ROBUST WEIGHTED SUBSPACE FITTING IN THE PRESENCE OF ARRAY MODEL ERRORS

Magnus Jansson[*], *A. Lee Swindlehurst*[†] *and Björn Ottersten*[*]

[*]Dept. of Signals, Sensors & Systems
Royal Inst. of Technology (KTH)
SE-100 44 Stockholm, SWEDEN

[†]Dept. of Electrical and Computer Eng.
Brigham Young University
Provo, UT 84602, USA

ABSTRACT

Model error sensitivity is an issue common to all high resolution direction of arrival estimators. Much attention has been directed to the design of algorithms for minimum variance estimation taking only finite sample errors into account. Approaches to reduce the sensitivity due to array calibration errors have also appeared in the literature. Herein, a weighted subspace fitting method for a wide class of array perturbation models is derived. This method provides minimum variance estimates under the assumption that the prior distribution of the perturbation model is known. Interestingly enough, the method reduces to the WSF (MODE) estimator if no model errors are present. On the other hand, when model errors dominate, the proposed method turns out to be equivalent to the "model-errors-only subspace fitting method". Unlike previous techniques for model errors, the estimator can be implemented using a two-step procedure if the nominal array is uniform and linear, and it is also consistent even if the signals are fully correlated.

1. INTRODUCTION

All signal parameter estimation methods in array signal processing rely on information about the array response, and assume that the signal wavefronts impinging on the array have perfect spatial coherence (e.g., perfect plane waves). Unfortunately, an array cannot be perfectly calibrated, and analytically derived array responses relying on the array geometry and wave propagation are at best good approximations. Therefore, the limiting factor in the performance of array signal processing algorithms is most often not measurement noise, but rather perturbations in the array response model. Depending on the size of such errors, estimates of the directions of arrival (DOAs) and the source signals may be significantly degraded. A number of studies have been conducted to quantify the performance degradation due to model errors for both DOA and signal estimation.

A number of techniques have also been considered for improving the robustness of array processing algorithms. In one such approach, the array response is parameterized not only by the DOAs of the signals, but also by perturbation or "nuisance" parameters that describe deviations of the response from its nominal value. These parameters can include, for example, displacements of the antenna elements from their nominal positions, uncalibrated receiver gain and phase offsets, etc.. With such a model, a natural approach is to attempt to estimate the unknown nuisance parameters simultaneously with the signal parameters. Such methods are referred to as *auto-calibration* techniques, and have been proposed by a number of authors (see, e.g., the references in [1]). When auto-calibration techniques are employed, it is critical to determine whether both the signal and nuisance parameters are identifiable. In certain cases they are not; for example, one cannot uniquely estimate both DOAs and sensor positions unless of course additional information is available, such as sources in known locations, cyclostationary signals with two or more known cycle frequencies, or partial information about the phase response of the array. The identifiability problem can be alleviated if the perturbation parameters are assumed to be drawn from some known *a priori* distribution. While this itself represents a form of additional information, it has the advantage of allowing an optimal maximum *a posteriori* (MAP) solution to the problem to be formulated [6]. In [6] it is shown that, by using an asymptotically equivalent approximation to the resulting MAP criterion, the estimation of the signal and nuisance parameters can be decoupled, leading to a significant simplification of the problem. The MAP approach in [6] is asymptotically statistically efficient for very general error models. However, since it is implemented by means of *noise* subspace fitting [2], if the sources are highly correlated or closely spaced in angle, its finite sample performance may be poor. In fact, the method of [6] is not a consistent estimator of the DOAs if the signals are perfectly coherent.

In this paper, we develop a statistically efficient weighted *signal* subspace fitting (SSF) algorithm that holds for very general array perturbation models, that has much better finite sample performance than [6] when the signals are highly correlated or closely spaced, and that yields consistent estimates when coherent sources are present. An additional advantage of our SSF formulation is that if the array is nominally uniform and linear, a two-step procedure similar to that for MODE [3] can be used to eliminate the search for the DOAs.

2. PROBLEM FORMULATION

Assume that the output of an array of m sensors is given by the model

$$\mathbf{x}(t) = \mathbf{A}(\boldsymbol{\theta}, \boldsymbol{\rho})\mathbf{s}(t) + \mathbf{n}(t),$$

where $\mathbf{s}(t)$ is a complex d-vector containing the emitted signal waveforms and $\mathbf{n}(t)$ is an additive noise vector. The array steering matrix is defined as

$$\mathbf{A}(\boldsymbol{\theta}, \boldsymbol{\rho}) = \left[\bar{\mathbf{a}}(\boldsymbol{\theta}_1, \boldsymbol{\rho}) \quad \ldots \quad \bar{\mathbf{a}}(\boldsymbol{\theta}_d, \boldsymbol{\rho})\right],$$

A. L. Swindlehurst acknowledges support by the National Science Foundation under grant MIP-9408154.

where $\bar{\mathbf{a}}(\boldsymbol{\theta}_i, \boldsymbol{\rho})$ denotes the array response to a unit waveform associated with the signal parameter $\boldsymbol{\theta}_i$ (possibly vector valued, although we will specialize to the scalar case). The parameters in the real n-vector $\boldsymbol{\rho}$ are used to model the uncertainty in the array steering matrix. It is assumed that $\mathbf{A}(\boldsymbol{\theta}, \boldsymbol{\rho})$ is known for a nominal value $\boldsymbol{\rho} = \boldsymbol{\rho}_0$ and that the columns in $\mathbf{A}(\boldsymbol{\theta}, \boldsymbol{\rho}_0)$ are linearly independent as long as $\boldsymbol{\theta}_i \neq \boldsymbol{\theta}_j,\ i \neq j$. The model for $\mathbf{A}(\boldsymbol{\theta}, \boldsymbol{\rho}_0)$ can be obtained for example by physical insight, or it could be the result of a calibration experiment. One may then have knowledge about the sensitivity of the nominal response to certain variations in $\boldsymbol{\rho}$, which can be modeled by considering $\boldsymbol{\rho}$ as a random vector. The *a priori* information is then in the form of a probability distribution which can be used in the design of an estimation algorithm to make it robust with regard to the model errors.

In this paper we will use a stochastic model for the signals; more precisely, $\mathbf{s}(t)$ is considered to be a zero-mean Gaussian random vector with second moments

$$\mathrm{E}\{\mathbf{s}(t)\mathbf{s}^*(s)\} = \mathbf{P}\delta_{t,s}, \qquad \mathrm{E}\{\mathbf{s}(t)\mathbf{s}^T(s)\} = \mathbf{0},$$

where $(\cdot)^*$ denotes complex conjugate transpose, $(\cdot)^T$ denotes transpose and $\delta_{t,s}$ is the Kronecker delta. Let d' denote the rank of the signal covariance matrix $\mathbf{P}$. Special attention will be given to cases where $d' < d$, i.e., cases in which the signals may be fully correlated (coherent). In particular, we show that the proposed method is consistent even under such severe conditions. The noise is modeled as a zero-mean spatially and temporally white complex Gaussian vector with second order moments

$$\mathrm{E}\{\mathbf{n}(t)\mathbf{n}^*(s)\} = \sigma^2\mathbf{I}\delta_{t,s}, \qquad \mathrm{E}\{\mathbf{n}(t)\mathbf{n}^T(s)\} = \mathbf{0}.$$

The perturbation parameter vector $\boldsymbol{\rho}$ is modeled as a Gaussian random variable with mean $\mathrm{E}\{\boldsymbol{\rho}\} = \boldsymbol{\rho}_0$ and covariance

$$\mathrm{E}\{(\boldsymbol{\rho} - \boldsymbol{\rho}_0)(\boldsymbol{\rho} - \boldsymbol{\rho}_0)^T\} = \boldsymbol{\Omega}.$$

It is assumed that both $\boldsymbol{\rho}_0$ and $\boldsymbol{\Omega}$ are known. Similar to [6, 7], we consider small perturbations in $\boldsymbol{\rho}$ and more specifically assume that the effect of the array errors on the estimates of $\boldsymbol{\theta}$ is of comparable size to those due to the finite sample effects of the noise. To model this, we assume that $\boldsymbol{\Omega} = \bar{\boldsymbol{\Omega}}/N$, where N is the number of samples and $\bar{\boldsymbol{\Omega}}$ is independent of N. This somewhat artificial assumption concerning the model errors is made only for convenience in showing the statistical optimality of the algorithm presented later. An identical result can be obtained for small $\boldsymbol{\rho} - \boldsymbol{\rho}_0$ if a first order perturbation analysis is used, but this approach is somewhat less elegant.

3. GENERALIZED WEIGHTED SUBSPACE FITTING

In [6], the so called MAP-NSF method was derived and shown to be asymptotically statistically efficient. One drawback with MAP-NSF is that it is not consistent if the sources are fully correlated, that is, if $d' < d$. This problem can be overcome in the *signal* subspace formulation. In this section, we derive a new method in the signal subspace fitting (SSF) class of algorithms. The method will be shown to possess the same large sample performance as MAP-NSF and is thus also asymptotically efficient. A further advantage of the proposed SSF approach is that the cost function depends on the parameters in a relatively simple manner and, if the nominal array is uniform and linear, the non-linear minimization can be replaced by a rooting technique (see Section 5).

First, let us introduce the eigendecomposition of the sample covariance matrix

$$\hat{\mathbf{R}} \triangleq \frac{1}{N}\sum_{t=1}^{N}\mathbf{x}(t)\mathbf{x}^*(t) = \hat{\mathbf{E}}_s\hat{\boldsymbol{\Lambda}}_s\hat{\mathbf{E}}_s^* + \hat{\mathbf{E}}_n\hat{\boldsymbol{\Lambda}}_n\hat{\mathbf{E}}_n^*. \qquad (1)$$

Here, $\hat{\boldsymbol{\Lambda}}_s$ and $\hat{\boldsymbol{\Lambda}}_n$ are diagonal matrices containing the d' largest and, respectively, the $m - d'$ smallest eigenvalues, and $\hat{\mathbf{E}}_s$ and $\hat{\mathbf{E}}_n$ are composed of the corresponding eigenvectors.

According to the problem formulation, $\hat{\mathbf{R}}$ converges to

$$\mathbf{R} = \mathrm{E}\{\mathbf{x}(t)\mathbf{x}^*(t)\} = \mathbf{A}_0\mathbf{P}\mathbf{A}_0^* + \sigma^2\mathbf{I}$$

as N tends to infinity, where $\mathbf{A}_0 = \mathbf{A}(\boldsymbol{\theta}_0, \boldsymbol{\rho}_0)$, and $\boldsymbol{\theta}_0$ denotes the true DOA vector. The eigendecomposition of $\mathbf{R}$ is defined similar to (1) using a notation without "hats." This implies that $\hat{\mathbf{E}}_s \to \mathbf{E}_s = \mathbf{A}_0\mathbf{T}$ as $N \to \infty$, where $\mathbf{T}$ is a $d \times d'$ full rank matrix. The idea of the proposed method is to minimize a suitable norm of the residuals

$$\boldsymbol{\varepsilon} = \mathrm{vec}(\mathbf{B}^*(\boldsymbol{\theta})\hat{\mathbf{E}}_s),$$

where $\mathbf{B}(\boldsymbol{\theta})$ is an $m \times (m - d)$ full rank matrix whose columns span the null-space of $\mathbf{A}^*(\boldsymbol{\theta})$, and $\mathrm{vec}(\cdot)$ is the vectorization operation. This implies that $\mathbf{B}^*(\boldsymbol{\theta})\mathbf{A}(\boldsymbol{\theta}) = \mathbf{0}$ and $\mathbf{B}^*(\boldsymbol{\theta}_0)\mathbf{E}_s = \mathbf{0}$.

Remark 1. *For general arrays, there is no known closed form parameterization of $\mathbf{B}$ in terms of $\boldsymbol{\theta}$. However, this will not introduce any problems since the final criterion can be written as an explicit function of $\boldsymbol{\theta}$ (see Section 5 and [1]).*

For general array error models, we have to consider the real and imaginary parts of $\boldsymbol{\varepsilon}$ separately to get optimal (minimum variance) performance. Equivalently, we may study $\boldsymbol{\varepsilon}$ and its complex conjugate, which we denote by $\boldsymbol{\varepsilon}^c$.

We propose to estimate the signal parameters $\boldsymbol{\theta}$ as follows:

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} V_{\mathrm{GWSF}}(\boldsymbol{\theta}), \qquad (2)$$

$$V_{\mathrm{GWSF}}(\boldsymbol{\theta}) = \bar{\boldsymbol{\varepsilon}}^*(\boldsymbol{\theta})\mathbf{W}\bar{\boldsymbol{\varepsilon}}(\boldsymbol{\theta}), \qquad (3)$$

where $\bar{\boldsymbol{\varepsilon}} = \begin{bmatrix}\boldsymbol{\varepsilon}^* & \boldsymbol{\varepsilon}^T\end{bmatrix}^*$, and $\mathbf{W}$ is a positive definite weighting matrix. The method will in the sequel be referred to as the *generalized weighted subspace fitting (GWSF)* method. To derive the weighting that leads to minimum variance estimates of $\boldsymbol{\theta}$, we need to compute the residual covariance matrix. As shown in [1], the (asymptotic) second order moment of the residual vector $\bar{\boldsymbol{\varepsilon}}$ at $\boldsymbol{\theta}_0$ can be written as

$$\mathbf{C}_{\bar{\varepsilon}} \triangleq \lim_{N\to\infty} N\ \mathrm{E}\{\bar{\boldsymbol{\varepsilon}}\bar{\boldsymbol{\varepsilon}}^*\} = \bar{\mathbf{L}} + \bar{\mathbf{G}}\bar{\mathbf{G}}^*,$$

where we have defined

$$\bar{\mathbf{L}} = \begin{bmatrix}\mathbf{L} & \mathbf{0}\\ \mathbf{0} & \mathbf{L}^c\end{bmatrix}, \qquad \mathbf{L} = \left(\sigma^2\tilde{\boldsymbol{\Lambda}}^{-2}\boldsymbol{\Lambda}_s \otimes \mathbf{B}^*\mathbf{B}\right),$$

$$\bar{\mathbf{G}} = \begin{bmatrix}\mathbf{G}\\ \mathbf{G}^c\end{bmatrix}, \qquad \mathbf{G} = \left(\mathbf{T}^T \otimes \mathbf{B}^*\right)\mathbf{D}_{\boldsymbol{\rho}}\bar{\boldsymbol{\Omega}}^{1/2},$$

and

$$\mathbf{D}_{\boldsymbol{\rho}} = \left[\frac{\partial\mathbf{a}(\boldsymbol{\theta},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}_1} \quad \dots \quad \frac{\partial\mathbf{a}(\boldsymbol{\theta},\boldsymbol{\rho})}{\partial\boldsymbol{\rho}_n}\right],$$

$$\mathbf{a}(\boldsymbol{\theta}, \boldsymbol{\rho}) = \left[\bar{\mathbf{a}}^T(\boldsymbol{\theta}_1, \boldsymbol{\rho}) \quad \dots \quad \bar{\mathbf{a}}^T(\boldsymbol{\theta}_d, \boldsymbol{\rho})\right]^T.$$

Here, $\bar{\mathbf{\Omega}}^{1/2}$ is a (symmetric) square root of $\bar{\mathbf{\Omega}}$, $\widetilde{\mathbf{\Lambda}} = \mathbf{\Lambda}_s - \sigma^2\mathbf{I}$, $\mathbf{T} = \mathbf{A}_0^\dagger \mathbf{E}_s$, and $\otimes$ denotes Kronecker product. It is easy to see that $\mathbf{C}_{\bar{\varepsilon}}$ is positive definite since $\mathbf{L}$ is. It is then well known that the optimal choice of the weighting in terms of minimizing the parameter estimation error variance is

$$\mathbf{W}_{\text{GWSF}} = \mathbf{C}_{\bar{\varepsilon}}^{-1}. \tag{4}$$

The implementation of GWSF is discussed in Section 5. We conclude this section with two remarks regarding the GWSF formulation.

Remark 2. *If there are no model errors, then GWSF reduces to the WSF/MODE estimator [5]. This can easily be verified by setting* $\bar{\mathbf{\Omega}} = \mathbf{0}$ *and rewriting the GWSF criterion* (3)*.*

Remark 3. *For the case of model errors only (*$N \to \infty$ *or* $\sigma^2 \to 0$*), GWSF becomes the "model errors only" algorithm [4]. The results obtained for GWSF are also consistent with weightings given in [7] for special array error models.*

Thus, the GWSF method can be considered to be optimal in general, and not just when the model errors and the finite sample effects are of the same order.

4. PERFORMANCE ANALYSIS

In [1], the asymptotic properties of the GWSF estimates are analyzed. It is shown that the estimates are consistent and have a limiting Gaussian distribution. It is also shown that the asymptotic covariance matrix of the estimation error is equal to the CRB for the problem under study [6, 8]. In summary we have the following result.

Theorem 1. *If* $d < (m+d')/2$*, then the GWSF estimate* (2) *tends to* $\boldsymbol{\theta}_0$ *w.p.1 as* $N \to \infty$*, and*

$$\sqrt{N}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) \in \text{AsN}(\mathbf{0}, N\mathbf{CRB}_{\boldsymbol{\theta}}),$$

where $\text{AsN}(\cdot)$ *means asymptotically Gaussian distributed, and*

$$\mathbf{CRB}_{\boldsymbol{\theta}} \triangleq \frac{\sigma^2}{2N}\left[\mathbf{C} - \mathbf{F}_{\boldsymbol{\theta}}^T \mathbf{\Gamma}^{-1}\mathbf{F}_{\boldsymbol{\theta}}\right]^{-1}, \tag{5}$$

where

$$\begin{aligned}
\mathbf{C} &= \text{Re}\{\mathbf{D}_{\boldsymbol{\theta}}^* \mathbf{M} \mathbf{D}_{\boldsymbol{\theta}}\},\\
\mathbf{M} &= \mathbf{U}^T \otimes \mathbf{\Pi}_{\mathbf{A}}^{\perp},\\
\mathbf{U} &= \mathbf{A}^\dagger \mathbf{E}_s \widetilde{\mathbf{\Lambda}}^2 \mathbf{\Lambda}_s^{-1} \mathbf{E}_s^* \mathbf{A}^{\dagger *},\\
\mathbf{\Pi}_{\mathbf{A}}^{\perp} &= \mathbf{I} - \mathbf{A}(\mathbf{A}^*\mathbf{A})^{-1}\mathbf{A}^*,\\
\mathbf{D}_{\boldsymbol{\theta}} &= \left[\frac{\partial \mathbf{a}(\boldsymbol{\theta},\boldsymbol{\rho})}{\partial \boldsymbol{\theta}_1} \quad \cdots \quad \frac{\partial \mathbf{a}(\boldsymbol{\theta},\boldsymbol{\rho})}{\partial \boldsymbol{\theta}_d}\right],\\
\mathbf{F}_{\boldsymbol{\theta}} &= \text{Re}\{\mathbf{D}_{\boldsymbol{\rho}}^* \mathbf{M} \mathbf{D}_{\boldsymbol{\theta}}\},\\
\mathbf{\Gamma} &= \text{Re}\{\mathbf{D}_{\boldsymbol{\rho}}^* \mathbf{M} \mathbf{D}_{\boldsymbol{\rho}} + \frac{\sigma^2}{2}\bar{\mathbf{\Omega}}^{-1}\}.
\end{aligned}$$

The above expressions are evaluated at $\boldsymbol{\theta}_0$ *and* $\boldsymbol{\rho}_0$*.*

Notice that the CRB for the case with no model errors is $\sigma^2\mathbf{C}^{-1}/2N$. The result of the theorem is similar to what was derived for the MAP-NSF method in [6]. It implies that GWSF and MAP-NSF are asymptotically equivalent and efficient for large N and small $\mathbf{\Omega}$ ($= \bar{\mathbf{\Omega}}/N$). An advantage of GWSF as compared to MAP-NSF is that GWSF is consistent even if $\mathbf{P}$ is rank deficient ($d' < d$). Another advantage is that the minimization of $V_{\text{GWSF}}(\boldsymbol{\theta})$ can be performed without a search if the nominal array is uniform and linear. GWSF can be seen as a generalization of WSF [5] and MODE [3] that allows one to include *a priori* knowledge of the array perturbations into the estimation criterion.

5. IMPLEMENTATION

Here, we only sketch the idea behind the implementation of GWSF for uniform linear arrays. A detailed description is given in [1]. Define the $m \times (m-d)$ Toeplitz matrix $\mathbf{B}$ as

$$\mathbf{B}^* = \begin{bmatrix} b_d & \cdots & b_1 & b_0 & 0 & \cdots & 0 \\ 0 & b_d & \cdots & b_1 & b_0 & \ddots & 0 \\ \vdots & \ddots & \ddots & & \ddots & \ddots & \vdots \\ 0 & \ldots & 0 & b_d & \ldots & b_1 & b_0 \end{bmatrix},$$

where $\{b_i\}$ are defined by

$$b_0 z^d + b_1 z^{d-1} + \cdots + b_d = b_0 \prod_{k=1}^{d}(z - e^{j\omega_k}), \qquad b_0 \neq 0. \tag{6}$$

Here, $\omega_k = 2\pi\delta\sin(\boldsymbol{\theta}_k)$, where δ is the separation between the sensors measured in wavelengths. The idea is to re-parameterize the minimization problem in terms of the polynomial coefficients instead of $\boldsymbol{\theta}$. As will be seen below, this leads to a considerable computational simplification. It is readily verified that $\mathbf{B}^*\mathbf{A}(\boldsymbol{\theta}) = \mathbf{0}$, and since the rank of $\mathbf{B}$ is $m-d$, it follows that the columns of $\mathbf{B}$ span the null-space of $\mathbf{A}^*$. Observe that $\bar{\boldsymbol{\varepsilon}}$ is linear in $\{b_i\}$ and, hence, the minimization in (2) is a quadratic problem for a fixed $\mathbf{W}$. Schematically, GWSF is implemented in the following steps:

1. Minimize $\bar{\boldsymbol{\varepsilon}}^*\bar{\boldsymbol{\varepsilon}}$ to get an initial estimate of the polynomial coefficients $\{b_i\}$ and estimate $\boldsymbol{\theta}$ from the roots of the polynomial in (6).
2. Compute a consistent estimate of the weighting matrix $\mathbf{W}_{\text{GWSF}}$ in (4) based on sample data and $\hat{\boldsymbol{\theta}}$ from the first step. Minimize the quadratic function

   $$\bar{\boldsymbol{\varepsilon}}^* \mathbf{W}_{\text{GWSF}} \bar{\boldsymbol{\varepsilon}}$$

   to get refined estimates of $\{b_i\}$ and obtain the DOA estimates by rooting the polynomial in (6).

Thus, GWSF can be implemented in a very attractive manner if the nominal array is uniform and linear. In particular, the solution can be obtained in a "closed form" by solving two quadratic problems and rooting the polynomial in (6) (see [1] for the details). There is no need for an iterative optimization procedure like that necessary in, for example, MAP-NSF.

6. SIMULATION EXAMPLE

In this section, we illustrate the findings of this paper by means of a simulation example. MAPprox is an alternative method for the problem considered in this paper and was first proposed in [8]. A version of MAPprox, MAPprox2, is analyzed in [1]. Below, the MAP-NSF, MAPprox2 and GWSF methods are compared with

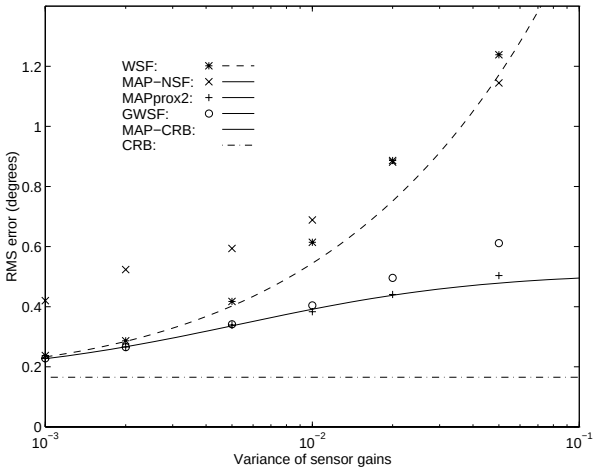

Figure 1: RMS errors for $\boldsymbol{\theta}_1$ versus the variance of the uncertainty in the gain of the sensors.

one another, as well as with methods like WSF/MODE that do not take the array perturbations into account.

Consider a uniform linear array consisting of $m = 10$ sensors separated by a half wavelength. Two signals impinge from the directions $\boldsymbol{\theta}_0 = [0^\circ \ 5^\circ]^T$ relative to broadside. The signals are uncorrelated and the SNR is 5 dB: $\mathbf{P}/\sigma^2 = 10^{5/10}\mathbf{I}_2$. The nominal unit gain sensors are perturbed by additive zero-mean Gaussian random variables. The sample size is fixed to $N = 200$ and the variance of the gain uncertainty is varied. In Figure 1, the RMS errors for $\boldsymbol{\theta}_1$ are plotted versus the variance of the sensor gains. (Only the RMS values for $\boldsymbol{\theta}_1$ are displayed; the results corresponding to $\boldsymbol{\theta}_2$ are similar.) In Figure 1, we show simulation results using different symbols, while the theoretical results are displayed by lines. The empirical RMS values are computed from 1000 independent trials. The curve denoted by CRB in Figure 1 is the Cramér-Rao lower bound for the case with no model errors and the MAP-CRB refers to (5). The MAP-NSF and MAPprox2 cost functions are minimized with Newton-type methods initialized by the WSF estimate. WSF (or, equivalently, MODE [3]) is implemented in the "rooting form" [3]. The GWSF method is implemented as described in [1]. The poor performance of MAP-NSF in Figure 1 is due to two main reasons. First, and most important, MAP-NSF fails to resolve the two signals in many cases. Second, the numerical search is sometimes trapped in a local minima. The performance of GWSF and MAPprox2 is excellent and close to the accuracy predicted by the asymptotic analysis. However, the numerical minimization of the MAPprox2 cost function is complicated. In this example, MAPprox2 produced "outliers" in about five percent of the 1000 trials. These outliers were removed before the RMS value for MAPprox2 was calculated. For small model errors, GWSF and WSF have a similar performance since the finite sample effects dominate. However, when calibration errors affect the performance, it can be seen that GWSF and MAPprox2 are significantly better than the standard WSF method. Notice that GWSF and MAPprox2 begin to differ for large model errors. The three methods (GWSF, MAPprox2 and MAP-NSF) are known to be asymptotically efficient estimators for large N and small model errors only; the performance for large model errors is not predicted by the theory in this paper.

7. CONCLUSIONS

This paper has studied the problem of developing robust weighted signal subspace fitting algorithms for arrays with calibration errors. It was shown that if the second-order statistics of the calibration errors are known, then asymptotically statistically efficient weightings can be derived for very general perturbation models. Earlier research had resulted in optimal weightings that were applicable only for very special cases. The GWSF technique derived herein unifies earlier work and also enjoys several advantages over the MAP-NSF approach, another statistically efficient technique developed for calibration errors. In particular, since GWSF is based on the signal rather than noise subspace, it is a consistent estimator even when the signals are perfectly coherent, and has better finite sample performance than MAP-NSF when the signals are highly correlated or closely spaced. In addition, unlike MAP-NSF, when the array is nominally uniform and linear, the GWSF criterion can be re-parameterized in such a way that the directions of arrival may be solved for by rooting a polynomial rather than via a gradient search. Our simulations indicate that, in the presence of calibration errors with known statistics, both algorithms can yield a significant performance improvement over techniques that ignore such errors.

8. REFERENCES

[1] M. Jansson, A. L. Swindlehurst, and B. Ottersten. "Weighted Subspace Fitting for General Array Error Models". Technical Report IR-S3-REG-9713, Dept. of Signals, Sensors & Systems, KTH, Stockholm, Aug. 1997. Submitted to IEEE Trans. on Signal Processing, Available from: http://www.s3.kth.se/control/.

[2] B. Ottersten, M. Viberg, P. Stoica, and A. Nehorai. "Exact and Large Sample ML Techniques for Parameter Estimation and Detection in Array Processing". In Haykin, Litva, and Shepherd, editors, *Radar Array Processing*, pages 99–151. Springer-Verlag, Berlin, 1993.

[3] P. Stoica and K. Sharman. "Maximum Likelihood Methods for Direction-of-Arrival Estimation". *IEEE Trans. ASSP*, ASSP-38:1132–1143, July 1990.

[4] A. Swindlehurst and T. Kailath. "A Performance Analysis of Subspace-Based Methods in the Presence of Model Errors – Part 2: Multidimensional Algorithms". *IEEE Trans. on Signal Processing*, 41(9):2882–2890, Sept. 1993.

[5] M. Viberg and B. Ottersten. "Sensor Array Processing Based on Subspace Fitting". *IEEE Trans. on Signal Processing*, SP-39(5):1110–1121, May 1991.

[6] M. Viberg and A. L. Swindlehurst. "A Bayesian Approach to Auto-Calibration for Parametric Array Signal Processing". *IEEE Trans. on Signal Processing*, 42(12):3495–3507, Dec. 1994.

[7] M. Viberg and A. L. Swindlehurst. "Analysis of the Combined Effects of Finite Samples and Model Errors on Array Processing Performance". *IEEE Trans. on Signal Processing*, 42(11):3073–3083, Nov. 1994.

[8] B. Wahlberg, B. Ottersten, and M. Viberg. "Robust Signal Parameter Estimation in the Presence of Array Perturbations". In *Proc. IEEE ICASSP*, pages 3277–3280, Toronto, Canada, 1991.