A Performance Analysis of Subspace-Based Methods in the Presence of Model Errors: Part II—Multidimensional Algorithms

A. Lee Swindlehurst, Member, IEEE, and Thomas Kailath, Fellow, IEEE

Abstract—This is the second of a two-part paper dealing with the performance of subspace-based algorithms for narrow-band direction-of-arrival (DOA) estimation when the array manifold and noise covariance are not correctly modeled. In Part I, the performance of the MUSIC algorithm was investigated. In Part II, we extend this analysis to multidimensional (MD) subspace-based algorithms including deterministic (or conditional) maximum likelihood, MD-MUSIC, weighted subspace fitting (WSF), MODE, and ESPRIT. A general expression for the variance of the DOA estimates is presented that can be applied to any of the above algorithms and to any of a wide variety of scenarios (e.g., gain/phase errors, mutual coupling, sensor position errors, noise covariance mismodeling, etc.). Optimally weighted subspace fitting algorithms are also presented for special cases involving random unstructured errors to the array manifold and noise covariance. In addition, it is shown that one-dimensional MUSIC outperforms all of the above MD algorithms for random angle-independent array perturbations.

I. INTRODUCTION

The subject of this two-part paper is the sensitivity of narrow-band direction-of-arrival (DOA) estimation algorithms to certain assumptions about the data model. In particular, we consider the effects of calibration errors in the array response (gain and phase perturbations, mutual coupling effects, sensor position errors, etc.) as well as incorrect models for the spatial covariance of the noise field. In Part I of this paper [1] the performance of the well-known MUSIC algorithm [2], [3] was investigated. In Part II, we extend this analysis to the class of so-called multidimensional subspace fitting algorithms, which includes deterministic (or conditional) maximum likelihood (DML) [4], [5], multidimensional (MD)-MUSIC [6], [7], weighted subspace fitting (WSF) [8], [9], and total least squares (TLS)-ESPRIT [10]. All algorithms in this class can be shown to implement the same generic least squares subspace fitting (SSF) minimization problem [8], each algorithm differing in the choice of weighting matrices used in the SSF criterion function.

As in [1], we examine algorithm performance in the limiting (infinite snapshot) case to focus on situations where the array and noise modeling errors dominate the finite sample effects due to additive noise. In addition to providing a closed-form equation for the estimation error of all algorithms in the SSF class for a wide variety of model perturbations, we show that the resulting optimal subspace weightings are quite different from those obtained when only finite sample effects alone are considered. In cases where array calibration errors are not identical from sensor to sensor, we show that a row (or sensor space) weighting of the subspace is needed in addition to the more common column (or DOA space) subspace weighting.

Two special random error models are considered in detail, one for perturbations to the array manifold, and one for errors in the model of the noise covariance. Conditions are derived under which these errors dominate those due to finite sample effects alone. For the special case of angle-independent array errors, we are able to show that the MUSIC algorithm has lower estimate variance than deterministic ML, MD-MUSIC, MODE, and WSF. An appropriately weighted SSF algorithm can, however, achieve performance equivalent to that of MUSIC. When the array errors are nonuniform from sensor to sensor, optimal performance is obtained by a weighted version of MUSIC and a row- and column-weighted SSF algorithm.

Several other authors have recently considered the sensitivity of various MD algorithms to errors in the model for the array response and noise field statistics. Friedlander [11] investigated the performance of deterministic ML for array calibration errors, while Li and Vaccaro have studied least squares (LS)-ESPRIT [12] and TAM [13] (among other algorithms) for both calibration [14] and noise modeling errors [15]. Soon and Huang [16] have also recently considered the sensitivity of LS-ESPRIT to random sensor errors. Besides providing the first performance analysis of MD-MUSIC, TLS-ESPRIT, WSF, and MODE for errors in both the array and noise models, the research presented herein differs from earlier work by focusing more on algorithm comparison and optimization.
II. General Error Expressions

To introduce notation, we begin by briefly revisiting the data model presented in [1]. Suppose \( d \) narrow-band plane-wave signals impinge upon an \( m \)-element sensor array, and let the complex response of the array to a signal with DOA \( \theta \) be defined as \( a(\theta) \in \mathbb{C}^m \). Assuming that the signals are narrow band and that the array elements are linear devices, the array output \( x(t) \) may be written as

\[
x(t) = A(\theta) s(t) + n(t)
\]

where \( s(t) \in \mathbb{C}^d \) is the amplitude and phase of the signals at time \( t \), \( n(t) \) is additive noise, \( \theta = [\theta_1, \cdots, \theta_d] \) represents the DOA’s and \( A(\theta) = [a(\theta_1) \cdots a(\theta_d)] \). Assuming the signals and noise are uncorrelated and that the noise is spatially white, the covariance of the array data under this model is given by

\[
R = E \{ x(t)x^*(t) \} = A(\theta)SA^*(\theta) + \sigma^2 I
\]

where \( S \) is the covariance of the signals, \( \sigma^2 \) is the noise power, and \( (\cdot)^* \) denotes complex conjugate transpose.

The eigendecomposition of \( R \) will be defined as follows:

\[
R = \sum_{\nu=1}^{\nu_{\min}} \lambda_{\nu} e_{\nu} e_{\nu}^* = E \Lambda E^* + \sigma^2 E e_{\nu} e_{\nu}^*
\]

where \( \nu_{\min} \) is the number of modes, \( \lambda_{\nu} \) are the diagonal elements of the covariance matrix \( R \). These perturbations will be modeled using the following expression for the array covariance matrix \( \tilde{R} \) first introduced in [1]:

\[
\tilde{R} = (I + \Delta)(A + \tilde{A}) S (A + \tilde{A})^* + \sigma^2 (I + \tilde{\Sigma}) (I + \Delta)^*,
\]

where \( \Delta \) and \( \tilde{\Sigma} \) are, respectively, row and column subspaces of the data model. The notation \( \tilde{\Sigma} \) is used to indicate that some of the algorithms under consideration (i.e., WSF and some of the optimal algorithms developed later) rely on estimates of an exact weighting \( W \) which is generally not available. Although the parameter vector \( \eta \) is often associated with the DOA’s \( \eta = \theta \), the estimated parameters for ESPRIT include not only the DOA’s, but also various other quantities (see [8]). Thus, to maintain generality, it is useful to make a distinction between the two cases.

Two notes are in order here with regard to DML, MODE, and the SSF cost function of (4). Although neither DML nor MODE is directly formulated as in (4), the error expressions derived below can still be applied to them by appropriate choice of the weighting matrices. To first order in the model perturbation, the DML approach can be shown to have equivalent performance to (4) implemented with \( W_r = I \) and \( W_r = \tilde{A} \), while MODE can be shown to be equivalent to WSF when \( d = d' \) (see [20] for details). Similar results have been presented for these algorithms when finite sample effects and model perturbations are the source of estimation error [21].

B. Estimation Error Covariance

In this analysis, the parameter estimate \( \hat{\theta} \) will be assumed to differ from the true parameter \( \theta \) because of various random perturbations to the above data model. These perturbations will be modeled using the following exact representation of the array covariance matrix \( \tilde{R} \) introduced in [1]:

\[
\tilde{R} = (I + \Delta)(A + \tilde{A}) S (A + \tilde{A})^* + \sigma^2 (I + \tilde{\Sigma}) (I + \Delta)^*.
\]

The matrix \( \Delta \) contains errors which affect both the signal and noise components of the data, such as gain imbalances in the receiver electronics, channel crosstalk, and mutual coupling effects. Errors in the nominal array response, including the effects of imprecisely known sensor locations, perturbations in the antenna amplitude and phase patterns, and signal-only mutual coupling are incorporated into \( \tilde{R} \) by the error term \( \tilde{\Sigma} \). In addition, the Hermitian matrix \( \tilde{\Sigma} \) represents any deviation in the spatial covariance of the additive noise from its nominal value of \( I \). All of the error terms \( \Delta, \tilde{\Sigma} \) may in general be \( \theta \)-dependent.

The effect of these model errors on the estimated signal subspace \( \hat{E} \) and on the parameter estimate \( \hat{\theta} \) can be quantified by means of a first-order perturbation analysis similar to that presented in [1]. The resulting error expressions will be given in terms of the following matrices:

\[
\Xi = [\xi_1 \cdots \xi_d] = [\tilde{A} + \Delta \tilde{A} + \sigma^2 (\tilde{\Sigma} + \tilde{\Sigma}^*) + \Delta \tilde{A}] T \tilde{A}^{-1} T^{-1}
\]

Since \( A^* S^{-1} = AT \tilde{A}^{-1} T^{-1} \) when \( S \) is full rank, our def-
The expression for the estimation error is summarized by the following theorem.

**Theorem 2.1:** For the SSF minimization of (4) and the perturbed covariance model of (5), the estimation error \( \eta - \eta_0 \) is zero mean with covariance given by

\[
C = \mathbb{E}\{(\eta - \eta_0)(\eta - \eta_0)^H\} = (V^*)^{-1}Q(V^*)^{-1}
\]

where

\[
V_{\eta}^* = 2\text{Re}\left[\text{tr}(A_0^* W_r^{*1/2} P^1 W_r^*/2 A_0 T W_r T^*)\right]
\]

\[
Q_{\eta} = 2\text{Re}\left[\sum_{p=1}^{d}\sum_{q=1}^{d} Y^{(p)} C_{\xi_1}^{pq} Y^{(q)*} + Y^{(p)} C_{\xi_2}^{pq} Y^{(q)*}\right]
\]

\[
Y^{(p)} = TW_r T^* A_0^* W_r^{*1/2} P^1 W_r^*/2
\]

\[
A_i = \frac{\partial A}{\partial \eta_i}\bigg|_{\eta - \eta_0}
\]

\[
\eta_i = i^{th} \text{element of } \eta
\]

\[
C_{\xi_1}^{pq} = \mathbb{E}\{\xi_p \xi_q^*\}
\]

\[
C_{\xi_2}^{pq} = \mathbb{E}\{\xi_p \xi_q^T\}
\]

**Proof:** A proof can be found in [20, sec. 4.3]

Theorem 2.1 states that given a perturbation to the data model that can be described by (5), (12), and (13), the expected variation in the estimates obtained from any of the algorithms in the SSF class may be evaluated by substituting appropriate values for \( W_r \) and \( W_c \) into (7)-(9). Expressions for the perturbation covariances in (12) and (13) were presented in [1] for several special error models (e.g., simple gain and phase errors, mutual coupling, sensor location errors, channel gain imbalances, etc.), and those may be applied directly here.

If \( \eta = 0 \) (which is the case for all multidimensional SSF algorithms except ESPRIT), then the following compact expressions for \( V^* \) and \( Q \) are possible:

\[
V^* = 2\text{Re}\left[(D^* W_r^{*1/2} P^1 W_r^*/2) \odot (T W_r T^*)^T\right]
\]

\[
Q = \sum_{p=1}^{d} \sum_{q=1}^{d} 2\text{Re}\left[(D^* W_r^{*1/2} P^1 W_r^*/2) \odot (T W_r T^*)^T\right]
\]

\[
\odot (T W_r T^* A_0^* W_r^{*1/2} P^1 W_r^*/2)\]

where \( D = [d(\theta_1), \ldots, d(\theta_d)] \),

\[
d(\theta_i) = \frac{\partial d(\theta)}{\partial \theta}\bigg|_{\theta = \theta_0}
\]

\( t_p \) and \( t_q \) denote the \( pth \) and \( qth \) rows of \( T \), respectively, \( \odot \) denotes an elementwise (Hadamard) product, and \( (\cdot)^T \) indicates the complex conjugate.

**III. ALGORITHM COMPARISONS AND OPTIMIZATION**

The expression for the error covariance derived above is extremely general. As a consequence of its applicability to several different algorithms and types of modeling errors, it is necessarily quite complex (though easily programmable) and it is difficult to draw any conclusions about the relative performance of SSF algorithms. While it is convenient to have such an expression for algorithm analysis, it is also desirable to consider some particularly simple error models for which algorithm comparison is possible. Two such models are considered in this section, one for errors to the array manifold \( A \), and one for errors to the nominal noise covariance \( \Sigma \).

**A. Errors in the Array Model**

Restricting our attention for the moment to errors in the nominal array response, we have \( \Xi = \hat{A} \), and the error covariances needed to evaluate \( C \) are

\[
C_{\xi_1}^{pq} = \mathbb{E}\{\hat{a}(\theta_p)a^*(\theta_q)\}
\]

\[
C_{\xi_2}^{pq} = \mathbb{E}\{\hat{a}(\theta_p)a^T(\theta_q)\}
\]

If we think of the perturbation to the array manifold \( \hat{a}(\theta) \) as a zero-mean multidimensional, complex random process in the variable \( \theta \), we see that a complete characterization of the second-order statistics of the process is necessary to determine \( C \). One such characterization that allows for angle dependent and angle correlated array errors is the following:

\[
\mathbb{E}\{\hat{a}(\theta_p)a^*(\theta_q)\} = v_{pq} B
\]

\[
\mathbb{E}\{\hat{a}(\theta_p)a^T(\theta_q)\} = 0 \quad \forall \ p, q
\]

where \( v_{pq} \) is a complex scalar representing the correlation between the errors in directions \( \theta_p \) and \( \theta_q \). As a simple example, it might be reasonable to assume that the array errors are exponentially correlated with \( \theta \), in which case \( v_{pq} \) might have the form

\[
v_{pq} = \sqrt{v_{pp}v_{qq}}e^{-\alpha|\theta_p - \theta_q|}
\]

for some constant \( \alpha \) and correlation phase \( \psi_{pq} \). The model of (16)-(17) is slightly more general than that presented in [1].
If we define the matrix $T$ with elements $T_{ij} = u_{ij}$, then using the statistical array perturbation model of (16)-(17) in Theorem 2.1 leads to the following simplified expressions for $V^* \eta$ and $Q$ when $\eta = \theta$:

$$V^* = 2\text{Re}[(D^* W_r^{1/2} P^ \perp W_l^{1/2} D) \odot (T W c T^*)]$$

$$Q = 2\text{Re}[(D^* W_r^{1/2} P^ \perp W_l^{1/2} BW_r^{1/2} P^ \perp W_l^{1/2} D) \odot (T W c T^*)]$$

Notice that if the row and column weightings are chosen as

$$W_r = B^{-1}$$

$$W_c = (T^* T^ T)^{-1}$$

then $Q = V^*$ and $C = (V^*)^{-1}$. The following theorem demonstrates that these are in fact the optimal subspace weightings which produce estimates of minimum variance.

**Theorem 3.1:** For the perturbed array manifold model of (16)-(17), the covariance of the DOA estimates obtained from the SSF minimization of (4) satisfies

$$C(B^{-1}, (T^* T^ T)^{-1} \leq C(W_r, W_c)$$

for all row and column weightings $W_r$ and $W_c$ when $\eta = \theta$.

**Proof:** A proof is given in [20].

Thus, when the array is subject to perturbations of the form (16)-(17) and finite sample effects may be neglected, the DOA estimates obtained from a SSF minimization using (21) and (22) will have lower variance than those obtained from WSF, MODE, DML, and MD-MUSIC. For convenience, in the sequel we will refer to this optimally weighted algorithm as robust subspace fitting (RSF) for array errors. Since the above result was derived assuming $\eta = \theta$, additional work is needed to analytically compare the performance of RSF with ESPRIT. However, all simulations performed to date indicate that the ESPRIT estimates have higher variance, though this variance may often be made to approach that of the RSF algorithm by proper subarray choice (see Example 4 of the following section).

If rank $(T^* T^ T) = p < d'$, then the optimal column weighting must be chosen as

$$W_c = U(U^* U)^{-2} U^* = U^* U^\dagger$$

where the $d' \times p$ matrix $U$ is defined as the square-root factor $T^* T^ T = U U^*$. The quantity $T^* T^ T$ will drop rank only if $T$ drops rank, which will in turn occur only if some subset of the array errors are 100% angle-correlated (e.g., this could result if the sensors experience a deviation in gain and phase that is uniform in $\theta$).

1) **Comparisons with MUSIC:** When the optimal weightings (21) and (22) are substituted into (19)-(20) above, the expression for the DOA estimation error covariance reduces to

$$C_{RSF} = \frac{1}{2} \{\text{Re}[(D^* B^{-1/2} P G B^{-1/2} D) \odot (T^*)^{-1}]\}^{-1}.$$  

where $P_G$ is the projection onto the column space of $G = B^{-1/2} A$, and where we have assumed that $S$ (and hence $T$) is full rank. The simple compact expression of (24) allows us to make an interesting comparison with the results obtained for the MUSIC algorithm in [1]. It was shown in [1] that in the presence of non-uniform array errors, the performance of MUSIC could be improved by employing the following weighted cost function:

$$V_{WMU} = \frac{a^*(\theta) E_n W_{MU} E_n^* a(\theta)}{a^*(\theta) a(\theta)}$$

where $W_{MU} = (E_n^* B E_n)^{-1}$ when the noise is spatially white. The following corollary demonstrates that when $T$ is diagonal in the simple array error model of (16)-(17), the one-dimensional weighted MUSIC algorithm has performance identical to the multidimensional RSF algorithm, and consequently has better performance than WSF, MODE, MD-MUSIC, and DML.

**Corollary 3.1:** For the array error model of (16)-(17), the covariance of the DOA estimates obtained from the weighted MUSIC algorithm presented in [1] is identical to that of the RSF algorithm given in (24) when $T$ is diagonal.

**Proof:** Substituting the model of (16)-(17) into the expression derived in [1] for the weighted MUSIC cost function leads to

$$C_{WMU} = \frac{1}{2} \{[D^* E_n (E_n^* B E_n)^{-1} E_n^* D] \odot (T^*)^{-1}\}^{-1}$$

when $T$ is diagonal. Since $E_n \perp A$, we have $B^{1/2} E_n \perp B^{-1/2} A$, and consequently

$$P_G = B^{1/2} E_n (E_n^* B E_n)^{-1} E_n^* B^{1/2}$$

when $G = B^{-1/2} A$. Substituting (26) into (24) then establishes the corollary.

When $B = I$ (i.e., when the array errors are uncorrelated from sensor to sensor) no weighting is necessary for MUSIC, and hence weighted MUSIC reduces to the standard MUSIC cost function. Corollary 3.1 implies that in such cases, regular MUSIC will outperform WSF, MODE, MD-MUSIC, and DML.

It is important to remember that the performance comparisons above apply only when finite sample effects due to noise can be neglected. In fact, our experience indicates that the situations where MUSIC and the RSF algorithm enjoy the greatest performance advantage are
precisely those that would place them at their greatest disadvantage were finite sample effects significant. This observation will be illustrated in the next section by a simulation example. In any case, additional work is needed to properly analyze algorithm performance under the combined effects of both the finite sample approximation and array perturbations.

2) Comparison with Finite Sample Effects: As in [1], we derive below conditions for which the errors in the DOA estimates due to array perturbations exceed those due to finite sample effects alone. When only one emitter is present, MUSIC and all SSF algorithms yield an identical DOA estimate, and hence the following condition derived in [1] for the single source case will apply to all of the algorithms under consideration:

$$v^2 > \frac{m \cdot \text{SNR} + 1}{mN \cdot \text{SNR}^2} = \frac{1}{N \cdot \text{SNR}}$$

where for simplicity we have assumed that $B = I$, $T = v^2I$, and the sensors have unity gain in the direction of the signal. As an example, with $N = 100$ snapshots and a relatively low SNR of 0 dB, array errors will dominate if the standard deviation in gain and phase exceeds only 0.1 and 5.7°, respectively.

In [8], the following expression was derived for the asymptotic DOA estimation error in the SSF minimization of (4) due to the finite sample approximation alone:

$$C_{\text{FS}} = \frac{\sigma^2}{2N} \left[ \text{Re}\{D^*P_A^+D\} \odot (TW,T^*)^T \right]^{-1}$$

$$\cdot \left[ \text{Re}\{D^*P_A^+D\} \odot (TW,\tilde{\Lambda},W_cT^*)^T \right]$$

$$\cdot \left[ \text{Re}\{D^*P_A^+D\} \odot (TW,T^*)^T \right]^{-1}$$

When $W_c = (T^*T)^{-1}$ and $S$ is full rank, we have

$$TW,\tilde{\Lambda},W_cT^* = (T\tilde{\Lambda}^2\Lambda_s^{-1}T^*)^{-1}$$

$$= S^{-1} + \sigma^2(\text{SA*AS})^{-1}$$

and hence the finite sample variance of the RSF algorithm is given by

$$C_{\text{RSF}} = \frac{\sigma^2}{2N} \left( D^*P_A^+D \odot I \right)^{-1} \left[ \text{Re}\{D^*P_A^+D\} \right.$$

$$\odot (S^{-1} + \sigma^2(\text{SA*AS})^{-1}T^* \{D^*P_A^+D \odot I \}^{-1}$$

which is identical to the corresponding expression obtained for the MUSIC algorithm in [22].

Table I summarizes the relative contributions of the finite sample approximation and array response errors in the estimation error of the SSF algorithms under consideration. In particular, the inequalities are obtained by comparing (19)–(20) with (27) to indicate what the magnitude of the variance $v^2$ must be for array perturbations to be the dominant source of error for the $i$th DOA estimate (assuming again that $B = I$).

**B. Errors in the Noise Model**

While it is relatively easy to link a particular model for $\tilde{\Lambda}$ with the physical process or mechanism that generated it, the same cannot be said for errors in the noise model. Consequently, in the discussion which follows, we will consider the following relatively nondescriptive model for the noise covariance perturbation $\tilde{\Sigma}$ introduced in [1]:

$$E\{\tilde{\Sigma}_{ik}\} = 0$$

$$E\{\tilde{\Sigma}_{ik},\tilde{\Sigma}_{kp}\} = 0 \quad \forall \, i \neq q, \, p \neq k$$

$$E\{\tilde{\Sigma}_{ik},\tilde{\Sigma}_{ik}\} = \mu^2.$$  

(29)

Thus, we assume $\tilde{\Sigma}$ is composed of zero-mean random variables with identical variance $\mu^2$. The diagonal elements are real and independent of all other elements of $\tilde{\Sigma}$, while the off-diagonal elements are complex, circular, and correlated only with their conjugate image across the diagonal (since $\tilde{\Sigma}$ must be Hermitian).

Assuming errors to the noise model of the form (29), we have $\tilde{\Sigma} = \sigma^2 \tilde{\Lambda}AT\tilde{\Lambda}^{-1}T^*$, and the covariances of the perturbation vectors $\tilde{\xi}_i$ become

$$C_{\tilde{\xi}_{1}} = \mu^2 \sigma_4 \tilde{\Lambda}^{-1}q_4(T^*T)^{-1} \tilde{\Lambda}^{-2}(T^*T)^{-1} t_t$$

$$C_{\tilde{\xi}_{2}} = \mu^2 \sigma_4 \tilde{\Lambda}^{-1}(T^*T)^{-1} \tilde{\Lambda}^{-1} T^*A^*$$

where $\tilde{\Lambda}^{-1} q_4$ is 1 only if $p = q$, and is zero otherwise. Substituting these expressions into (15) gives the following expressions for $Q$ when $\eta = 0$:

$$Q = 2\mu^2 \sigma_4 \tilde{\Lambda}^{-1} \tilde{\Lambda}^{-2} W_cP_A^+W_cP_A^+$$

$$\odot (TW,\tilde{\Lambda}^{-2}W_cT^*)^T.$$  

(30)

As formalized in the theorem below, yet another set of optimal weightings results in minimum variance estimates for this error model.

**Theorem 3.2:** For the perturbed noise covariance model of (29), the covariance of the DOA estimates obtained from the SSF minimization of (4) satisfies

$$C(I, \tilde{\Lambda}) \leq C(W_c, W_c)$$

for all row and column weightings $W_c$ and $W_r$ when $\eta = 0$.

**Proof:** The proof is essentially identical to that given for Theorem 3.1.

The covariance of the estimation error when $W_c = I$ and $W_r = \tilde{\Lambda}$ simplifies to

$$C_{\text{RSF}} = \frac{\mu^2 \sigma_4}{2} \left\{ \text{Re}\{D^*P_A^+D\} \odot (\text{SA*AS})^T \right\}^{-1}$$

(31)
and the resulting SSF algorithm will be referred to as robust subspace fitting (RSF) for noise model errors. While (31) is the minimum possible covariance for multidimensional algorithms of the form (4), an additional step is needed to demonstrate that such a weighting scheme outperforms MUSIC. To show this, note that using 

as well.

and the resulting SSF algorithm will be referred to as robust subspace fitting (RSF) for noise model errors. While (31) is the minimum possible covariance for multidimensional algorithms of the form (4), an additional step is needed to demonstrate that such a weighting scheme outperforms MUSIC. To show this, note that using 

as well.

and the resulting SSF algorithm will be referred to as robust subspace fitting (RSF) for noise model errors. While (31) is the minimum possible covariance for multidimensional algorithms of the form (4), an additional step is needed to demonstrate that such a weighting scheme outperforms MUSIC. To show this, note that using 

as well.

and the resulting SSF algorithm will be referred to as robust subspace fitting (RSF) for noise model errors. While (31) is the minimum possible covariance for multidimensional algorithms of the form (4), an additional step is needed to demonstrate that such a weighting scheme outperforms MUSIC. To show this, note that using 

as well.

and the resulting SSF algorithm will be referred to as robust subspace fitting (RSF) for noise model errors. While (31) is the minimum possible covariance for multidimensional algorithms of the form (4), an additional step is needed to demonstrate that such a weighting scheme outperforms MUSIC. To show this, note that using 

as well.

and the resulting SSF algorithm will be referred to as robust subspace fitting (RSF) for noise model errors. While (31) is the minimum possible covariance for multidimensional algorithms of the form (4), an additional step is needed to demonstrate that such a weighting scheme outperforms MUSIC. To show this, note that using 

as well.

and the resulting SSF algorithm will be referred to as robust subspace fitting (RSF) for noise model errors. While (31) is the minimum possible covariance for multidimensional algorithms of the form (4), an additional step is needed to demonstrate that such a weighting scheme outperforms MUSIC. To show this, note that using 

as well.

and the resulting SSF algorithm will be referred to as robust subspace fitting (RSF) for noise model errors. While (31) is the minimum possible covariance for multidimensional algorithms of the form (4), an additional step is needed to demonstrate that such a weighting scheme outperforms MUSIC. To show this, note that using 

as well.
illustrate the performance advantage of the RSF algorithms. As in [1], finite sample effects are ignored by directly computing the exact perturbed covariance $\hat{R}$ for each Monte Carlo trial. A total of 1000 trials were conducted for each example, with $\hat{R}$ generated using the nominal covariance $R$ and the distribution of the perturbation $A$ or $\Sigma$. The sample RMS error of the DOA estimates was then calculated and compared to that predicted by the corresponding theoretical expressions. In all of the following examples, the nominal gain of all sensors was assumed to be unity in the direction of the impinging signals, and it was assumed that the number of emitters $d$ had been correctly determined.

**Example 1:** For this case, a 10 element uniform linear array (ULA) with half wavelength interelement spacing was assumed. Two emitters were simulated, one at 0° broadside and the other at 7°. The signal-to-noise ratio (SNR) for the 7° source was 20 dB, while that of the broadside source was varied between 0 and 12 dB, and the two sources were 90% correlated with 0° correlation phase of the first sensor. The covariance of (16) was assumed to be of the form $\nu^2 I$, with $\nu = 0.01$. This corresponds to a $-40$ dB gain error and a 0.57° phase error standard deviation. Fig. 1 shows a plot of the standard deviation of the 0° source estimates for MD-MUSIC, DML, WSF, and RSF using the weighting of (22) and an estimate of $T$ obtained from initial conditions generated using ESPRIT. The connected lines indicate the theoretical predictions and the symbols represent the results of the simulations. Note the excellent agreement between the predicted and measured DOA errors, and also the improvement achieved by the optimal RSF weighting relative to WSF and DML (in this case, $W_c = I$). Though not shown on the plot, the one-dimensional MUSIC algorithm has virtually identical performance to that of MD-MUSIC and RSF.

Fig. 2 shows the predicted finite sample performance of the algorithms for $N = 250$ snapshots under the assumption of no array errors. Note that the relative performance of the algorithms is reversed compared to Fig. 1; in this case WSF and DML outperform RSF and MD-MUSIC (only two curves seem to be present since the DOA error is virtually identical for WSF and ML, and also for RSF and MD-MUSIC). This phenomenon has been observed in other cases as well; i.e., in scenarios involving array errors where RSF has the greatest advantage over WSF, it has relatively poor finite sample performance. This illustrates the need for study of combined weightings that optimally account for both sources of error.

**Example 2:** As an example of a case where RSF outperforms MD-MUSIC, the algorithms were compared as a function of the separation between two sources. For variety, a 10 element circular array with 1 $\lambda$ radius and uniformly spaced elements was assumed. Two 30-dB uncorrelated sources were simulated, and the separation between the sources was varied from 2° to 20°. Array errors were generated using the model of (16)-(17) with covariance $(0.005)^2 I$, which corresponds to a $-46$ dB gain error and 0.3° phase error standard deviation. Fig. 3 shows the results of the simulation. In this case, the performance of WSF, DML, and RSF is virtually identical, while that of MD-MUSIC is significantly degraded as the sources move closer together. The predicted finite sample performance of WSF and RSF is also identical for this example, and is significantly better than that of MD-MUSIC.

**Example 3:** This example illustrates the effects of errors in the model for the noise covariance and demonstrates the advantage of using the optimal column weighting $W_{RSF} = \hat{\Lambda}^2$. A 12 element half wavelength spaced ULA and two correlated sources at 10° and 15° were simulated. The SNR for each source was 0 dB, and the standard deviation of the additive noise covariance perturbation
in (29) was $\mu = 0.01$. The actual and predicted RMS error performance of MUSIC, WSF, DML, and RSF are plotted for the source at $15^\circ$ versus signal correlation in Fig. 4. As predicted, the RSF algorithm with column weighting $W_r = \hat{A}^2$ achieves the lowest root mean square (rms) estimation error. Note that while the sensitivity of MUSIC increases dramatically at higher levels of correlation, the multidimensional methods remain relatively unaffected. In fact, the performance of WSF and RSF actually improves as the signals become more highly correlated. The finite sample performance of WSF and RSF were not significantly different in this example.

**Example 4:** The parameters of this simulation run were identical to that of example 1, except that $\nu$ was varied from 0.0001 to 0.4, and the performance of RSF and root-MUSIC was compared with that of ESPRIT. The results of this case are plotted in Fig. 5. Each of the three implementations of ESPRIT referenced in Fig. 5 corresponds to a different choice of identical subarrays. The variable $\Delta$ indicates the distance between the subarrays in terms of the wavelength $\lambda$. As the name implies, two interleaved ESPRIT subarrays are obtained by separating the even- and odd-numbered elements of the ULA. For overlapping subarrays, the first $m - k$ elements are grouped in one subarray, and the last $m - k$ elements in another. In the figure, $\Delta = \lambda / 2$ and $\Delta = 3\lambda / 2$ correspond to $k = 1$ and $k = 3$, respectively. Note also that, as predicted by Corollary 3.1, the performance of RSF and root-MUSIC is identical under this error model.

There are several important conclusions that can be drawn from this example. First, there is excellent agreement between predicted and measured algorithm performance. This is true even at the relatively large value of $\nu = 0.4$, which corresponds to a standard deviation in gain of 0.4 (relative to unity gain) and a standard deviation in phase of roughly 24°. Second, ESPRIT tends to degrade quite gracefully as the degree of perturbation increases. For example, at $\nu = 0.1$ corresponding to a gain and phase standard deviation of 0.1 and 6° (which is well beyond the tolerance of many commercial sensing devices), the standard deviation of the ESPRIT DOA estimates varies from only 0.2° to 1°, depending on the subarray separation. Performance improves as $\Delta$ increases because of the increased baseline between the subarrays. However, at larger values of $\Delta$, the diminished subarray size begins to play against this advantage and the DOA error will increase.

Another interesting observation is that, of the five algorithms implemented, root-MUSIC and RSF were the least sensitive to this type of sensor error, though ESPRIT with $\Delta = 3\lambda / 2$ is only slightly worse. It was shown in Section III that for this error model, RSF and MUSIC (and hence root-MUSIC) achieved the lowest possible estimate variance of all SSF algorithms for which $\eta = 0$. However, none of the results in Section III can be directly applied to ESPRIT since it assumes a different parameter-
ization for $\eta$. Although in this particular example it appears that the performance of ESPRIT is bounded below by the performance of RSF and MUSIC, there is as yet no analytical result which would preclude ESPRIT from having performance superior to these algorithms. Even if $C_{SE} \geq C_{ESP}$ in general, we see that in this case at least, the much more computationally efficient ESPRIT algorithm is able to achieve near optimal performance for an appropriate choice of subarrays.

REFERENCES


A. Lee Swindlehurst (S’83-M’84) was born in Boulder City, NV, in 1960. He received the B.S. (summa cum laude) and M.S. degrees in electrical engineering from Brigham Young University, Provo, UT, in 1985 and 1986, respectively, and the Ph.D. degree in electrical engineering from Stanford University, Stanford, CA, 1991.

From 1983 to 1984 he was employed with Eyring Research Institute of Provo, UT, as a Scientific Programmer writing software and documentation for a Minuteman missile simulation system. During 1984-1986, he was a Research Assistant in the Department of Electrical Engineering, at Brigham Young University. He was awarded an Office of Naval Research Graduate Fellowship for 1985-1988, and during that time was affiliated with the Information Systems Laboratory at Stanford University. From 1986 to 1990, he was also employed at ESL, Inc., Sunnyvale, CA, where he was involved in the design of algorithms and architectures for a variety of radar and sonar signal processing systems. He is currently an Assistant Professor in the Department of Electrical and Computer Engineering at Brigham Young University. His research interests include sensor array signal processing, detection and estimation theory, system identification and control.

Thomas Kailath (S’57-M’62-F’70) was educated in Poona, India, and received the S.M. and Sc. D. degrees from the Massachusetts Institute of Technology in 1959 and 1961, respectively. From October 1961 to December 1962, he worked with the Jet Propulsion Laboratories, Pasadena, CA, where he also taught part time at the California Institute of Technology. He joined Stanford University as an Associate Professor of Electrical Engineering in 1963. He has been a Full Professor since 1968, served as Director of the Information Systems Laboratory from 1971 through 1980, was Associate Department Chairman from 1981 to 1987, and currently holds the Hitachi America Professorship in Engineering. He has also held short-term appointments at several institutions around the world. He has worked in a number of areas including information theory, communications, computation, control, linear systems, statistical signal processing, stochastic processes, linear algebra, and operator theory; his recent research interests include array processing, fast algorithms for nonstationary signal processing, and the design of special purpose computing systems. He is the author of *Linear Systems* (Prentice-Hall, 1980), and *Lectures on Wiener and Kalman Filtering* (Springer-Verlag, 1981).

Dr. Kailath has held Guggenheim, Churchill, and Royal Society fellowships, among others, and received awards from the IEEE Information Theory Society, the IEEE Signal Processing Society, and the American Control Council. He served as President of the IEEE Information Theory Society in 1975, and was awarded an honorary doctorate from Linköping University, Sweden, in 1990. He is a Fellow of the Institute of Mathematical Statistics and is a member of the National Academy of Engineering.