ANALYSIS OF A SUBSPACE-BASED SPATIAL FREQUENCY ESTIMATOR

M. Jansson *B. Göransson* *B. Ottersten*

Dept. of Signals, Sensors & Systems, Royal Institute of Technology (KTH), S-100 44 Stockholm, Sweden
B. Ottersten is currently on leave at ArrayComm Inc., 3141 Zanker Rd, San Jose, CA 95134.

ABSTRACT

In a previous paper we have presented a novel method for spatial and temporal frequency estimation assuming that the sources are uncorrelated. The current contribution analyzes this method in the case of spatial frequency estimation. In particular an optimal weighting matrix is derived and it is shown that the asymptotic variance of the frequency estimates coincides with the relevant Cramér-Rao lower bound. This means that the estimator is in large samples an efficient subspace-based spatial frequency estimator. The proposed method thus utilizes the a priori knowledge about the signal correlation as opposed to previously known subspace estimators. Moreover, when a uniform linear array is employed, it is possible to implement the estimator in a non-iterative fashion.

1. INTRODUCTION

Estimating frequencies from uniformly sampled data has been an active research area for decades. A number of, so called, high resolution algorithms or eigenstructure methods have been presented and analyzed in the literature, e.g., [1–4]. One disadvantage with these subspace based methods is that it is difficult to incorporate knowledge of the source correlation into the eigendecomposition. In [5] we proposed an estimator which combines ideas from subspace and covariance matching methods. The objective was to find a frequency estimator which uses the knowledge of the signal correlation. This method uses the geometrical properties of the eigendecomposition of the data covariance matrix and is valid for a large class of problems such as spatial and temporal frequency estimation. In this paper we specialize to spatial frequency or direction of arrival estimation and the large sample performance is analyzed.

2. MODEL DESCRIPTION

The well known problem of estimating spatial frequencies from uniformly sampled data corrupted by additive white noise can be reduced to the problem of determining the parameters in the following model of the data covariance matrix

$$\mathbf{R} = \mathbf{A}(\boldsymbol{\omega})\mathbf{S}\mathbf{A}^*(\boldsymbol{\omega}) + \sigma^2\mathbf{I} \ . \tag{1}$$

The $d \times d$ matrix $\mathbf{S}$ denotes the unknown diagonal signal covariance matrix, σ^2 is the unknown noise variance and the $m \times d$ matrix $\mathbf{A}(\boldsymbol{\omega})$ is the sensor array steering matrix where $m > d$ denotes the number of sensors. If a uniform linear array (ULA) of identical omni-directional sensors is employed, $\mathbf{A}(\boldsymbol{\omega})$ takes the form

$$\mathbf{A}(\boldsymbol{\omega}) = \begin{pmatrix} 1 & \cdots & 1 \\ e^{i\omega_1} & \cdots & e^{i\omega_d} \\ \vdots & & \vdots \\ e^{i(m-1)\omega_1} & \cdots & e^{i(m-1)\omega_d} \end{pmatrix} \tag{2}$$

where $\boldsymbol{\omega} = [\omega_1, \cdots, \omega_d]^{\mathrm{T}}$. Notice that $\mathbf{A}$ is full column rank when the frequencies are distinct. In the spatial frequency estimation problem, the matrix $\mathbf{A}$ is often parameterized by the direction of arrivals (DOAs) denoted by $\boldsymbol{\theta}$. For a ULA, the relationship between $\boldsymbol{\omega}$ and $\boldsymbol{\theta}$ is given by $\omega_k = 2\pi\Delta\sin(\theta_k)$, where Δ is the element spacing measured in wavelengths, and where θ_k is measured relative to the array broadside. Throughout the paper it is assumed that each column in $\mathbf{A}$ depends on a single parameter.

3. FREQUENCY ESTIMATION

This paper focuses on the estimation of the frequencies $\boldsymbol{\omega} = [\omega_1, \cdots, \omega_d]^{\mathrm{T}}$. In doing this we would like to use the knowledge that the signals are uncorrelated to improve estimation accuracy.

The subspace estimation techniques rely on the properties of the eigendecomposition of (1). Let

$$\mathbf{R} = \mathbf{E}_s\boldsymbol{\Lambda}_s\mathbf{E}_s^* + \mathbf{E}_n\boldsymbol{\Lambda}_n\mathbf{E}_n^* \ . \tag{3}$$

be a partitioned eigendecomposition, where $\boldsymbol{\Lambda}_s$ is a diagonal matrix containing the d largest eigenvalues and where the columns of $\mathbf{E}_s$ are the corresponding eigenvectors. Similarly, $\boldsymbol{\Lambda}_n$ contains the $m-d$ smallest eigenvalues and $\mathbf{E}_n$ is built of the remaining eigenvectors. Since $\mathbf{A}$ is assumed to be full rank and since $\mathbf{S}$ is positive definite, it follows that $\boldsymbol{\Lambda}_n = \sigma^2\mathbf{I}$. Using the fact $\mathbf{E}_n\mathbf{E}_n^* = \mathbf{I} - \mathbf{E}_s\mathbf{E}_s^*$, it follows from (1) and (3) that

$$\mathbf{A}\mathbf{S}\mathbf{A}^* = \mathbf{E}_s\boldsymbol{\Lambda}\mathbf{E}_s^* \ , \tag{4}$$

where $\boldsymbol{\Lambda} = \boldsymbol{\Lambda}_s - \sigma^2\mathbf{I}$. By using the vec-operator ($\mathrm{vec}(\mathbf{D})$ is a vector obtained by stacking the columns of $\mathbf{D}$), (4) can be written as $\left(\mathrm{vec}(\mathbf{XYZ}) = \left(\mathbf{Z}^{\mathrm{T}} \otimes \mathbf{X}\right)\mathrm{vec}(\mathbf{Y})\right)$

$$\left(\mathbf{A}^c \otimes \mathbf{A}\right)\mathrm{vec}(\mathbf{S}) = \left(\mathbf{E}_s^c \otimes \mathbf{E}_s\right)\mathrm{vec}(\boldsymbol{\Lambda}), \tag{5}$$

where $\otimes$ denotes the Kronecker matrix product, and where $(\cdot)^c$ denotes complex conjugation. Since $\mathbf{S}$ and $\boldsymbol{\Lambda}$ are diagonal matrices, there exists a $(d^2 \times d)$ selection matrix $\mathbf{L}$ such that $\mathrm{vec}(\mathbf{S}) = \mathbf{L}\mathbf{s}$ and $\mathrm{vec}(\boldsymbol{\Lambda}) = \mathbf{L}\boldsymbol{\lambda}$, where $\mathbf{s}$ and $\boldsymbol{\lambda}$ are vectors consisting of the diagonal entries of $\mathbf{S}$ and $\boldsymbol{\Lambda}$, respectively. Notice that $\left(\mathbf{A}^c \otimes \mathbf{A}\right)\mathbf{L} = \left(\mathbf{A}^c \circ \mathbf{A}\right)$ where $\circ$ denotes the Khatri-Rao matrix product which is column-wise Kronecker product.

Let $\hat{\mathbf{R}}$ denote the usual sample estimate of the theoretical covariance matrix, i.e., the average of the outer products of the array output vectors, and let

$$\hat{\mathbf{R}} = \hat{\mathbf{E}}_s \hat{\mathbf{\Lambda}}_s \hat{\mathbf{E}}_s^* + \hat{\mathbf{E}}_n \hat{\mathbf{\Lambda}}_n \hat{\mathbf{E}}_n^* \tag{6}$$

be the eigendecomposition of $\hat{\mathbf{R}}$ similar to (3). Replacing $\mathbf{E}_s$ by $\hat{\mathbf{E}}_s$ and $\mathbf{\Lambda}$ by $\hat{\mathbf{\Lambda}}_s - \hat{\sigma}^2\mathbf{I}$, where $\hat{\sigma}^2 = \mathrm{Tr}\{\hat{\mathbf{\Lambda}}_n\}/(m-d)$, in (5) yields

$$(\mathbf{A}^c \circ \mathbf{A})\,\mathbf{s} \approx \left(\hat{\mathbf{E}}_s^c \circ \hat{\mathbf{E}}_s\right)\hat{\boldsymbol{\lambda}} \tag{7}$$

or with obvious definitions

$$\mathbf{B}(\boldsymbol{\omega})\mathbf{s} \approx \hat{\mathbf{f}}\,. \tag{8}$$

The least-squares estimate of $\mathbf{s}$ is

$$\hat{\mathbf{s}} = \mathbf{B}^{\dagger}(\boldsymbol{\omega})\hat{\mathbf{f}}\,, \tag{9}$$

where $\mathbf{B}^{\dagger}$ denotes the Moore-Penrose pseudo-inverse of $\mathbf{B}$. We now suggest to estimate the frequencies by minimizing the weighted norm of the residuals obtained by substituting $\hat{\mathbf{s}}$ back into (8), that is,

$$\hat{\boldsymbol{\omega}} = \arg\min_{\boldsymbol{\omega}} V(\boldsymbol{\omega}), \tag{10}$$

$$V(\boldsymbol{\omega}) = \|\mathbf{P}_{\mathbf{B}}^{\perp}(\boldsymbol{\omega})\hat{\mathbf{f}}\|_{\mathbf{W}}^2 = \hat{\mathbf{f}}^*\mathbf{P}_{\mathbf{B}}^{\perp}(\boldsymbol{\omega})\mathbf{W}\mathbf{P}_{\mathbf{B}}^{\perp}(\boldsymbol{\omega})\hat{\mathbf{f}}\,, \tag{11}$$

where $\mathbf{P}_{\mathbf{B}}^{\perp} = \mathbf{I} - \mathbf{B}\mathbf{B}^{\dagger}$ is the orthogonal projector onto the null-space of $\mathbf{B}^*$. In (11), $\mathbf{W}$ is a Hermitian positive definite weighting matrix.

In the next section the asymptotic properties of the estimates given by (10) are analyzed. The implementation of the estimator is discussed in Section 5.

4. ANALYSIS

The asymptotic behavior of the estimates (10) is analyzed in this section. To simplify notation, we write $\mathbf{P}^{\perp}$ in lieu of $\mathbf{P}_{\mathbf{B}}^{\perp}$. We also use $\boldsymbol{\omega}_0$ to distinguish the true frequency vector from a generic vector $\boldsymbol{\omega}$.

In [7] we prove that $\hat{\boldsymbol{\omega}}$ given by (10) is strongly consistent, that is, $\hat{\boldsymbol{\omega}} \to \boldsymbol{\omega}_0$ with probability one as N tends to infinity. After establishing consistency, the asymptotic distribution of the estimates from (10) can be derived through a Taylor series expansion approach; see e.g. [4, 6]. By definition $V'(\hat{\boldsymbol{\omega}}) = 0$ and since $\hat{\boldsymbol{\omega}}$ is consistent, an expansion of $V'(\hat{\boldsymbol{\omega}})$ around $\boldsymbol{\omega}_0$ leads to

$$\tilde{\boldsymbol{\omega}} \simeq -\mathbf{H}^{-1}V'(\boldsymbol{\omega}_0), \tag{12}$$

where $\simeq$ denotes equality in probability up to first order, and where $\mathbf{H} = \lim_{N\to\infty} V''(\boldsymbol{\omega}_0)$ and $\tilde{\boldsymbol{\omega}} = \hat{\boldsymbol{\omega}} - \boldsymbol{\omega}_0$. The derivative of (11) with respect to ω_i is

$$\begin{aligned} V_i &= \hat{\mathbf{f}}^*\mathbf{P}_i^{\perp}\mathbf{W}\mathbf{P}^{\perp}\hat{\mathbf{f}} + \hat{\mathbf{f}}^*\mathbf{P}^{\perp}\mathbf{W}\mathbf{P}_i^{\perp}\hat{\mathbf{f}} \\ &\simeq -2\,\mathrm{Re}\{\hat{\mathbf{f}}^*\mathbf{B}^{\dagger^*}\mathbf{B}_i^*\mathbf{P}^{\perp}\mathbf{W}\mathbf{P}^{\perp}\hat{\mathbf{f}}\}, \end{aligned} \tag{13}$$

since $\mathbf{P}_i^{\perp} = -\mathbf{B}^{\dagger^*}\mathbf{B}_i^*\mathbf{P}^{\perp} - \mathbf{P}^{\perp}\mathbf{B}_i\mathbf{B}^{\dagger}$. It is shown in Appendix A that, asymptotically, $\hat{\mathbf{f}} = \mathbf{M}\,\mathrm{vec}\{\hat{\mathbf{R}}\}$ for a certain transformation $\mathbf{M}$. Since the elements in $\sqrt{N}(\hat{\mathbf{R}} - \mathbf{R})$ are asymptotically Gaussian distributed, the same is true for $\sqrt{N}V'$ and $\sqrt{N}\tilde{\boldsymbol{\omega}}$. Hence, we have the following result.

Theorem 1. *The estimate $\hat{\boldsymbol{\omega}}$ from (10) is a consistent estimate of $\boldsymbol{\omega}_0$ and the normalized estimation error is asymptotically Gaussian distributed according to*

$$\sqrt{N}(\hat{\boldsymbol{\omega}} - \boldsymbol{\omega}_0) \in \mathrm{AsN}(0, \mathbf{\Gamma}) \tag{14}$$

where

$$\mathbf{\Gamma} = \mathbf{H}^{-1}\mathbf{Q}\mathbf{H}^{-1}. \tag{15}$$

The matrices $\mathbf{H}$ and $\mathbf{Q}$ are given by

$$\mathbf{H} = 2\,\mathrm{Re}\left\{\mathbf{S}\mathbf{D}^*\mathbf{P}_{\mathbf{B}}^{\perp}\mathbf{W}\mathbf{P}_{\mathbf{B}}^{\perp}\mathbf{D}\mathbf{S}\right\}, \tag{16}$$

$$\mathbf{Q} \triangleq \lim_{N\to\infty} N\,\mathrm{E}\{V'V'^T\} = 2\,\mathrm{Re}\left\{\mathbf{U}^*\bar{\mathbf{C}}\mathbf{U}^c + \mathbf{U}^*\mathbf{C}\mathbf{U}\right\}. \tag{17}$$

Here,

$$\mathbf{D} = (\bar{\mathbf{D}}^c \circ \mathbf{A}) + (\mathbf{A}^c \circ \bar{\mathbf{D}})\,, \tag{18}$$

$$\bar{\mathbf{D}} = \left[\left.\frac{\partial \mathbf{a}(\omega)}{\partial \omega}\right|_{\omega_1} \quad \cdots \quad \left.\frac{\partial \mathbf{a}(\omega)}{\partial \omega}\right|_{\omega_d}\right], \tag{19}$$

$$\mathbf{U} = \mathbf{P}_{\mathbf{B}}^{\perp}\mathbf{W}\mathbf{P}_{\mathbf{B}}^{\perp}\mathbf{D}\mathbf{S}, \tag{20}$$

where $\mathbf{a}(\omega)$ denotes a column of $\mathbf{A}(\boldsymbol{\omega})$. The two covariance matrices $\mathbf{C}$ and $\bar{\mathbf{C}}$ are defined as

$$\mathbf{C} = \lim_{N\to\infty} N\,\mathrm{E}\{\tilde{\mathbf{f}}\tilde{\mathbf{f}}^*\}, \tag{21}$$

$$\bar{\mathbf{C}} = \lim_{N\to\infty} N\,\mathrm{E}\{\tilde{\mathbf{f}}\tilde{\mathbf{f}}^T\}, \tag{22}$$

where $\tilde{\mathbf{f}} = \hat{\mathbf{f}} - \mathbf{f}$. Explicit expressions for $\mathbf{C}$ and $\bar{\mathbf{C}}$ can be found in Appendix A. All quantities are evaluated in $\boldsymbol{\omega}_0$.

Proof. The expressions for $\mathbf{H}$ and $\mathbf{Q}$ are derived in Appendix B. □

The above result is valid for any Hermitian positive definite weighting $\mathbf{W}$. The optimal choice of $\mathbf{W}$ in terms of minimizing the asymptotic covariance matrix $\mathbf{\Gamma}$ is provided by the following corollary.

Corollary 1. *Let*

$$\mathbf{W} = \mathbf{W}_{opt} \triangleq [\mathbf{P}_{\mathbf{B}}^{\perp}\tilde{\mathbf{C}}\mathbf{P}_{\mathbf{B}}^{\perp} + \mathbf{B}\mathbf{B}^*]^{-1}, \tag{23}$$

where

$$\begin{aligned} \tilde{\mathbf{C}} &= \mathbf{C} + \sigma^4\left(\mathbf{P}_{\mathbf{A}}^{\perp^T} \otimes \mathbf{P}_{\mathbf{A}}^{\perp}\right) \\ &= \left(\mathbf{R}^T \otimes \mathbf{R}\right) + \frac{\sigma^4}{m-d}\,\mathrm{vec}\{\mathbf{P}_{\mathbf{A}}\}\,\mathrm{vec}^*\{\mathbf{P}_{\mathbf{A}}\}, \end{aligned} \tag{24}$$

and where $\mathbf{P}_{\mathbf{A}} = \mathbf{I} - \mathbf{P}_{\mathbf{A}}^{\perp} = \mathbf{E}_s\mathbf{E}_s^*$. Then

$$\mathbf{\Gamma} = \mathbf{CRB}_{\boldsymbol{\omega}}, \tag{25}$$

where $\mathbf{CRB}_{\boldsymbol{\omega}}$ is the Cramér Rao lower bound on the estimation error variance of $\boldsymbol{\omega}$ corresponding to the prior knowledge that the signals are uncorrelated.

Proof. The proof can be found in [7]. □

The result of the corollary implies that the estimator (10) with the weighting (23) is asymptotically equivalent to the maximum likelihood estimator that also uses information about the signal correlation. The gain in using (10) lies in the possibility that it may be easier to minimize (11) than the likelihood function. As described in the next section it is in fact possible to solve the minimization of (11) non-iteratively if a ULA is employed.

5. IMPLEMENTATION

One may directly observe that the optimal weighting (23) depends on unknown quantities. However, it can be shown that $\mathbf{W}_{opt}$ can be replaced with a consistent estimate without changing the asymptotic properties; c.f. [4]. For general arrays, (10) can be solved by a Newton-type method. In the following we will however discuss a way to avoid the nonlinear minimization that usually is necessary. If a ULA is employed, a technique similar to the one used in MODE [2, 8] can be utilized. The idea is to find a basis for the null-space of $\mathbf{B}^*$ that depends linearly on a minimal set of parameters. For this purpose, introduce the following polynomial

$$g_0 z^d + g_1 z^{d-1} + \cdots + g_d = g_0 \prod_{k=1}^{d} \left(z - e^{i\omega_k}\right) \qquad (26)$$
$$g_0 \neq 0 .$$

From the definition of $\mathbf{B}(\omega)$ it follows that the kth column of $\mathbf{B}$ is given by

$$\mathbf{B}_k = [1 \; z_k \cdots z_k^{m-1} \vdots z_k^{-1} \; 1 \cdots z_k^{m-2} \vdots z_k^{-2} \cdots z_k^{m-3} \vdots \cdots$$
$$\cdots \vdots z_k^{-(m-1)} \cdots 1]^T \qquad (27)$$

where $z_k = e^{i\omega_k}$. The goal is to find a full rank matrix $\mathbf{G}$ of dimension $m^2 \times (m^2 - d)$ such that $\mathbf{G}^*\mathbf{B}(\omega_0) = 0$. Below we give two simple examples from which a general parameterization easily follows. For illustration purposes we permute the rows of $\mathbf{B}(\omega_0)$, and thus the columns of $\mathbf{G}^*$, such that the permuted kth column of $\mathbf{B}(\omega_0)$ reads

$$[z_k^{-m+1} \; z_k^{-m+2} \; z_k^{-m+2} \; z_k^{-m+3} \; \cdots \; z_k^{m-1}]^T. \qquad (28)$$

This permutation will highlight the generalization of the parameterization of $\mathbf{G}^*$ given in the examples that follow. In the first example of a permuted $\mathbf{G}^*$ matrix, $\tilde{\mathbf{G}}^*$, let $m = 3$ and $d = 1$ which implies that we need $m^2 - d = 8$ independent rows. One such $\tilde{\mathbf{G}}^*$ is

$$\tilde{\mathbf{G}}^* = \left[\begin{array}{c|cc|ccc|ccc} g_1 & g_0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ g_1 & 0 & g_0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & g_1 & 0 & g_0 & 0 & 0 & 0 & 0 & 0 \\ 0 & g_1 & 0 & 0 & g_0 & 0 & 0 & 0 & 0 \\ 0 & g_1 & 0 & 0 & 0 & g_0 & 0 & 0 & 0 \\ 0 & 0 & 0 & g_1 & 0 & 0 & g_0 & 0 & 0 \\ 0 & 0 & 0 & g_1 & 0 & 0 & 0 & g_0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & g_1 & 0 & g_0 \end{array}\right],$$

which is easily seen to be full rank. In the second example we let $m = 3$ and $d = 2$, implying that $m^2 - d = 7$ independent rows are needed. In this case one may take

$$\tilde{\mathbf{G}}^* = \left[\begin{array}{c|cc|ccc|ccc} 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ g_2 & g_1 & 0 & g_0 & 0 & 0 & 0 & 0 & 0 \\ g_2 & g_1 & 0 & 0 & g_0 & 0 & 0 & 0 & 0 \\ g_2 & g_1 & 0 & 0 & 0 & g_0 & 0 & 0 & 0 \\ 0 & g_2 & 0 & g_1 & 0 & 0 & g_0 & 0 & 0 \\ 0 & g_2 & 0 & g_1 & 0 & 0 & 0 & g_0 & 0 \\ 0 & 0 & 0 & g_2 & 0 & 0 & g_1 & 0 & g_0 \end{array}\right].$$

In the general case there are $d(d-1)/2$ rows with ± 1 and $m^2 - d^2/2 - d/2$ rows with g-coefficients. Altogether this becomes $m^2 - d$ rows and by construction these rows are linearly independent. Observe that the polynomial (26) should have its roots on the unit circle. For our purposes, this can be realized by imposing the conjugate symmetry constraint $g_i = g_{d-i}^c$, $i = 0, ..., d$; see [2, 7] for details.

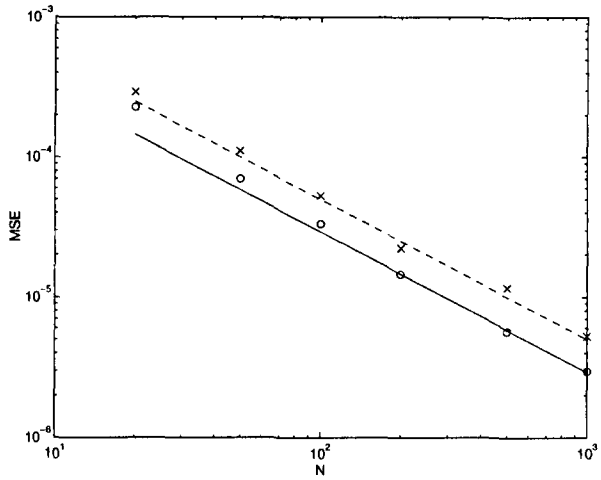

Figure 1. MSE for θ_2 versus the number of snapshots, N: 'o' – proposed method, 'x' – root-MUSIC. The solid line represents the CRB when the correlation structure of the sources is known and the dashed line is the CRB without this knowledge.

Next notice that $\mathbf{P}_{\mathbf{B}}^{\perp} = \mathbf{G}\mathbf{G}^{\dagger}$ and rewrite (11) as

$$\hat{\mathbf{f}}^* \mathbf{P}_{\mathbf{B}}^{\perp} \mathbf{W}_{opt} \mathbf{P}_{\mathbf{B}}^{\perp} \hat{\mathbf{f}} = \hat{\mathbf{f}}^* \mathbf{G}(\mathbf{G}^* \tilde{\mathbf{C}} \mathbf{G})^{-1} \mathbf{G}^* \hat{\mathbf{f}}. \qquad (29)$$

Since $\mathbf{G}^*\mathbf{f} = 0$, it is possible to show that the inverse in (29) can be replaced with a consistent estimate without altering the asymptotic properties of the estimates. We thus propose to estimate ω by minimizing

$$\hat{\mathbf{f}}^* \mathbf{G}(\hat{\mathbf{G}}^* \hat{\tilde{\mathbf{C}}} \hat{\mathbf{G}})^{-1} \mathbf{G}^* \hat{\mathbf{f}} \qquad (30)$$

over the d free real parameters in $\mathbf{G}$ (the parameters are real and imaginary parts of g_i under the conjugate symmetry constraint). Since these parameters enter linearly in $\mathbf{G}$, the problem can be solved by the solution to an over-determined set of linear equations. Once the polynomial coefficients are given, $\hat{\omega}$ is obtained by rooting the polynomial (26). In (30), $\hat{\tilde{\mathbf{C}}}$ is an estimate of $\tilde{\mathbf{C}}$ computed from sample data and $\hat{\mathbf{G}}$ is constructed from a consistent estimate of ω_0, for example, the root-MUSIC estimate.

6. SIMULATION EXAMPLE

To illustrate that the asymptotic expressions may be valid for quite modest sample sizes we provide an example. Consider the direction of arrival estimation of two waves impinging from angles $\theta_1 = 0°$ and $\theta_2 = 10°$ on a ULA with 5 elements separated by a half wavelength. The uncorrelated signal sources are modeled as white and circularly symmetric complex Gaussian distributed with a variance of 3 and 10, respectively. The additive noise is spatially and temporally white circularly symmetric complex Gaussian with variance $\sigma^2 = 1$. The mean-square-error (MSE) for different data lengths are calculated for the proposed method and for root-MUSIC [1, 9]. Each MSE is based on 200 independent trials. The MSE for θ_2 is depicted in Fig. 1. It is seen that the new method performs similar to root-MUSIC but has less variance for large samples.

7. CONCLUSIONS

In this paper the method proposed in [5] was analyzed. The asymptotic distribution was derived and the asymptotic variance was shown to coincide with the Cramér-Rao

lower bound including knowledge of uncorrelated signals. It was also shown that the estimator can be implemented in a non-iterative fashion for uniform linear arrays. This makes the method quite attractive since it provides minimum variance frequency estimates without having to resort to non-linear minimization.

REFERENCES

[1] P. Stoica and A. Nehorai, "MUSIC, Maximum Likelihood and Cramér-Rao Bound", *IEEE Trans. ASSP*, **ASSP-37**:720–741, May 1989.

[2] P. Stoica and K. Sharman, "Novel Eigenanalysis Method for Direction Estimation", *Proc. IEE*, **F**:19–26, Feb 1990.

[3] P. Stoica and T. Söderström, "Statistical Analysis of MUSIC and Subspace Rotation Estimates of Sinusoidal Frequencies", *IEEE Trans. ASSP*, **ASSP-39**:1836–1847, Aug. 1991.

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

[5] B. Göransson, M. Jansson, and B. Ottersten, "Spatial and Temporal Frequency Estimation of Uncorrelated Signals Using Subspace Fitting", In *Proceedings of 8th IEEE Signal Processing Workshop on Statistical Signal and Array Processing*, pages 94–96, Corfu, Greece, June 1996.

[6] L. Ljung, *System Identification: Theory for the User*, Prentice-Hall, Englewood Cliffs, NJ, 1987.

[7] Magnus Jansson, Bo Göransson, and Björn Ottersten, Frequency estimation of uncorrelated signals using subspace fitting, Technical Report IR-S3-REG-9616, Signals, Sensors & Systems, KTH, Stockholm, 1996.

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

[9] B. D. Rao and K. V. S. Hari, "Performance Analysis of Root-Music", *IEEE Trans. ASSP*, **ASSP-37**(12):1939–1949, Dec. 1989.

A ASYMPTOTIC RELATIONS FOR THE RESIDUAL

In this appendix we relate the residual $\tilde{\mathbf{f}} = \hat{\mathbf{f}} - \mathbf{f}$ to the sample covariance matrix and derive the asymptotic covariances of $\tilde{\mathbf{f}}$. From the definition of $\hat{\mathbf{f}}$ we get

$$\begin{aligned}\hat{\mathbf{f}} &= \mathrm{vec}\{\hat{\mathbf{E}}_s(\hat{\mathbf{\Lambda}}_s - \hat{\sigma}^2\mathbf{I})\hat{\mathbf{E}}_s^*\} \\ &= \mathrm{vec}\{\hat{\mathbf{R}} - \hat{\sigma}^2\mathbf{I} - \hat{\mathbf{E}}_n(\hat{\mathbf{\Lambda}}_n - \hat{\sigma}^2\mathbf{I})\hat{\mathbf{E}}_n^*\} \\ &\simeq \mathrm{vec}\{\hat{\mathbf{R}} - \hat{\sigma}^2\mathbf{I} - \mathbf{E}_n(\hat{\mathbf{\Lambda}}_n - \hat{\sigma}^2\mathbf{I})\mathbf{E}_n^*\} \\ &= \mathrm{vec}\{\hat{\mathbf{R}} - \hat{\sigma}^2(\mathbf{I} - \mathbf{E}_n\mathbf{E}_n^*) - \mathbf{E}_n\hat{\mathbf{\Lambda}}_n\mathbf{E}_n^*\}.\end{aligned} \tag{31}$$

Recall that the noise variance is estimated as the average of the noise eigenvalues in $\hat{\mathbf{\Lambda}}_n$ and notice that

$$\hat{\mathbf{\Lambda}}_n \simeq \mathbf{E}_n^*\hat{\mathbf{R}}\mathbf{E}_n. \tag{32}$$

We thus asymptotically have

$$\begin{aligned}\hat{\sigma}^2 &= \frac{1}{m-d}\mathrm{Tr}\{\hat{\mathbf{\Lambda}}_n\} \simeq \frac{1}{m-d}\mathrm{Tr}\{\mathbf{E}_n^*\hat{\mathbf{R}}\mathbf{E}_n\} \\ &= \frac{1}{m-d}\mathrm{vec}^*\{\mathbf{E}_n\mathbf{E}_n^*\}\,\mathrm{vec}\{\hat{\mathbf{R}}\},\end{aligned} \tag{33}$$

since $\mathrm{Tr}\{\mathbf{AB}\} = \mathrm{vec}^*\{\mathbf{A}^*\}\,\mathrm{vec}\{\mathbf{B}\}$. After some calculations we obtain

$$\begin{aligned}\hat{\mathbf{f}} &= [\mathbf{I} - \left(\mathbf{P}_\mathbf{A}^{\perp T} \otimes \mathbf{P}_\mathbf{A}^\perp\right) - \frac{1}{m-d}\mathrm{vec}\{\mathbf{P}_\mathbf{A}\}\,\mathrm{vec}^*\{\mathbf{P}_\mathbf{A}^\perp\}]\,\mathrm{vec}\{\hat{\mathbf{R}}\} \\ &\triangleq \mathbf{M}\,\mathrm{vec}\{\hat{\mathbf{R}}\},\end{aligned} \tag{34}$$

where $\mathbf{P}_\mathbf{A}^\perp = \mathbf{E}_n\mathbf{E}_n^* = \mathbf{I} - \mathbf{E}_s\mathbf{E}_s^* = \mathbf{I} - \mathbf{P}_\mathbf{A}$. Using (34), it is straightforward to derive $\mathbf{C}$ and $\bar{\mathbf{C}}$. Let us start with $\bar{\mathbf{C}}$ in (22)

$$\begin{aligned}\bar{\mathbf{C}} &= \lim_{N\to\infty} N\,\mathrm{E}\{\mathbf{M}\,\mathrm{vec}\{\tilde{\mathbf{R}}\}\,\mathrm{vec}^T\{\tilde{\mathbf{R}}\}\mathbf{M}^T \\ &= \mathbf{M}\left[\mathrm{vec}(\mathbf{R})\,\mathrm{vec}^T(\mathbf{R})\right]^{BT}\mathbf{M}^T,\end{aligned} \tag{35}$$

where the superscript 'BT' denotes block-transpose and means that each $m \times m$ block is transposed. Similarly we get for $\mathbf{C}$

$$\begin{aligned}\mathbf{C} &= \lim_{N\to\infty} N\,\mathrm{E}\{\mathbf{M}\,\mathrm{vec}\{\tilde{\mathbf{R}}\}\,\mathrm{vec}^*\{\tilde{\mathbf{R}}\}\mathbf{M}^* \\ &= \mathbf{M}\left(\mathbf{R}^T \otimes \mathbf{R}\right)\mathbf{M}^*.\end{aligned} \tag{36}$$

Using the expression for $\mathbf{M}$ given in (34), some tedious calculations lead to the following expression for $\mathbf{C}$:

$$\begin{aligned}\mathbf{C} &= \left(\mathbf{R}^T \otimes \mathbf{R}\right) - \sigma^4\left(\mathbf{P}_\mathbf{A}^{\perp T} \otimes \mathbf{P}_\mathbf{A}^\perp\right) \\ &\quad + \frac{\sigma^4}{m-d}\mathrm{vec}\{\mathbf{P}_\mathbf{A}\}\,\mathrm{vec}^*\{\mathbf{P}_\mathbf{A}\}.\end{aligned} \tag{37}$$

The matrix is (as expected) singular and the null space is spanned by the columns of $(\mathbf{E}_n^c \otimes \mathbf{E}_n)$. The dimension of the null space is $(m-d)^2$ and consequently the dimension of the range space of $\mathbf{C}$ is $m^2 - (m-d)^2 = 2md - d^2$. Notice that this equals the number of real parameters in the Cholesky-factorization of $\mathbf{ASA}^*$.

B DERIVATION OF Q AND H

Consider first the $\mathbf{Q}$ matrix. Define

$$\mathbf{u}_i = \mathbf{P}^\perp\mathbf{W}\mathbf{P}^\perp\mathbf{B}_i\mathbf{s} \tag{38}$$

which is nothing but the ith column of $\mathbf{U}$ in (20). The ijth element of $\mathbf{Q}$ is then, in view of (13), given by

$$\begin{aligned}\mathbf{Q}_{ij} &= \lim_{N\to\infty} N\,\mathrm{E}\{V_iV_j\} = \lim_{N\to\infty} N\,\mathrm{E}\{2\,\mathrm{Re}\{\mathbf{u}_i^*\tilde{\mathbf{f}}\}2\,\mathrm{Re}\{\mathbf{u}_j^*\tilde{\mathbf{f}}\} \\ &= \lim_{N\to\infty} N\,\mathrm{E}\{2\,\mathrm{Re}\{\mathbf{u}_i^*\tilde{\mathbf{f}}\tilde{\mathbf{f}}^*\mathbf{u}_j + \mathbf{u}_i^*\tilde{\mathbf{f}}\tilde{\mathbf{f}}^T\mathbf{u}_j^c\} \\ &= 2\,\mathrm{Re}\{\mathbf{u}_i^*\mathbf{C}\mathbf{u}_j + \mathbf{u}_i^*\bar{\mathbf{C}}\mathbf{u}_j^c\},\end{aligned} \tag{39}$$

which is the ijth element of the expression given in (17).

From (13) we get

$$\mathbf{H}_{ij} = \lim_{N\to\infty} V_{ij} = 2\,\mathrm{Re}\{\mathbf{s}^*\mathbf{B}_i^*\mathbf{P}^\perp\mathbf{W}\mathbf{P}^\perp\mathbf{B}_j\mathbf{s}\} \tag{40}$$

which is readily seen to equal the ijth element of (16).