega_i)^2_{\text{FS}} = \frac{1}{2NK^2} \left( \sum_{p=1}^{K} \sum_{q=1}^{K} (\beta_i^H R_{p\beta_i})(\epsilon_i^H N_{q\epsilon_i}) ight) \]

\[
- \text{Re} \left( \sum_{p=1}^{K} \sum_{q=1}^{K} 2 (\beta_i^H N_{pq\epsilon_i})(\beta_i^H N_{q\epsilon_i}) \right). \tag{26}
\]

**Proof:** The result can be easily obtained as in Property 2, starting from (21) and (7).

**Remark:** The above equation can also be used to compute the MSE in DOA for the FBS approach by using the corresponding \(\beta_i\). This is true because for the FBS approach (follows from Appendix C by using (48) and (49)).

\[
\text{Re} \left( \sum_{p=1}^{K} \sum_{q=1}^{K} 2 (\beta_i^H N_{pq\epsilon_i})(\beta_i^H N_{q\epsilon_i}) \right) \]

\[
= 2 \sum_{p=1}^{K} \sum_{q=1}^{K} \sum_{q \neq p} |\beta_i^H N_{pq\epsilon_i}|^2. \]

**Property 4:** For the FS approach, for small \(K\) or high SNR

\[
(\Delta \omega_i)^2_{\text{FS}} \approx \frac{1}{2} |\Delta z_i|^2_{\text{FS}}.
\]

**Proof:** Follows from Property 3 and (20) by neglecting the second term (26).

**Property 5:** For uncorrelated sources, when a centrohermitian weighting matrix is used

\[
(\Delta \omega_i)^2_{\text{FBS}} = (\Delta \omega_i)^2_{\text{FS}}.
\]

**Proof:** Follows from remark following Property 3 since in the uncorrelated case \(R_{\beta_i} = R_{\beta_i}\) and hence the \(\beta_i\) are equal.

### D. Optimum Weighting Matrix

Here we concentrate on the FBS approach because of its apparent superiority over the FS approach. The weighting matrix is restricted to be centrohermitian as this enables using Property 2, and hence requires minimization of only the error in the signal zeros simplifying the optimization process.
Theorem 4: For the FBS approach, the optimum centrohermitian weighting matrix \( W_{\text{opt}} \) for weighted state space method/ESPRIT which minimizes the MSE in a given DOA is given by
\[
W_{\text{opt}} = (TB_{\text{f,b}}T^H)^{-1}.
\]
The MSE corresponding to \( W_{\text{opt}} \) is then given by
\[
(\Delta \omega_j)_{\text{opt}}^2 = l_j^H (A_j^H (TB_{\text{f,b}}T^H)^{-1}A_j) l_j.
\]

Proof: Only an outline of the proof is provided here. The proof that \( W_{\text{opt}} \) minimizes \( \Gamma_{\mu \delta \mu \delta} \) can be shown in a manner similar to that in [15]. The fact that \( W_{\text{opt}} \) is centrohermitian can be shown using arguments similar to those used to establish Property 2.
By virtue of Property 1, the optimum matrix for the forward only case is also given by (27).

Property 6: The optimum weighting matrix for the FO and FB approaches is the same and is given by
\[
W_{\text{opt}} = (TT^H)^{-1}.
\]

Proof: Corresponding to FO and FB estimates, \( B_{\text{f,b}} \approx I \). Since scaling does not affect the MSE, substituting \( I \) for \( B_{\text{f,b}} \) in (27), the optimum matrix for FO and FB approaches is given by (29). The result corresponding to the FO approach was first shown in [15].

IV. COMPARISON WITH WEIGHTED MUSIC

A. Weighted MUSIC

In spectral weighted MUSIC, the DOA are found by locating the peaks of the spatial spectrum \( S(\omega) = 1/D(e^{j\omega}) \), where \( D(e^{j\omega}) \), the weighted null spectrum, is defined as
\[
D(e^{j\omega}) = V^H(\omega)P_nWP_nV(\omega)
\]
\( W \) being an \( L \times L \) matrix. \( W = I \) corresponds to the usual MUSIC procedure. Again simple scalar scaling of the weighting matrix does not change the DOA estimate. In [31], a slightly different form of weighted MUSIC was used. The null spectrum was defined as \( D(e^{j\omega}) = V^H(\omega)E_nWE_n^H V(\omega) \). For a given \( W \), the same null spectrum is obtained by choosing \( W = E_nWE_n^H \). Hence (30) is more general, and turns out to be more convenient for our purposes.

B. Analysis of Weighted MUSIC

Based on the results presented in connection with the weighted state space methods, and previous work on unweighted MUSIC [5], [3], [31], the analysis and properties of the weighted MUSIC algorithm can be readily derived [25]. Hence most of the results are presented without proof except for Property 7.

Theorem 5: The statistics of the error in the DOA for weighted MUSIC are given by
\[
(\Delta \omega_j)_{\text{opt}}^2 = \frac{\Re \Gamma_{\mu \delta \mu \delta}}{2(V^H(\omega)P_nWP_nV(\omega))^2}(\Delta \omega_j)_{\text{opt}}^2
\]

\[
(\Delta \omega_j)_{\text{opt}}^2 = \frac{\Re \Gamma_{\mu \delta \mu \delta}}{(V^H(\omega)P_nWP_nV(\omega))^2}
\]

with \( V_1(\omega) = (\partial V(\omega)/\partial \omega), \mu_i^H = V^H(\omega_i)P_nWP_n, \) and \( \beta_i = R^+V(\omega_i) \).

Property 7: For the FBS approach, the performance obtained using any symmetric positive semidefinite matrix \( W \) can be obtained with a centrohermitian matrix \( W_1 \)

\[
W_{\text{opt}} = (P,B_{\text{f,b}}P_n)^{+}
\]

The optimum weighting matrix for the FBS approach, which minimizes the MSE in a given DOA, is stated next.

Theorem 6: Given the FBS covariance matrix estimate, an optimum weighting matrix \( W_{\text{opt}} \) for weighted MUSIC is given by

\[
(\Delta \omega_j)_{\text{opt}}^2 = V^H(\omega)P_nWP_nV(\omega)
\]

\[
= \frac{\Re \Gamma_{\mu \delta \mu \delta}}{(V^H(\omega)P_nWP_nV(\omega))^2}
\]

\[
W_{\text{opt}} = (P,B_{\text{f,b}}P_n)^{+}
\]

The MSE in DOA corresponding to \( W_{\text{opt}} \) is given by

\[
(\Delta \omega_j)_{\text{opt}}^2 = \frac{1}{V^H(\omega)P_nWP_nV(\omega)}
\]

Remark: For the forward only estimate, the MSE in DOA has a form similar to (33), and hence the result is also valid for the FO approach.

Property 9: The optimum weighting matrix is the same for both FO and FB approaches.

Remark: For the FO and FB cases, \( W_{\text{opt}} = P_n^+P_n \). This choice of \( W \) is equivalent to the choice of \( W = I \) derived in [31].

C. Comparison

In the FO approach, it was shown that MUSIC was superior to ESPRIT [17], [15] with the difference becoming significant for large array size. However, we will show that this is not the case in the context of linear arrays. In
The result is also true for the FO estimates.

V. SPECIAL CASES

The above properties already lend some insight into the effect of spatial smoothing and weighting on the DOA estimates. Here the one and the two source cases are analyzed in detail to obtain more insight. Only eigen-based state space methods are examined because unweighted MUSIC has been studied in [5], [3], and optimally weighted MUSIC is the same as optimally weighted state space methods (cf. Theorem 7).

A. Unweighted State Space Methods

Here \( W = I \), and the results are specialized to the one and two source case.

1) One Source: For the one source case the covariance matrix is independent of the approach used and is given by

\[
R = R_s + \sigma_n^2 I
\]

where \( R_s = (LP_1) V(\omega_1) V^H(\omega_1) P_1 \).

Note that \( L = L_K \), and \( L \) is being used for notational simplicity. Also, from Property 5, both the FS and FBS approach have identical performance. It can be shown that the mean-squared error expression in this special case simplifies to

\[
\frac{(\Delta \omega_1)^2}{\alpha_{opt}} = \frac{1}{N KL K^2 P_1} \left( 1 + \frac{2(L_K - K)}{L_K} \frac{\sigma_n^2}{L_K P_1} \right). \tag{35}
\]

The details are presented in Appendix F. It is useful to note that though \((\Delta \omega_1)^2\) does not depend on \(\omega_1\), \((\Delta \theta_1)^2\) does depend on \(\theta_1\) by virtue of (22). From (35), the variance depends inversely on the number of subarrays \(K\) simulating the effect of \(K\) identical, but uncorrelated arrays. However, unlike uncorrelated arrays, in this case \(L_K\) also depends on \(K\) affecting the DOA variance, and so \(K\) cannot be made arbitrarily large. The optimum choice of \(K\) is between \(L_1/3\) and \(L_1/4 \cdot L_1/3\) is a better choice for high SNR and \(L_1/4\) at lower SNR. Using \(K = L_1/3\) for high SNR

\[
\left( \frac{(\Delta \omega_1)^2}{\alpha_{opt}} \right)_{opt} = \frac{1}{N 4 L_1^2 P_1} 27 \frac{\sigma_n^2}{\alpha_{opt}}.
\]

This is a significant improvement over the case when no smoothing was used [8], [15], and is close to the Cramer-Rao lower bound given by \((1/N)(1/6L)^2(\sigma_n^2/P_1)\). Also note that the performance is comparable to MUSIC [14], [17].

2) Two Source Case: Here we consider two sources that are closely spaced giving rise to the high resolution problem. The forward only covariance matrix is given by

\[
R_f = APA^H + \sigma_n^2 I \quad \text{where} \quad P = LP_1 \begin{bmatrix}
1 & \alpha \rho \alpha^2 \\
\alpha \rho & \alpha^2
\end{bmatrix}
\]

with \( \rho = |\rho| e^{j\beta} \), \( 0 \leq |\rho| \leq 1 \), \( 0 \leq \alpha \leq 1 \), and \( A = [V(\omega_1), V(\omega_2)] \). The FS and FBS covariance matrix have similar structure except that the \( P \) has to be modified appropriately. \( \alpha \) measures the relative power of the sources and \( \rho \) is the correlation coefficient. The closeness of the sources is measured by \( \epsilon \) where

\[
\epsilon = \epsilon F (\omega_1) V(\omega_2) = \frac{1}{L} e^{-j(L - 1)\epsilon \omega_2} \sin (L\omega_2) \frac{\sin (\omega_2)}{\sin (\omega_2)}.
\]

If it is assumed that \((L\omega_2)^2 << 1\), then [11]

\[
\epsilon = e^{-j(L - 1)\epsilon \omega_2} \left[ 1 - \frac{1}{8} \frac{L^2 \omega_2^2}{120} \frac{L^4 \omega_2^4}{4} + \cdots \right]
\]

with \( \omega_2 = (\omega_1 - \omega_2)/2 \). For the high resolution case \(|\epsilon| \approx 1\) and \(1 - |\epsilon| \approx \Delta^2/2\), where \( \Delta = (L\omega_2)/\sqrt{3} \).

Only \( \rho \) changes when the other covariance estimates are used, the details are discussed in [3]. For the FS and FBS approaches we have

\[
\rho_{fs} = \rho e_2 \quad \text{and} \quad \rho_{fs} = |\rho| e^{jL(1-1)\epsilon \omega_2} \sin (L(1 - 1)\epsilon \omega_2 - \xi)
\]

where

\[
\epsilon K = \frac{1}{K} e^{-j(L - 1)\epsilon \omega_2} \sin (K\omega_2) \frac{\sin (\omega_2)}{\sin (\omega_2)}.
\]

Just as in the one source case, it can be shown that (details are in Appendix B)

\[
\frac{\Delta \omega_2}{\alpha_{opt}} = \frac{3}{2NK L_2 (1 - |\epsilon|)} \left[ \sigma_n^2 I H P^{-1} I_1 \\
+ \sigma_n^2 I_H A^{-1} P^{-1} I_1 \right]. \tag{37}
\]

Equation (37) is valid for the general resolution scenario. The different covariance approaches can be accommodated by using the appropriate \( P \) matrix. For instance, for the FBS approach, in the case of equipower sources \((\alpha = 1)\), and real \( \rho \), (37) can be simplified to

\[
\frac{(\Delta \omega_2)^2}{\alpha_{opt}} = \frac{3}{2NK L_1 (1 - |\epsilon|) ASNR} \left( \frac{1}{(1 - |\rho_{fs}|^2)} + \frac{1}{\text{ASNR} \Delta^2 (1 - |\rho_{fs}|^2)} \right)
\]
where ASNR = \( (L_1 P_1 / \sigma^2_0) \). Ignoring\(^7\) the dependence of \( \rho_{n0} \) on \( K \), the optimum choice of \( K \) is \( L_1 / 5 \) for high SNR, tending to \( L_1 / 8 \) for lower SNR.

Other interesting observations that follow from (37) are

1. The mean-squared error in the estimates depends only on the separation \( \omega_d = (\omega_1 - \omega_2) / 2 \) and not on the actual values \( \omega_1 \) and \( \omega_2 \). Such an observation was first made in the context of sinusoidal frequency estimation [32].

2. The phase of the correlation coefficient \( \rho_1 \), i.e., \( \xi \), plays an important role in the effective correlation coefficient (cf. 36), and hence on the quality of the DOA estimate. From (36) and (37), it is clear that the method performs well for values of \( \xi \) which satisfy \( (L_1 - 1) \omega_d - \xi = n(\pi / 2) \), for nonzero integer \( n \), and poorly for values of \( \xi \) which satisfy \( (L_1 - 1) \omega_d - \xi = m\pi \), for integer \( m \). This behavior is similar to that of the initial phase of the sinusoids in the sinusoidal frequency estimation case [21], [32].

B. Optimally Weighted State Space Methods

The MSE expression is quite complicated even for the one and two source case and so only explicit expressions for the FO and FB approach (\( K = 1 \)) are provided. Exploiting Theorem 7, the MSE expression for the FO approach is given by

\[
\Delta \omega^T = \frac{1}{2N^2P_1^2} \left[ \frac{\sigma^2_1}{L_1} P_1^{-1} + \frac{\sigma^2_2}{P_1} \right].
\]

The expression for the FB approach is obtained by replacing \( P \) by \((P + JP^*J) / 2\) in (38). In the one source case, FO and FB approaches have the same performance and the expression can be further specialized to

\[
\Delta \omega^T = \frac{6}{L_1(L_1 - 1)} \left[ \frac{\sigma^2_1}{L_1 P_1} + \frac{\sigma^2_2}{P_1} \right].
\]

Comparing it to the unweighted case, which can be obtained from (35) by setting \( K = 1 \), it can be seen that the optimal weighting significantly improves the performance particular if the length of the array \( L_1 \) is large. Hence, optimal weighting provides another effective way to improve the performance. Similar conclusions also hold in the two source case.

\(^7This seems to be reasonable except in the extreme cases of correlation and resolution.

VI. NON-GAUSSIAN SOURCE AMPLITUDES AND THE FREQUENCY ESTIMATION PROBLEM

A. Non-Gaussian Source Amplitudes

In this section, we shall show that the mean squared error in the subspace estimates is independent of the source amplitude distribution, i.e., Assumption A2 is not necessary. For this we need to examine \( \Gamma_{\alpha_1,\gamma_1} \Gamma_{\alpha_1,\gamma_2} \) and \( \Gamma_{\gamma_1,\gamma_2} \gamma_2 \) where \( \alpha_i, i = 1, 2 \) lie in the noise subspace and \( \gamma_i, i = 1, 2 \) lie in the signal subspace (cf. (21)). Similar observations were also made in the context of subspace fitting methods in [16], [18], [19].

Theorem 8: Under Assumptions A1 and A3, and that the source amplitudes belong to a distribution having finite second-order moments, \( \Gamma_{\alpha_1,\gamma_1} \Gamma_{\alpha_1,\gamma_2} \) and \( \Gamma_{\gamma_1,\gamma_2} \gamma_2 \) where \( \alpha_i, i = 1, 2 \) lie in the noise subspace and \( \gamma_i, i = 1, 2 \) lie in the signal subspace, are dependent only on the second-order moments of the source amplitudes.

Proof: By definition, we have

\[
\begin{align*}
\Gamma_{\alpha_1,\gamma_1} \Gamma_{\alpha_1,\gamma_2} & = \text{cov} \{ \alpha_1^H \hat{R}_{\alpha_1}, \alpha_1^H \hat{R}_{\alpha_1} \} \\
& = \{ \alpha_1^H (\hat{R} - \hat{R} \gamma_1) \} \{ \alpha_1^H (\hat{R} - \hat{R} \gamma_2) \}^H.
\end{align*}
\]

For simplicity, only the FS approach is shown, and the same ideas readily extend to the FBS approach. For the FS approach, from (1)

\[
\hat{R}_a = \frac{1}{K} \sum_{k=1}^K \sum_{n=1}^N Y_p(n) Y_p^H(n) = \frac{1}{N} \sum_{n=1}^N \left( X_p(n) + N_p(n) \right)^H \left( X_p(n) + N_p(n) \right)
\]

where

\[
Y_p(n) = \frac{1}{\sqrt{K}} \left[ Y_1(n), Y_2(n), \ldots, Y_K(n) \right] = X_p(n) + N_p(n).
\]

and are defined in a manner similar to \( Y_p(n) \):

\[
R_a = \frac{1}{N} \sum_{n=1}^N \left( X_p(n) X_p(n)^H \right) + \sigma^2_2 I = R_n + \sigma^2_2 I.
\]

Using the orthogonality of \( \alpha_1 \) and \( \gamma_1 \), and that \( \alpha_1 \) is a noise eigenvector,

\[
\alpha_1^H (\hat{R}_a - R_a) \gamma_1 = \alpha_1^H \hat{R}_a \gamma_1 - \alpha_1^H \hat{R}_a \gamma_1 = \alpha_1^H \hat{R}_a \gamma_1.
\]

Since \( \alpha_1 \) is the noise subspace, \( \alpha_1^H X_p(n) = 0 \ \forall n \). Thus from (40)

\[
\alpha_1^H \hat{R}_a \gamma_1 = \frac{1}{N} \sum_{n=1}^N \alpha_1^H (N_p(n) X_p(n))^H \gamma_1 + \alpha_1^H N_p(n) N_p(n)^H \gamma_1.
\]

Thus using (42) and (41) in (39)

\[
\Gamma_{\alpha_1,\gamma_1} \Gamma_{\alpha_1,\gamma_2} = \frac{1}{N} \sum_{n=1}^N \left( \alpha_1^H N_p(n) X_p(n)^H \gamma_1 \right) \left( \alpha_1^H N_p(n) N_p(n)^H \gamma_1 \right)^H + \alpha_1^H N_p(n) N_p(n)^H \gamma_1.
\]

Therefore, \( \Gamma_{\alpha_1,\gamma_1} \Gamma_{\alpha_1,\gamma_2} \) and \( \Gamma_{\gamma_1,\gamma_2} \gamma_2 \) are independent of the source amplitude distribution, i.e., Assumption A2 is not necessary.
Noting that the odd moments of the Gaussian noise vectors are zero, we have

$$
\Gamma_{\eta_1, \eta_2} = \frac{1}{N^2} \sum_{n=1}^{N} \sum_{m=1}^{N} \{ (\alpha^H \eta_1(n) X_F(n)^H \eta_2) (\alpha^H \eta_2(n) X_F(m) \eta_1)^H \\
+ (\alpha^H \eta_1(n) X_F(n) \eta_1)(\alpha^H \eta_2(n) X_F(m) \eta_2)^H \}.
$$

The result follows since only the first term depends on the signal amplitudes and only on their second-order properties. One can provide a similar result for $$\Gamma_{\eta_1, \eta_2}$$.

Remark: Note that the Gaussian assumption on the noise can be relaxed in the above theorem to random variables having finite fourth moments [33]. However, to obtain tractable expressions the Gaussian assumption is useful.

B. Deterministic Sources and Frequency Estimation

Another important case that deserves special attention is the case where the amplitudes are deterministic. The noise is still assumed to be zero mean and circularly Gaussian. Theorem 8 is still valid, and all that needs to be done is to carefully define the covariance matrix $$R$$. For the case of deterministic sources, for the FBS approach we define

$$
R_{na} = \frac{1}{N} \sum_{n=1}^{N} \left( X_F(n), JX_F^T(n) \right) \left( X_F(n), JX_F^T(n) \right)^H + \sigma_a^2 I
$$

and

$$
R_{pa} = \frac{1}{N} \sum_{n=1}^{N} X_F(n) X_F^H(n) + N_{pa}.
$$

In addition to the direct dependence of the performance on the number of snapshots $$N$$ (cf. the denominator of (24)), $$\beta_i = R_{ni} V(\omega_i)$$ and $$R_{pa}$$ also now depend on $$N$$. Asymptotically, they can be replaced by $$R_{na}^\infty$$ and $$R_{pa}^\infty$$ where

$$
R_{na}^\infty = \lim_{N \to \infty} R_{na}
$$

and

$$
R_{pa}^\infty = \lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} X_F(n) X_F^H(n) + \sigma_a^2 I,
$$

and $$R_{pa}^\infty$$ is defined in a similar manner. However, it is useful to retain the dependence of $$\beta_i$$ and $$R_{pa}$$ on $$N$$ as the expression in (24) is then a valid first-order perturbation expression for all $$N$$. This allows for the transition from the asymptotic arguments to the finite number of snapshots case. The special case of $$N = 1$$ corresponds to the time series case and the results can be used to understand the time series methods [9], [10].

C. Comparison with Previous Results

The results when applied to the frequency estimation problem are no longer asymptotic results and are valid for high SNR situations. However, they are more general than existing results. Normally, the time series analysis results start with data matrices $$[Y_F(1), JY_F^T(1)]$$, and treat $$[N_F(1), JN_F^T(1)]$$ as a perturbation and obtain expressions for $$\Delta z$$ that are linear in $$[N_F(1), JN_F^T(1)]$$ [32]. Consequently, the mean-squared error depends linearly on the noise to signal ratio. The same is true of similar analysis in the array processing area [34]. In contrast, in this paper, by combining (42) and (19), we have for the FBS case

$$
\frac{\Delta z}{z} = \frac{1}{2} \varepsilon^H \left( [N_F(1), JN_F^T(1)] (X_F(1), JX_F^T(1))^H \right) \beta_i
$$

$$
+ \frac{1}{2} \varepsilon^H \left( [N_F(1), JN_F^T(1)] (N_F(1), JN_F^T(1))^H \right) \beta_i.
$$

The first term corresponds to the result one obtains using a first-order perturbation of the data matrix. The presence of the second term (quadratic noise term) in the above equation results in the mean-squared error being dependent on the square of the noise-to-signal ratio also. Consequently the expressions derived in this paper are more general and are valid for lower signal-to-noise ratios compared to the analysis based on a linear perturbation of the data matrices.

Most of the properties derived can be shown to hold true. For instance, for state space based frequency estimation methods, Properties 2-5 are still true. In fact, Property 4 is approximate because of the inclusion of the higher order terms and is exact if only the linear approximation is used as in [32]. Property 5 can be assumed to hold approximately if the length of the array $$L_K$$ is held fixed, and the number of subarrays is increased, i.e., $$K$$ and $$L$$ are increased. This is because the situation approximately corresponds to $$R$$ being Toeplitz [17], i.e., uncorrelated sources.

VII. Simulations

In this section we provide some numerical experiments and they all support our theoretical claims. In the examples considered, computer experiments were carried out and the statistics computed from 200 independent trials. They agree closely with our theoretical predictions and hence only the theoretical results (cf. (22)) are shown.

Example 1: A uniform linear array with $$L = 24, M = 1$$, DOA = 18°, $$N = 100$$ is considered. The unweighted state space method employing a FBS covariance matrix is used. Figs. 2(a) and (b) show the mean-squared error in the DOA estimates for SNR = 10 dB and SNR = -20 dB as a function of $$K$$, the number of subarrays.
It is observed that there exists an optimum $K$ which is in agreement with the discussion in the previous section for the high and low signal-to-noise ratios. In this example, for SNR = 10 dB, $K_{opt} = 8$ and for SNR = -20 dB, $K_{opt} = 6$.

Example 2: A uniform linear array with $L_1 = 24$, $M = 2$, DOA = 18°, 20°, $N = 100$, SNR = 10 dB is considered. Figs. 3(a) and (b) show the mean-squared error in the DOA estimates using the unweighted state space method corresponding to 18°, for various real values of the correlation coefficient $\rho$ as a function of $K$. Fig. 3(a) shows the performance when the FS covariance matrix is used and Fig. 3(b) shows the case when FBS covariance matrix is used.

1) It is seen that for uncorrelated sources ($\rho = 0$), both FS and FBS result in the same MSE as indicated by Property 4.

2) It is seen that the performance deteriorates as $\rho$ increases, with the deterioration being more severe for the FS approach.

3) An optimum value for $K$ exists for the FBS approach which is approximately independent of $\rho$ and agrees with the theoretical expressions.

The results corresponding to DOA 18° as a function of the phase $\xi$ (0 to $\pi$) of a complex $\rho$ are shown in Fig. 3(c). For this experiment $K = 5$. As expected, the performance depends on $\xi$, with the worst performance being for a value of $\xi = (L_1 - 1) \omega_f$.

Example 3: The results for the optimally weighted state space method are shown in Fig. 4. The scenario considered is exactly the same as that in Example 2 with a real correlation coefficient $\rho$. Note that the performance of the method improves, and the improvement is particularly significant for $K = 1$. 

---

**Fig. 2.** One source case: Mean-squared error in DOA as a function of the number of subarrays $K$. Parameters in this example are $L_1 = 24$, $M = 1$, $N = 100$, DOA = 18°. (a) SNR of 10 dB, and (b) SNR of -20 dB.
Example 4: Here we consider the single snapshot case, i.e., $N = 1$, for the scenario described in Example 2. This corresponds to a time series frequency estimation problem with the frequencies of the complex exponentials being given by $\omega_1 = 0.31\pi$ and $\omega_2 = 0.34\pi$. Fig. 5 shows the mean-squared error in the angular frequency corresponding to $0.31\pi$ as a function of $K$ for the FBS approach (theory and computer simulations). There is a good agreement between them (to within 1 dB) supporting the discussion in Section VI.

VIII. CONCLUSIONS

In this paper, expressions for the mean-squared error in the estimates of the signal zeros and the DOA estimates obtained using the weighted eigen-based state space methods/ESPRIT and weighted MUSIC were derived. Based on these expressions, some general properties of the estimates and optimal weighting matrices were derived. An important outcome of the analysis is that state space methods/ESPRIT can be significantly improved by properly choosing the number of subarrays and by using proper weighting matrices. Another important result presented in the paper is that the assumption that the source amplitudes are circularly Gaussian is not necessary for deriving the asymptotic expressions, and that they depend only on the second-order properties of the source amplitudes. This enables the development of an unified framework for dealing with DOA estimation using ULA's and the time series frequency estimation problem.
RAO AND HARI: WEIGHTED SUBSPACE METHODS AND SPATIAL SMOOTHING

Fig. 3. (Continued) (c) Two source case: The performance of the FBS state space method as a function of the phase of the correlation coefficient $\rho$ is shown. The values of $|\rho|$ used are 0.25, 0.5, 0.75, and 1.0. The parameters of the example are $L_1 = 24$, $N = 100$, DOA equal to $18^\circ$ and $20^\circ$, and $L_F = 20$.

Fig. 4. The mean-squared error in the DOA estimate corresponding to $18^\circ$ obtained using the optimally weighted state space method are shown as a function of $\rho$, and the number of subarrays $K$. Parameters used in this example are $L_1 = 24$, $M = 2$, $N = 100$, DOA equal to $18^\circ$ and $20^\circ$, and $\rho$ equals 0, 0.25, 0.5, 0.75, and 1.0.

APPENDIX A

First we present a proof of Theorem 1. For ease of readability, the result is restated.

Theorem 1: For a given number of subarrays $K$, in the noise free case the FBS covariance matrix is better conditioned than the FS covariance matrix, i.e.,

$$(\lambda_1)_F \leq (\lambda_1)_S \quad \text{and} \quad (\lambda_M)_F \geq (\lambda_M)_S$$

resulting in

$$\kappa(R_{FS}) \leq \kappa(R_{FS})$$

where $\kappa(R) = \|R\|_2\|R^+\|_2 = \lambda_1/\lambda_M$ denotes the condition number of $R$.

Proof: Let

$$R_S = \bar{E}_S \bar{\Lambda}_S E_S^H \quad \text{and} \quad R_{FS} = E_S \bar{\Lambda}_F E_S^H.$$ 

Since the signal eigenvectors of both matrices span the same signal subspace there must exist a unique, nonsingular, unitary matrix $\varphi$ which satisfies

$$E_S \varphi = \bar{E}_S.$$ 

$$\kappa(R_{FS}) \leq \kappa(R_{FS})$$

(43)
The covariance matrix of the amplitudes is \( P \) whose elements are \( P_{ij} \), where \( P_{ij} = (p_i(n)p_j^*(n)) \). When the covariance matrix of the amplitudes \( P \) is singular, the sources are said to be coherent.

A3) The \( N \) snapshot vectors are assumed to be independent and the signal eigenvalues of \( R \) are distinct.

**Appendix C**

The proof of Property 2 is presented in this Appendix. Some identities that are necessary for the proof are discussed first. From the fact that

\[
V(\omega) = J^*V^*(\omega)e^{jL-1}\omega \tag{46}
\]

it can be deduced that

\[
J^*A_i = A_i^*e^{-j(L-1)\omega} \quad \text{and} \quad JP_n^*J = P_n \tag{47}
\]

where \( J \) is a \((L-1) \times (L-1)\) exchange matrix. Also for the FBS approach

\[
J\beta_\omega^* = e^{-j(L-1)\omega}\beta_\omega. \tag{48}
\]

Now we present a proof of Property 2:

**Property 2:** For the FBS approach using a centrohermitian weighting matrix,

\[
(\Delta \omega_\theta)_\text{ reacting error} = |\Delta \omega_\theta|_{\text{FBS}}
\]

and therefore

\[
|\Delta \omega_\theta|_{\text{FBS}} = |\Delta \omega_\theta|_{\text{FBS}} = \Gamma_{\text{react}} \Delta \omega_\theta
\]

where \( N_{\text{FBS}} = N_{\text{FBS}}(n)N_{\text{FBS}}(n) \).

**Proof:** Since \( \Delta \omega_\theta = \text{Im}(\Delta z_i/z_i) \) [32], one needs to prove

\[
\left( \frac{\Delta z_i}{z_i} \right)_{\text{FBS}} = -\left( \frac{\Delta z_i}{z_i} \right)_{\text{FBS}}^*
\]

Rewriting the expression for the error in the signal zero from (19)

\[
\left( \frac{\Delta z_i}{z_i} \right)_{\text{FBS}} = \epsilon_i^*\hat{R}_{\text{FBS}}^*\beta_i
\]

\[
= \epsilon_i^*JJ\hat{R}_{\text{FBS}}^*J\beta_i = -z_i^{-(L-1)}\epsilon_i^*\hat{R}_{\text{FBS}}^*z_i^{(L-1)}
\]

\[
= -(\epsilon_i^*\hat{R}_{\text{FBS}}^*\beta_i)^* = -\left( \frac{\Delta z_i}{z_i} \right)_{\text{FBS}}^*
\]

The above simplification has been carried out using (49), (48), and (3). To obtain the simplified expression for
using an optically weighted MUSIC estimate, i.e., the optimal centrohermitian matrix is equal to that obtained from (46) it can be shown that $J_{c}^{+}$ is also a vector in the noise subspace. This results in $R_{np}^{+} = N_{np} J_{c}^{+}$. Because the noise is white, $N_{np}$ is a real Toeplitz matrix, and satisfies $JN_{np} J = N_{np}$. Combining these facts with (48) and (49), (5) reduces to (24).

**APPENDIX D**

The proof of Property 7 is presented here.

**Property 7:** For the FBS weighted MUSIC approach, the performance obtained using any symmetric positive semidefinite matrix $W$ can be obtained with a centrohermitian matrix $W_{i} = (W + JW*J) / 2$.

**Proof:** Following steps similar to that in [5], it can be shown that

$$
\Delta \omega_{i}^{2} = \frac{-1}{2} \text{Re} \left[ (J_{c}^{+})^{T} (e^{j \omega_{i}}) W_{i}^{*} (J_{c}^{+}) (e^{j \omega_{i}}) \right]
$$

where

$$
\tilde{D}_{i} (e^{j \omega_{i}}) = -2 \text{Re} \left( \mu_{n}^{H} \tilde{\mathcal{R}}_{i} \right) = -2 \text{Re} \left( V_{n}^{H} \mathcal{P}_{n} W_{i} \tilde{\mathcal{R}}_{i} \right)
$$

with $\mu_{n}^{H} = V_{n}^{H} \mathcal{P}_{n} W_{i} \tilde{\mathcal{R}}_{i}$ is being used for notational simplicity to represent $V_{i} (\omega_{i})$. Using the identities $P_{n} = \mathcal{P}_{n} J$ (cf. (47)), $\beta_{i} = \delta_{i}^{l-1} J \beta_{i}^{l}$ (cf. (48)), $V_{n}^{H} P_{n} = (z_{i}^{T})^{(l-1)} V_{n}^{H} J$ [5], it can be shown that

$$
V_{n}^{H} P_{n} W_{i} \tilde{\mathcal{R}}_{i} = V_{n}^{H} P_{n} \left( W + JW*J \right) P_{n} \tilde{\mathcal{R}}_{i}
$$

From (50) this implies that the MSE in the DOA estimate is the same if $W$ is replaced by $(W + JW*J) / 2$.

**APPENDIX E**

To prove Theorem 7, some preparatory results are needed which are stated below.

**Lemma 1** [36]: If $H$ is rectangular and $S$ is symmetric and nonsingular,

$$(SH)^{+} = H^{+} S^{-1} [I - (SQS^{-1})^{+} (QS^{-1})]$$

where $Q = (I - HH^{+})$.

**Lemma 2:** For $T$ defined in (15)

$$
T^{+} T = I - V(\omega_{i}) V^{H}(\omega_{i}).
$$

**Proof:** This readily follows from the fact that the $(L - 1) \times L$ matrix $T$ has rank $L - 1$, and that $TV(\omega_{i}) = 0$.

**Theorem 7:** The MSE in the DOA obtained using the weighted space method/ESPRIT approach with an optimal centrohermitian matrix is equal to that obtained using an optimally weighted MUSIC estimate, i.e.,

$$
\left( \Delta \omega_{i}^{2} \right)_{\text{music}} = \left( \Delta \omega_{i}^{2} \right)_{\text{opt}}
$$

**Proof:** From (24)

$$
\left( \Delta \omega_{i}^{2} \right)_{\text{music}} = \epsilon_{i}^{H} P_{n} \mathcal{P}_{n} \epsilon_{i}
$$

and from (34)

$$
\left( \Delta \omega_{i}^{2} \right)_{\text{opt}}^{2} = \frac{1}{V_{n}^{H} (P_{n} \mathcal{P}_{n} P_{n}^{H})^{+} P_{n} V_{i}}.
$$

$V_{i} = V_{i} (\omega_{i})$, and $B$ is used instead of $B_{n}$ for notational simplicity. Let the SVD of $P_{n} \mathcal{P}_{n} P_{n}^{H}$ be $U_{n} \Sigma_{n} U_{n}^{H}$. Then

$$
\left( \Delta \omega_{i}^{2} \right)_{\text{opt}}^{2} = (\epsilon_{i}^{H} P_{n} \mathcal{P}_{n} \epsilon_{i}) (\epsilon_{i}^{H} P_{n} \mathcal{P}_{n} P_{n}^{H})^{+} P_{n} V_{i}
$$

$= \| U_{n} \Sigma_{n}^{1/2} \|^{2} \| P_{n} U_{n} \Sigma_{n}^{-1/2} \|^{2}

\geq \| U_{n} \Sigma_{n}^{1/2} P_{n} V_{i} \|^{2} = \| V_{i} \|^{2} = 1. \ (51)

The last equality was shown in [15]. This shows that optimal MUSIC always has a smaller variance than state space methods. Here we will show that optimal state space methods have the same mean squared error as optimal MUSIC. From (51), it is equivalent to showing that

$$
\epsilon_{i}^{H} U_{n} \Sigma_{n}^{1/2} = c_{1} \epsilon_{i}^{H} P_{n} U_{n} \Sigma_{n}^{-1/2}
$$

or

$$
P_{n} \mathcal{P}_{n} \epsilon_{i} = c_{1} \epsilon_{i} P_{n} V_{i}, \ (52)
$$

where $c_{1}$ is a scalar quantity whose exact value is not important for the purposes of this proof. In this proof, $c_{1}$, $c_{2}$, etc., will be used to represent scalars. We will show that (52) holds when the optimum choice of $W$ is made in the state space case, i.e., when $(TTT)^{-1}$ is substituted for $W$. In the expression for $\epsilon_{i}$ in (14) resulting in

$$
(\epsilon_{i})_{\text{opt}} = T^{H} (TTT)^{-1} A_{1} (A_{1}^{H} (TTT)^{-1} A_{1})^{-1} I_{j}
$$

where $T = TB_{i}^{H}$, with $B_{i}$ being a square root of $B$, i.e., $B = B_{i} B_{i}^{H}$. Since $B$ is symmetric, $B_{i}$ can be chosen to be symmetric, and is assumed to be symmetric for the purpose of simplifying the proof. It can then be shown that

$$
A_{1}^{H} (T_{j}^{H}) B_{i} \epsilon_{i} = I_{j}, \ \text{for all} \ W. \ (53)
$$

Note that $\epsilon_{i}$ is contained in the noise subspace and depends on $W$. By selecting $W$ properly, $(L - M)$ independent vectors that form a basis for the noise subspace can be generated. Therefore, it follows that

$$
P_{n} B_{i} T_{j}^{H} V(\omega_{i}) = 0, \ \text{if} \ j \neq i, \ \text{and} \ j = 1, 2, \cdots M \ (54)
$$

where $A_{i} = [V(\omega_{i}), V^{H}(\omega_{i}),\cdots, V^{H}(\omega_{M})]$ with $V^{H}(\omega) = 1 / \sqrt{L} [1, e^{j \omega}, \cdots, e^{j \omega (L-2 \omega)]}$, which is similar to $V(\omega)$ except that it is a vector of length $(L - 1)$. Using (53) and (54), it can be shown that

$$
P_{n} B_{i} \epsilon_{i} = c_{2} P_{n} B_{i} T_{j}^{H} V(\omega_{i}). \ (55)
$$

Note that this not true if a maximally overlapping configuration is not used.
Lemma 1 is used to expand $T_i^1$ in (55). Then the resulting expression is simplified with the help of Lemma 2, and the fact that $TV_i = c_1V^T(\omega_i)$. This results in
\[ P_nB(\epsilon_i)_{mp} = c_1P_nV_i, \]
This completes the proof.

**APPENDIX F**

Here we provide details of the steps and the approximations used in simplifying the expressions for the special cases considered in Section V. Similar simplifications were used in our earlier work [8].

**One Source Case:** From (14) it can be shown that for the one source case
\[ \epsilon^H = \begin{bmatrix} L^{-1} [1, 0, 0, \ldots, 0, -(\epsilon^*_i)^L-1] \end{bmatrix}. \]

Also since $R = (LP_i)V(\omega_1)V^H(\omega_1)$,
\[ \beta_1 = R^0 V(\omega_1) = \frac{1}{LP_i} V(\omega_1). \]

It will be assumed that $L \gg 1$ and the approximations $L \pm 1 = L$ will be used. For $K \leq L$, using the sparsity of $\epsilon_1$, the first term of (24) reduces to
\[ \frac{1}{2NK^2} \sum_{p=1}^{K} \sum_{q=1}^{K} (\beta^H R_{pp} \beta_i)(\epsilon^H Q_0 \epsilon_1) \]
\[ = \frac{1}{2NK^2} \sum_{p=1}^{K} (\beta^H R_{pp} \beta_i) \|\epsilon_1\|^2 \sigma_n^2 \]
\[ = \frac{1}{NKL^2} \frac{\sigma_n^2}{P_1} \left(1 + \frac{\sigma_n^2}{P_1} \right) \]  
(56)

and the second term reduces to
\[ \frac{1}{2NK^2} \sum_{p=1}^{K} \sum_{q=1}^{K} \sum_{q \neq p} (\beta^H R_{pp} \beta_i) \|\epsilon_1\|^2 \sigma_n^2 \]
\[ = \frac{K-1}{2NK} \frac{\sigma_n^2}{P_1} \]  
(57)

Combining (56) and (57) we have the simplified expression of (35).

**Two Source Case:** Just as in the one source case, it can be shown that $\epsilon_i$ is sparse and that the first and fast entry of the vector dominate, i.e.,
\[ \epsilon^H_i = [\epsilon^*_i(1), 0, \ldots, 0, \epsilon^*_i(L - 1)], \quad i = 1, 2 \]
and
\[ \|\epsilon\|^2 = \frac{3}{L(1 - |\epsilon|)}. \]

$\epsilon_i(k)$ is an element of the vector $\epsilon_i$. This is obtained starting from (14), and retaining the dominant terms. Now to determine the error in the DOA estimate, for simplicity, the second term in (24) is dropped. However, this seems to be a reasonable approximation. It is clearly valid at high SNR. Also, from the one source case comparing (56) and (57) it can be seen that (56) dominates even at low SNR and this seems to be the case even for this two source scenario. Using the sparsity of $\epsilon_1$, as in the one source case, the double sum reduces to a single sum, i.e.,
\[ \frac{\Delta \omega_1^2}{\Delta \omega_1^2} = \frac{1}{2NK^2} \sum_{p=1}^{K} (\beta^H R_{pp} \beta_i) \|\epsilon_1\|^2 \sigma_n^2 \]
\[ = \frac{1}{2NK} \beta^H R_{pp} \beta_i \|\epsilon_1\|^2 \sigma_n^2. \]

Note that this equation is valid for the general high resolution two source scenario. $\beta_i$ depends on the covariance matrix $\Sigma$. For the FBS case, using (48) it can be shown that $\beta^H R_{pp} \beta_i = \beta^H R_{pp} \beta_i$. Substituting for $\beta_i = R^0 V(\omega_1) = (A^H)^* P^{-1}_i$ and $\|\epsilon_1\|^2$, we have the simplified expression of (37).

**REFERENCES**


Bhaskar D. Rao (S’80–M’83–SM’91) received the B.Tech. degree in electronics and electrical communication engineering from the Indian Institute of Technology, Kharagpur, India, in 1979, and the M.S. and Ph.D. degrees from the University of Southern California, Los Angeles, in 1981 and 1983, respectively.

Since 1983, he has been with the University of California, San Diego, where he is currently an Associate Professor in the Department of Electrical and Computer Engineering. His interests are in the areas of digital signal processing, estimation theory, and parallel processing.

Dr. Rao is a member of the IEEE Statistical Signal and Array Processing Committee.

K. V. S. Hari obtained the B.E. degree from Osmania University, Hyderabad, India, in 1983, the M.Tech. degree from Indian Institute of Technology, Delhi, in 1985, and the Ph.D. degree from the University of California, San Diego, in 1990.

From January 1985 to July 1987, he was a Scientist at the Defence Electronics Research Laboratory, Hyderabad, India, and is presently an Assistant Professor in the Department of Electrical Communication Engineering at the Indian Institute of Science, Bangalore, India. His research interests are in the fields of array signal processing and spectral estimation algorithms.