

Tunable Line Spectral Estimators Based on State-Covariance Subspace Analysis

Ali Nasiri Amini, *Member, IEEE*, and Tryphon T. Georgiou, *Fellow, IEEE*

Abstract—Subspace methods for spectral analysis can be adapted to the case where state covariance of a linear filter replaces the traditional Toeplitz matrix formed out of a partial autocorrelation sequence of a time series. This observation forms the basis of a new framework for spectral analysis. The goal of this paper is to quantify potential advantages in working with state-covariance data instead of the autocorrelation sequence. To this end, we identify tradeoffs between resolution and robustness in spectral estimates and how these are affected by the filter dynamics. The approach leads to a novel *tunable* high-resolution frequency estimator.

Index Terms—Harmonic decomposition, spectral estimation, state-covariance, subspace methods.

I. INTRODUCTION

THERE are numerous engineering applications that require estimating the power spectrum of a stochastic process from a finite observation record. For instance, in communications, radar, sonar, and geophysical seismology, spectral analysis of a recorded signal/echo is essential for data compression, speaker recognition, target identification, or the identification of underlying geophysical morphology. To this end, spectral estimation and analysis techniques occupied a central role in signal processing research over the past three decades. While traditional Fourier and periodogram-based techniques remained the workhorse in most areas, their limited resolution capabilities stimulated research in nonlinear methods. Indeed, over the past three decades, this research has led to a wide spectrum of “high-resolution” methods and, among them, subspace methods which offer high resolution with low-computational complexity. This paper is about a new formalism for spectral analysis and corresponding analogs of subspace methods.

Standard nonlinear spectral estimation methods [1], [2] typically rely on an estimated partial covariance sequence. The power spectrum is then identified as being consistent with such statistics while satisfying side conditions (e.g., maximum entropy, composed of sinusoids, etc.). In this spirit, subspace methods are based on a suitable eigendecomposition of the estimated covariance matrix of the process, so as to identify the dominant eigendirections which characterize embedded

harmonics. The new formalism begins with estimating the covariance of state vector of a linear filter. The dynamics of such a filter, may be dictated either by sensing apparatus, or the geometry of antenna array, or simply by “virtual dynamics,” specified in an algorithmic stage. The filter can be also conceived as a filter-bank made of several subcomponents.

The state-covariance can be treated and analyzed very much like the usual covariance matrix in standard subspace methods [3], [4]. The essence of the new framework is that it allows a degree of freedom in tuning the statistics (state covariance), via suitable choice of the filter, to affect the resolution when estimating the power spectrum over different frequency bands. The underlying idea is akin to beamspace methods for array processing [5, p. 1062], with an important difference—the state covariance has a known algebraic structure which can be exploited.

In Section II, we briefly survey classical subspace spectral estimation. In Section III we introduce the new setting for spectral analysis which is based on state-covariances and we explain how the concepts of the *signal subspace* and the *noise subspace* carry over in this new framework. We show how the parameters of the IS filter affect the frequency estimator and its resolution over different frequency bands. The degree of freedom afforded by the choice of such parameters allows to trade resolution in different part of the spectrum.

A rigorous statistical analysis has been carried out in Section IV to quantify relevant tradeoffs. More specifically, we study the variability of frequency estimates of a time series which consist of sinusoids in white noise. We derive an explicit expression for the variance of canonical angles between estimated and true signal subspaces when the variance of the estimated state-covariance is taken into account. Later in Section V, we discuss and compare the performance of this approach with more traditional subspace analysis. Finally, we discuss a “suboptimal” IS filter design and demonstrate the performance of relevant frequency estimator via simulation studies. Proofs of the propositions and the lemmas are given in Section VII.

II. SUBSPACE METHODS

Consider the following classical problem: estimate the power spectral density function of a stationary stochastic process $y(k)$ from a finite observation record $\{y(1), y(2), \dots, y(N)\}$, where it is known that $y(k)$ consists of a number of complex sinusoids (i.e., cisoids) embedded in white noise. More precisely, let

$$y(k) = \sum_{i=1}^m s_i(k) + v(k) \quad (1)$$

Manuscript received May 21, 2004; revised June 15, 2005. This work was supported in part by NSF and AFSOR. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Steven L. Grant.

The authors are with Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN 55455 USA (email: nasiri@ece.umn.edu; tryphon@ece.umn.edu).

Digital Object Identifier 10.1109/TSP.2006.874397

where

$$s_i(k) := \rho_i e^{j(\omega_i k + \phi_i)}$$

$$\Omega := \{\omega_1, \dots, \omega_m\}$$

and the following standard assumptions hold: ϕ_i s are independent random variables uniformly distributed in $[-\pi, \pi]$; $v(k)$ is zero mean white Gaussian noise with $E\{v(k)v(k)^*\} = \sigma^2$; and the number of complex sinusoids, i.e., m , is known, where “*” denotes complex conjugate transpose. Note that if Ω is known, determining the remaining of the signal parameters (i.e., phase and amplitude for each sinusoid) is a simple linear regression problem [2, ch. 4]. Therefore, the challenging part of the problem is estimating Ω and we focus on this.

A popular class of techniques for estimating harmonics relies on exploring the covariance structure of the stochastic process. To this end, let $\{r_0, r_1, \dots, r_{n-1}\}$ be the partial covariance sequence of $y(k)$ (i.e., $r_\ell = E\{y(k)y^*(k-\ell)\}$), and

$$T = \begin{bmatrix} r_0 & r_1^* & \dots & r_{n-1}^* \\ r_1 & r_0 & \dots & r_{n-2}^* \\ \vdots & \vdots & \ddots & \vdots \\ r_{n-1} & r_{n-2} & \dots & r_0 \end{bmatrix} \quad (2)$$

the associated Toeplitz covariance matrix. Under the earlier hypothesis on the composition of $y(k)$

$$T = \mathcal{G}(\Omega)R\mathcal{G}^*(\Omega) + \sigma^2 I \quad (3)$$

where I is identity matrix of appropriate dimension $R := \text{diag}(\rho_1^2, \dots, \rho_m^2)$ and

$$\mathcal{G}(\Omega) := \begin{bmatrix} 1 & 1 & \dots & 1 \\ e^{-j\omega_1} & e^{-j\omega_2} & \dots & e^{-j\omega_m} \\ \vdots & \vdots & \ddots & \vdots \\ e^{-j(n-1)\omega_1} & e^{-j(n-1)\omega_2} & \dots & e^{-j(n-1)\omega_m} \end{bmatrix}.$$

At the same time a singular value decomposition (SVD) of T gives that

$$T = U_s \Lambda_s U_s^* + U_n \Lambda_n U_n^* \quad (4)$$

where $[U_s \ U_n]$ is a unitary matrix and the above summands correspond to splitting the contribution of the largest m singular values from the rest. It is easy to show [2, pp. 155–157], that

$$\text{Range}(\mathcal{G}(\Omega)) = \text{Range}(U_s). \quad (5)$$

This range, which is a m -dimensional subspace, is referred to as the *signal subspace* while its complement $\text{Range}(U_n)$ is referred to as the *noise subspace*. Note that the signal subspace is fully specified by Ω . In fact,

$$\Omega \xrightarrow{\mathbf{g}} \text{Range}(\mathcal{G}(\Omega)) \quad (6)$$

is bijective (see [2, pp. 155–157]). In other words, the set of frequencies, Ω , can be determined uniquely from the signal subspace. Thus, estimating Ω amounts to estimating corresponding signal subspace.

The subspace methodology for estimating Ω can be viewed as a two-step approach: First a basis for the signal subspaces is

obtained via a SVD on the signal statistics T in (4). Then the frequencies are determined by $\Omega = g^{-1}(\text{Range}(U_s))$.

In practice, we work with the sample covariance matrix \hat{T} , and, hence, signal and noise subspaces can only be approximated. Therefore, a set of frequencies that corresponds to the estimated subspaces can not be found in general. The variety in subspace algorithms (e.g., MUSIC, ESPRIT) stems from different approximation schemes that they use to fit a set of frequencies to the estimated subspaces. In the sequel, we show a natural extension of the concept of signal statistics T and accordingly, the generalization of subspace methods for spectral estimation.

III. STATE-COVARIANCE SUBSPACE METHODS

The basic idea in retrieving information from the structure of T extends to the case where T is replaced by the state-covariance of a known linear filter which is driven by the stationary process $y(k)$ (see [3]). Here we briefly sketch this idea.

Consider the linear filter:

$$x(k) = Ax(k-1) + By(k), \quad k \in Z \quad (7)$$

where $y(k)$ is a scalar input and the complete state x is available at the output. In particular, assume $A \in \mathbb{C}^{n \times n}$ and $B \in \mathbb{C}^{n \times 1}$ is a controllable pair, A is a stable matrix (i.e., its eigenvalues are inside the unit disk), and the following holds:

$$AA^* + BB^* = I. \quad (8)$$

Equation (8) is a normalization condition that causes no loss in generality but simplifies a number of derivations. We refer to this system as an *input-to-state filter (IS filter)*. Its frequency response is

$$G(\omega) = (I - e^{-j\omega}A)^{-1}B$$

and the eigenvalues of A , denoted by $\{a_1, \dots, a_n\}$, are referred to as the *modes of the filter*.

For an input $y(k)$ as in (1), the value of the state vector at steady state is given by

$$x(k) = \sum_{i=1}^m G(\omega_i)s_i(k) + \xi(k) \quad (9)$$

where $\xi(k)$ denotes the response to the additive noise $v(k)$. If

$$\mathcal{G}(\Omega) := [G(\omega_1) \ G(\omega_2) \ \dots \ G(\omega_m)]$$

$$s(k) := [s_1(k) \ s_2(k) \ \dots \ s_m(k)]^T \quad (10)$$

(9) becomes

$$x(k) = \mathcal{G}(\Omega)s(k) + \xi(k). \quad (11)$$

Since $y(k)$ is wide-sense stationary so, is $x(k)$. The *state-covariance* $P := E\{x(k)x^*(k)\}$ can be written as

$$P = \mathcal{G}(\Omega)R\mathcal{G}(\Omega)^* + \sigma^2 I \quad (12)$$

where $R := \text{diag}(\rho_1^2, \dots, \rho_m^2)$. Equation (12) follows from the fact that $\xi(k)$ and $s(k)$ are independent random vectors,

$\xi(k)$ has zero mean, and the normalization in (8) implies that $E\{\xi(k)\xi(k)^*\} = \sigma^2 I$. On the other hand a SVD of P gives

$$P = U_s \Lambda_s U_s^* + U_n \Lambda_n U_n^*. \quad (13)$$

The two summands in (13) correspond to splitting the contribution of the largest m singular values from the rest. In the sequel, the singular values of P are denoted by $\lambda_1^2 \geq \lambda_2^2 \geq \dots \geq \lambda_n^2$ and their corresponding singular vectors by $[u_1 \dots u_n]$. It is easy to verify that

$$\lambda_{m+1} = \dots = \lambda_n = \sigma.$$

In complete analogy with (3) and (4), (12) and (13) imply

$$\text{Range}(\mathcal{G}(\Omega)) = \text{Range}(U_s).$$

This range is again the *signal subspace*. It is proven in [4] that

$$\Omega \xrightarrow{\mathcal{G}} \text{Range}(U_s)$$

is injective. Hence, estimating the signal subspace allows for estimating Ω as well.

In practice, P should be estimated from the sample data. To this end, the sample state process $\hat{x}(k)$, i.e.

$$\{\hat{x}(0), \hat{x}(1), \dots, \hat{x}(N-1)\}$$

is obtained using (7) and zero initial condition. This initial condition assumption has a transient effect on $\hat{x}(k)$. However, after a certain time constant, denoted by N_s , this effect becomes negligible. Therefore, state-covariance matrix can be estimated as

$$\hat{P} = \frac{1}{N_x} \sum_{k=N_s}^{N-1} \hat{x}(k)\hat{x}(k)^* \quad (14)$$

where $N_x = N - N_s$. It can be shown that (14) is a consistent estimator of the state-covariance (for details, see [6]). As in traditional subspace methods, working with a sample state-covariance requires approximation techniques to fit a set of frequencies to the estimated signal subspace. Analogs of MUSIC and ESPRIT algorithms have been developed in [4]. For instance, a scheme analogous to the ESPRIT algorithm is outlined in the following proposition.

Proposition 1: Suppose μ is a $1 \times m$ row vector and Φ is a $m \times m$ matrix such that

$$U_s = [B \quad AU_s] \begin{bmatrix} \mu \\ \Phi \end{bmatrix} \quad (15)$$

then μ and Φ are unique and eigenvalues of Φ are precisely $e^{j\omega_i}$ for $i = 1, \dots, m$.

Proof: The proof is given in [4, p. 36].

The frequencies can be retrieved as the angles of the eigenvalues of Φ . In practice, since only an estimate of the signal subspace (\hat{U}_s) is available, an estimate of the matrix Φ can be obtained by a least squares solution of

$$[B \quad A\hat{U}_s] \begin{bmatrix} \hat{\mu} \\ \hat{\Phi} \end{bmatrix} \stackrel{\text{LS}}{=} \hat{U}_s. \quad (16)$$

It turns out that IS filters with identical modes lead to identical frequency estimators. In fact, as stated below, the state space matrices of such filters relate through a unitary transformation.

Proposition 2: Suppose (A, B) and (\tilde{A}, \tilde{B}) define IS filters with A and \tilde{A} having identical eigenvalues. Then there exists unitary matrix V such that

$$(A, B) = (V^* \tilde{A} V, V^* \tilde{B}). \quad (17)$$

Proof: See Section VII.

It follows that $G(\omega) = V \tilde{G}(\omega)$ and $U_s = V \tilde{U}_s$ with $\tilde{G}(\omega)$ and \tilde{U}_s denote the corresponding quantities for (\tilde{A}, \tilde{B}) . By substituting these transformed objects in (15), V cancels out. Therefore, the frequency estimator only depends on the modes of the filter.

To estimate spectral lines one needs to consider the following:

$$\text{I) } \{y(k)\} \xrightarrow{\text{IS filter}} \{x(k)\} \longrightarrow \text{Range}(U_s) \quad (18)$$

$$\text{II) } \text{Range}(U_s) \xrightarrow{\mathcal{G}^{-1}} \Omega \quad (19)$$

where in the first, one needs to estimate the signal subspace and in the second, to fit spectral lines to it. The classical framework of Toeplitz covariance matrices is a special case and corresponds to the choice

$$A = \begin{bmatrix} 0 & 0 & \dots & 0 \\ 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \dots & 1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \quad (20)$$

In this case, all of the filter modes are at the origin and the state covariance is precisely T given in (2), while the choice of $N_s = n - 1$ in (14) gives the classical estimator of the Toeplitz covariance matrix. The main point of this paper is to show that a well designed IS filter can improve the performance of the frequency estimator.

Remark 1: To get some intuition on how the IS filter parameters affect the frequency estimation we consider the case $m = 1$, i.e., when there is only one cisoid $s_1(k) = \rho_1 e^{j(\omega_1 k + \phi_1)}$. The state process vector is given by

$$x(k) = G(\omega_1) s_1(k) + \xi(k).$$

The summands correspond to the signal and the noise, respectively. The signal and noise power can be defined, respectively, as

$$E\{G(\omega_1) s_1(k) s_1(k)^* G(\omega_1)^*\} = \rho_1^2 G(\omega_1) G(\omega_1)^* \\ E\{\xi(k)\xi(k)^*\} = \sigma^2 I.$$

Intuitively large signal-to-noise ratio (SNR) must lead to the small variance of the error in estimating ω_1 . Note that the noise power is not affected by the choice of IS filter since the normalization in (8) keeps the noise contribution constant. Thus, the estimation error may be reduced by increasing $\|G(\omega_1) G(\omega_1)^*\|$ with a suitable selection of the IS filter modes.

An expression for $\|G(\omega)G(\omega)^*\|$, derived in the Section VII, is

$$\|G(\omega)\|^2 = \sum_{i=1}^n \frac{1 - r_i^2}{1 + r_i^2 - 2r_i \cos(\omega - \theta_i)} \quad (21)$$

where $\{a_i = r_i e^{j\theta_i}\}_{i=1}^n$ are the modes of the IS filter. It is evident that if $r_i > 0$ then the corresponding summand in the above expression has a bandpass shape around θ_i and the Q-factor is determined by r_i . Furthermore, it is easy to verify that $\int_{-\pi}^{\pi} \|G(\omega)\|^2 d\omega = n$. This integral bound highlights the main tradeoff between different frequency bands, where increasing the gain $\|G(\omega)\|$ over one frequency interval necessarily reduces $\|G(\omega)\|$ elsewhere. Thus it is expected that placing the IS filter modes close to $e^{j\omega_1}$ improves the performance of the frequency estimator around ω_1 at the expense of the performance elsewhere. There is a price to be paid by placing modes of the filter close to unit circle. \square

Remark 2: IS filters with modes close to the unit circle has a large time constant which adversely affects the variance of \hat{P} . To mediate the transient effects we select N_s in (14) such that $\|A^{N_s}\| < \epsilon$, where ϵ is a sufficiently small constant. We refer to ϵ as the *transient error constant*. Large N_s reduces N_x in (14), which in turn leads to large variance in the estimated state-covariance. Next we present a statistical performance analysis that guides the design methodology to balance these tradeoffs. \square

IV. STATISTICAL PERFORMANCE ANALYSIS

We analyze the effect of the IS filter on the frequency estimator performance. In particular, first we derive an expression for the variance of the error in \hat{U}_s , and then use it to evaluate the error in the estimated frequencies.

A. Statistical Analysis of the Estimated Subspace

To evaluate the statistical properties of the estimated subspaces, a notion of distance between subspaces is needed. A natural notion of the distance between subspaces is defined via the so-called *canonical angles*. Consider the ‘‘angle operator’’ between U_s and \hat{U}_s

$$\mathcal{S} := U_n^* \hat{U}_s$$

where $[U_s, U_n]$ is a unitary matrix. The singular values of \mathcal{S} are the sines of the canonical angles between subspaces spanned by U_s and \hat{U}_s (see [7, p. 43] for details). We use the second-order statistics of \mathcal{S} to evaluate the error in \hat{U}_s . To compute the statistics of \mathcal{S} , we introduce some notation and derive an approximation of \mathcal{S} which is linear in \hat{P} . This approximation follows analogous steps as in [8] which deals with Toeplitz covariance matrices.

The second-order statistics of \mathcal{S} , i.e., ESS^* , denoted by $C(\Theta)$ where $\Theta := (\omega_1, \dots, \omega_m, (\rho_1^2/\sigma^2), \dots, (\rho_m^2/\sigma^2))$. We should point out that $C(\Theta)$ also depends on the IS filter and the length of the sample state process $N - N_s$. The autocorrelation function of $x(k)$ can be expressed as

$$P_\ell = \text{E}\{x(k)x(k - \ell)^*\} = \text{GRD}^\ell \mathcal{G}^* + Q_\ell \quad (22)$$

where the two summands correspond to the signal and the noise contributions and

$$D := \text{diag}(e^{j\omega_1}, \dots, e^{j\omega_m})$$

$$Q_\ell := \text{E}\{\xi(k)\xi^*(k - \ell)\} = \begin{cases} \sigma^2 A^\ell & \text{if } \ell \geq 0 \\ \sigma^2 (A^*)^\ell & \text{if } \ell < 0. \end{cases}$$

The following lemma gives a linear approximation of \mathcal{S} in terms of \hat{P} .

Lemma 1: For the small perturbation, i.e., $(\|P - \hat{P}\|/\|P\|) \ll 1$, the following holds

$$\mathcal{S} \simeq U_n^* \hat{P} U_s \Lambda_\sigma^{-1},$$

where $\Lambda_\sigma := \Lambda_s - \sigma^2 I$.

Proof: See Section VII.

This lemma is the key to deriving $C(\Theta)$, which is given in the following proposition.

Proposition 3: For $N_x \gg 1$

$$C(\Theta) \simeq \Lambda_\sigma^{-1} \left(\sum_{\ell=-N_x}^{N_x} \frac{(N_x - |\ell|)}{N_x^2} \text{tr}(U_n^* Q_\ell U_n) (U_s^* P_\ell^* U_s) \right) \Lambda_\sigma^{-1}. \quad (23)$$

Proof: See Section VII.

The large sample assumption guarantees the approximation in Lemma 1. Extensive Monte Carlo simulations in [6] and numerical examples in the next section show that (23) generates accurate values even for moderately sized samples. An insightful comparison of (23) with an analogous expression in [9] for antenna array processing is given in the Appendix.

B. A Performance Measure for IS Filter Design

The performance of the frequency estimator depends on the actual frequencies, their amplitudes and the noise power (i.e., Θ) which are not known *a priori*. However, in many applications there is *a priori* knowledge of the range of the unknown frequencies, or there is a particular interest in a specific frequency band. In such cases the IS filter can be tuned to improve the performance in selected frequency bands. We define a convenient performance measure to be used in the IS filter tuning as

$$f(\omega) := \frac{\gamma(\omega)}{C(\Theta)^{\frac{1}{2}}} \quad (24)$$

where

$$\gamma(\omega) := \left\| \frac{d \frac{G(\omega)}{\|G(\omega)\|}}{d\omega} \right\| \quad (25)$$

and $\Theta = (\omega, \text{SNR})$. This is defined to capture the performance of the frequency estimator in retrieving a single cisoid at the particular frequency ω when the signal-to-noise ratio is equal to SNR. Below we explain the intuition behind this definition.

In retrieving a single cisoid at frequency ω , it is natural to quantify the performance by $1/\sqrt{\text{E}\{(\omega - \hat{\omega})^2\}}$. The variance of the error in the signal subspace, which is given in (23), is a scalar quantity when $m = 1$. Therefore, it is used in the denominator of (24) to represent the error in the estimated

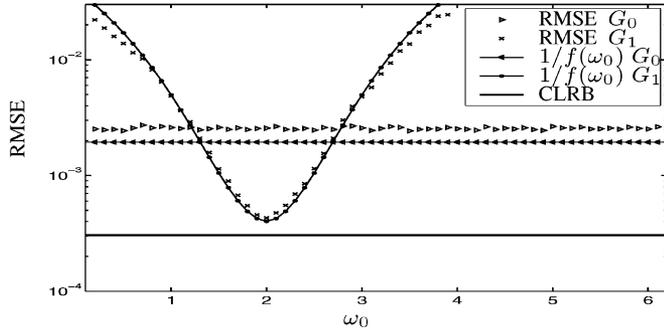


Fig. 1. Comparing performance of the frequency estimators based on G_0 and G_1 .

signal subspace. The numerator $\gamma(\omega)$, defined in (25), is the rate of changes in the signal subspace due to small changes in the ω . The factor $\gamma(\omega)$ is the sensitivity of the signal subspace to the frequency changes at the particular frequency ω . Thus, it is reasonable to scale the error in the estimated signal subspace by a factor $1/\gamma(\omega)$ to evaluate the error in the frequency estimates. As we see in Section V, $f(\omega)$ well approximates $1/\sqrt{\mathbb{E}\{(\hat{\omega} - \omega)^2\}}$. Thus, $f(\omega)$ can be used as a convenient performance measure in the tuning of the IS filter.¹

The derivative in (25) which is a function of the IS filters has been calculated in Section VII. The performance measure depends on the sample data characteristics (i.e., $\text{SNR} = \rho^2/\sigma^2$ and N) and the IS filter modes. In Section V we show how the IS filter modes can be designed to increase $f(\omega)$.

V. NUMERICAL ANALYSIS

It has been discussed in Remark 1 that placing the modes of the IS filter close to $e^{j\theta}$, for a particular θ , increases $G(\omega)$ in the frequency band around θ . Intuitively this should enhance the performance of the frequency estimator in this frequency band. This is highlighted in the following numerical examples. Throughout this section we use the extension of the well-known ESPRIT algorithm as stated in (16) to retrieve frequencies from the estimated signal subspace. It should be noted that using other techniques such as MUSIC produces similar results (see [6]).

Example 1: Consider the case of retrieving a single complex sinusoid with the frequency ω_0 embedded in noise when $N = 400$, $\text{SNR} = 0$ dB, and $\epsilon = 0.001$. We study three different frequency estimators built upon three IS filters of order $n = 10$. The first IS filter G_0 has all of its modes at the origin, thus, it corresponds to the classical ESPRIT. The second IS filter G_1 has one mode at $0.7e^{2j}$ with multiplicity 10. The third one G_2 has two modes at $0.7e^{1.5j}$ and $0.7e^{2.5j}$ each with multiplicity 5. Note that N_s is obtained as 10, 83, and 55 to satisfy $\epsilon = 0.001$, respectively G_0 , G_1 , and G_2 . For explanation on computing (A, B) corresponding to a set of prescribed modes see the Section VII-F.

¹The exact error in the frequency estimate depends on the specific technique (e.g., MUSIC, ESPRIT) which is employed to identify frequency from the estimated signal subspace. However, these techniques have similar statistical performances with only slight differences [2, p. 143] Therefore, instead of analyzing each technique separately, we resort to this rather general perturbation analysis.

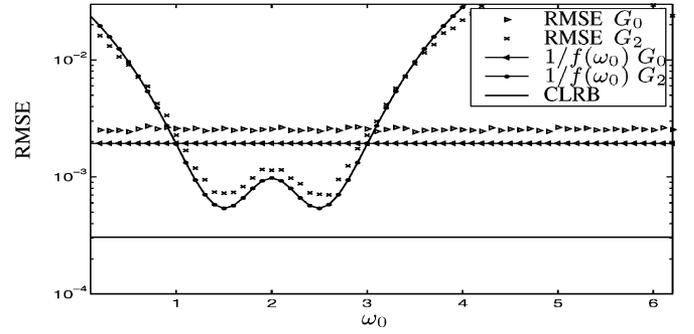


Fig. 2. Comparing performance of the frequency estimators based on G_0 and G_2 .

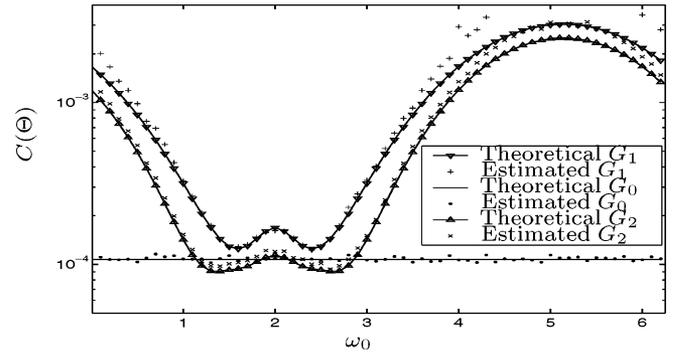


Fig. 3. Comparing the theoretical expression in (23) with Monte Carlo simulation results; $\Theta = (\omega_0, 1)$.

For a number of $\omega_0 \in [0, 2\pi]$ we evaluated the performance of these three frequency estimators by numerical Monte Carlo study. For each ω_0 we simulated 1000 realizations. The frequency estimates based on applying G_0 , G_1 , and G_2 were computed and the square root of the mean squared error (RMSE) of these estimates over the realization set have been evaluated. In Figs. 1 and 2 we have illustrated these results and compared them with the corresponding $1/f(\omega_0)$ and the Cramer-Rao lower bound (CRLB). The results confirm that $1/f(\omega)$ is a reasonable performance criterion, since its behavior agrees with the RMSE. Also, as expected, G_1 and G_2 outperform G_0 (i.e., classical ESPRIT) in the frequency bands around their modes. Note that G_1 almost achieves CRLB if the sought frequency resides close to $\omega_0 = 2.0$. We also estimated $C(\omega_0, \text{SNR})$ for the three IS filters in these simulations to verify the theoretical result in (23). The theoretical result and simulation result are shown and compared in Fig. 3. \square

This example suggests that a well designed IS filter can be tailored to specific applications to improve the performance within the selected frequency bands. Here we focus on designing the IS filter to optimize the performance in a specified frequency band, e.g., $\omega_l \leq \omega \leq \omega_u$. This is an abstraction of practical situations when some *a priori* knowledge is available about the sought frequencies. Such *a priori* information may be available from a fast but not accurate preprocessing or from a prior estimate in a tracking problem.²

²See [11] where IS filters have been applied for remote estimation of tissue temperature.

The space of design parameters is all of the combinations of $\{a_i\}_{i=1}^n$ inside the unit circle. Searching the whole space of design parameters for an optimal IS filter would be an immense task. Therefore, we resort to a suboptimal design by limiting our search to IS filters with single mode at $a = re^{j\theta}$ with multiplicity n . We choose the minimum of $f(\omega)$ in the frequency band of interest as the design objective (other possibilities are $\int_{\omega_l}^{\omega_u} f(\omega)d\omega$ and $\int_{\omega_l}^{\omega_u} f^2(\omega)d\omega$). More precisely, our choice for a is

$$a_o = \arg \max_{|a|<1} \min_{\omega_l \leq \omega \leq \omega_u} f(\omega). \quad (26)$$

Note that for IS filter with a single mode at $re^{j\theta}$ the performance measure $f(\omega)$ decreases symmetrically from its maximum at $\omega = \theta$. This implies that the angle of a_o (i.e., θ_o) should be chosen as $\theta_o = (\omega_l + \omega_u)/2$. Consequently the minimum of $f(\omega)$ in the frequency interval of interest is $f(\omega_l) = f(\omega_u)$. Therefore, the design of the suboptimal IS filter reduces to a search for the optimal r and can be summarized in these steps:

- 1) $\theta_o = (\omega_l + \omega_u)/2$;
- 2) $r_o = \arg \max_{0 \leq r < 1} f(\omega_l)$;
- 3) Follow Section VII-E to derive (A, B) with modes at $a_1 = a_2 = \dots = a_n = r_o e^{j\theta_o}$.

It should be noted that $f(\omega)$ depends on the SNR which is unknown in practice. Furthermore, when we face multiple cisoids with different amplitudes, each of them has its own SNR. In such cases we should rely on a rough estimate of the SNR or the average SNR. Fortunately r_o is practically constant for a wide range of the SNR as we see in the following example. Thus, only a rough assessment of the SNR suffices to determine the appropriate IS filter. The performance of the frequency estimators built upon the suboptimal IS filters are evaluated and compared with the classical ESPRIT algorithm in the next examples.

Example 2: We consider the case of retrieving three complex sinusoids from a sample data of length $N = 400$ where $\sigma^2 = 1$, $\omega_1 = \omega_0 - 5(2\pi/N)$, $\omega_2 = \omega_0$, $\omega_3 = \omega_0 + 5(2\pi/N)$, $\rho_1 = 8$, $\rho_2 = 4$, $\rho_3 = 2$, $n = 20$, and $\epsilon = 0.001$. Note that the cisoids are well separated from each other in the frequency domain compared to the resolution limit of the Fourier based techniques $2\pi/N$. The frequency band of interest is chosen as $[1.75, 2.25]$. The proposed algorithm is used to determine the suboptimal IS filter corresponding to this frequency band. In order to compute r_o we plotted in Fig. 4 $f(1.75) = f(2.25)$ versus r for three SNR levels corresponding to the three cisoids, i.e., $\text{SNR}_i = \rho_i^2/\sigma^2$. Note that all of these curves attains their maximum practically at $r = 0.58$. The suboptimal IS filter has one mode at $0.58e^{2j}$ with multiplicity $n = 20$. Note that satisfying $\epsilon = 0.001$ requires $N_s = 97$.

For a number of $\omega_0 \in [1, 3]$ we evaluated the performance of this frequency estimator by numerical Monte Carlo study. For each ω_0 we simulated 1000 realizations. The frequency estimates based on applying the classical ESPRIT and the suboptimal IS filter were computed. The RMSE of these estimates over the realization set have been evaluated and illustrated in Fig. 5. Since these cisoids are separated with $d > (2\pi/N)$ the CRLBs for them are well approximated by $\sqrt{6/N^3 \text{SNR}_i}$ (see [2]). Note that when the cisoids are within the frequency band of

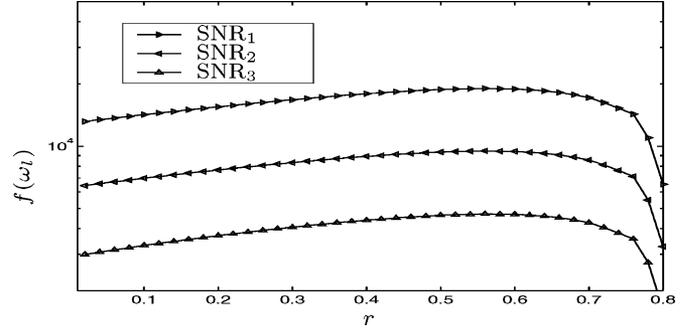


Fig. 4. $f(\omega_l)$ versus r ; $\omega_l = 1.75$, $\text{SNR}_1 = 64 = 18$ dB, $\text{SNR}_2 = 16 = 12$ dB, $\text{SNR}_3 = 4 = 6$ dB.

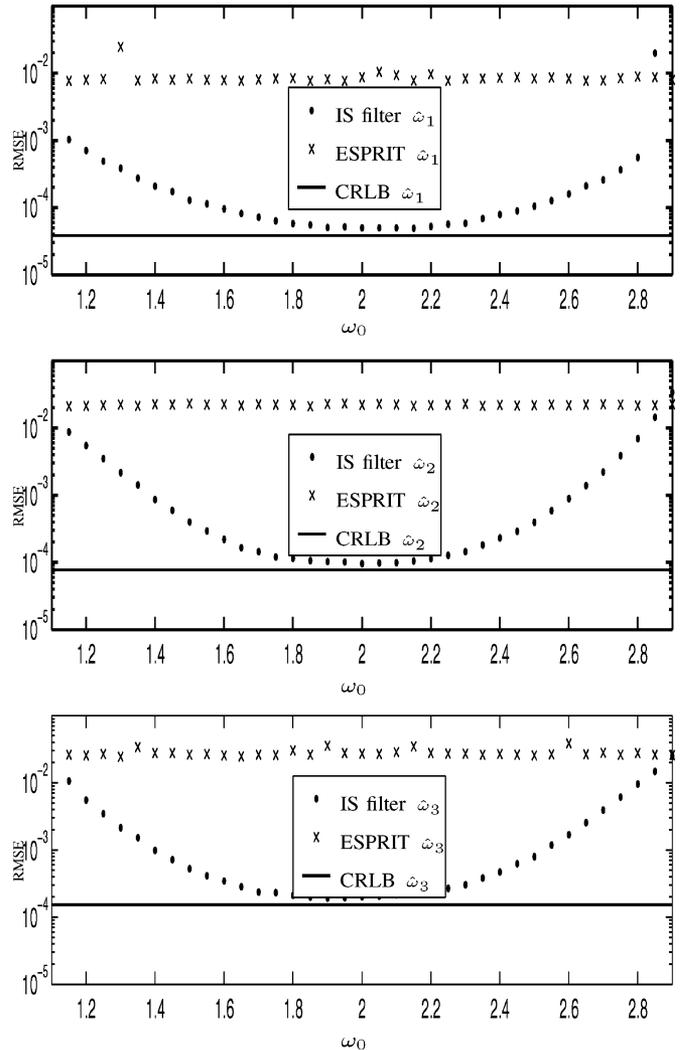


Fig. 5. Comparison of the frequency estimator based on the suboptimal IS filter and the classical ESPRIT algorithm.

interest, i.e., $[1.75, 2.25]$, the performance of the designed frequency estimator is very close to the CRLB. \square

Example 3: Subspace methods are well known for their ability to resolve closely spaced sinusoids. In this example we study the problem of resolving two closely spaced cisoids from a sample data of length $N = 400$ where $\sigma^2 = 1$, $\omega_1 = \omega_0 - (\pi/N)$, $\omega_2 = \omega_0 + \pi/N$, $\rho_1 = \rho_2 = 5$, $n = 20$, and $\epsilon = 0.001$. Similar to the previous example the frequency

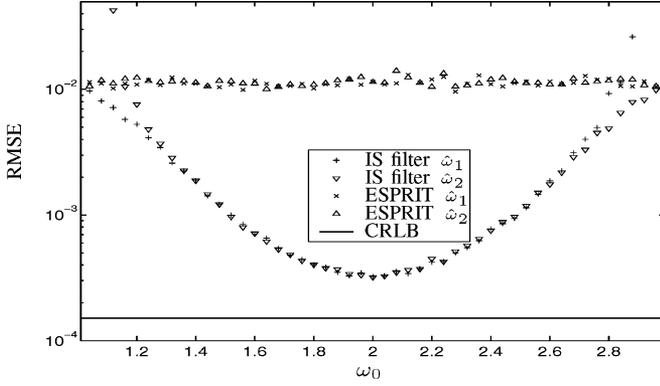


Fig. 6. Comparison of the frequency estimator based on the suboptimal IS filter and the classical ESPRIT algorithm.

band of interest is chosen as $[1.75, 2.25]$. Since the SNRs in this example are close to the previous example, the suboptimal IS filter remains the same.

For several values of $\omega_0 \in [1, 3]$ we evaluated the performance of this frequency estimator by numerical Monte Carlo study. For each ω_0 we simulated 1000 realizations. The frequency estimates based on applying classical ESPRIT and the suboptimal IS filter were computed (the smaller frequency estimates is assigned to ω_1 and the other one is assigned to ω_2). The RMSE of these estimates over the realization set have been evaluated and illustrated in Fig. 6. Since the cisoids are closely spaced the CRLB depends on the actual phase difference between cisoids (see [12, pp. 413–416]). The CRLB plotted in Fig. 6 corresponds to the case of zero phase difference (i.e., $\phi_1 = \phi_2$). However, in simulations phases are random variable uniformly distributed over $[0, 2\pi]$. The improvement is clear especially in the frequency band of interest the RMSE reduced by a factor of 20. \square

VI. CONCLUDING COMMENTS

We have explained how the IS-filtering can be used to estimate the spectral content of signals. This amounts to a generalization of the classical subspace techniques. This generalization introduces tuning parameters that can be used to improve the spectral estimators. In order to analyze various tradeoffs a statistical performance analysis has been carried out. This leads to a design methodology that uses prior information or preprocessing for tuning the spectral estimator. Numerical studies have shown advantages of the proposed algorithms as compared to standard high-resolution algorithms.

VII. PROOFS AND DERIVATIONS

A. Proof of Proposition 2

Since both pairs (A, B) and (\tilde{A}, \tilde{B}) are controllable, share the same eigenvalues, and B, \tilde{B} are column vectors, they relate through a similarity transformation

$$(\tilde{A}, \tilde{B}) = (T^{-1}AT, T^{-1}B)$$

where T is an invertible matrix. (To see that this is so, simply note that they can be both brought into the same canonical form, e.g., a control canonical form, by such a similarity transformation.) Thus, we have

$$I = \tilde{A}\tilde{A}^* + \tilde{B}\tilde{B}^*,$$

$$I = (T^{-1}AT)(T^*A^*T^{-1*}) + (T^{-1}B)(B^*T^{-1*}).$$

If we multiply by T and T^* from the left and right, respectively, we have

$$(TT^*) = A(TT^*)A^* + BB^*.$$

The last equation is a Lyapunov equation, and since A is stable and $BB^* \geq 0$ it has a unique solution. Also $AA^* + BB^* = I$ which implies that the unique solution is the identity and $TT^* = I$. \square

B. Proof of (21)

Normalization in (8) implies that

$$\begin{aligned} G(\omega)G(\omega)^* &= (I - Ae^{-j\omega})^{-1}BB^*(I - A^*e^{j\omega})^{-1} \\ &= (I - Ae^{-j\omega})^{-1}(I - AA^*)(I - A^*e^{j\omega})^{-1} \\ &= I + \sum_{k=1}^{\infty} [c_k e^{-jk\omega} + c_k^* e^{jk\omega}] \end{aligned}$$

where $c_k = A^k \sum_{i=0}^{\infty} A^i (I - AA^*) A^{*i}$. Since A is a stable matrix

$$\sum_{i=0}^{\infty} A^i (I - AA^*) A^{*i} = I.$$

Thus $c_k = A^k$ and we have that

$$\begin{aligned} G(\omega)^*G(\omega) &= \text{tr}(G(\omega)G(\omega)^*) \\ &= \text{tr}((I - Ae^{-j\omega})^{-1}BB^*(I - A^*e^{j\omega})) \\ &= \text{tr}\left(I + \sum_{k=1}^{\infty} [A^k e^{-jk\omega} + A^{*k} e^{jk\omega}]\right) \\ &= n + \sum_{k=1}^{\infty} \sum_{i=1}^n [\lambda_i^k e^{-jk\omega} + \bar{\lambda}_i^k e^{jk\omega}] \\ &= n + \sum_{i=1}^n \left[\frac{\lambda_i e^{-j\omega}}{1 - \lambda_i e^{-j\omega}} + \frac{\bar{\lambda}_i e^{j\omega}}{1 - \bar{\lambda}_i e^{j\omega}} \right] \\ &= \sum_{i=1}^n \frac{1 - \lambda_i \bar{\lambda}_i}{(1 - \lambda_i e^{-j\omega})(1 - \bar{\lambda}_i e^{j\omega})}. \end{aligned}$$

If we set $\lambda_i = r_i e^{j\theta_i}$, we obtain

$$\|G(\omega)G(\omega)^*\| = \sum_{i=1}^n \frac{1 - r_i^2}{1 + r_i^2 - 2r_i \cos(\omega - \theta_i)}.$$

C. Proof of Lemma 1

Below, we repeatedly use the following identities:

$$U_n U_n^* = I - U_s U_s^*, \quad U_s^* U_s = I, \quad \text{and} \quad U_s^* U_n = 0.$$

The following holds:

$$\begin{aligned}
 U_s^* \hat{P} U_n &= U_s^* \left(\hat{U}_s \hat{\Lambda}_s \hat{U}_s^* + \hat{U}_n \hat{\Lambda}_n \hat{U}_n^* \right) U_n \\
 &= \left(U_s^* \hat{\Lambda}_s \right) \hat{\Lambda}_s \left(\hat{U}_s^* U_n \right) + \left(U_s^* \hat{U}_n \right) \hat{\Lambda}_n \left(\hat{U}_n^* U_n \right) \\
 &= \left[\left(U_s^* \hat{\Lambda}_s \right) \hat{\Lambda}_s - \Lambda_s \right] \left(\hat{U}_s^* U_n \right) + \Lambda_s \left(\hat{U}_s^* U_n \right) \\
 &\quad + \left(U_s^* \hat{U}_n \right) \left(\hat{\Lambda}_n - \sigma^2 I \right) \left(\hat{U}_n^* U_n \right) + \sigma^2 U_s^* \hat{U}_n \hat{U}_n^* U_n \\
 &= \Lambda_s \left(\hat{U}_s^* U_n \right) - \sigma^2 U_s^* \hat{U}_s \hat{U}_s^* U_n \\
 &\quad + \left[\left(U_s^* \hat{\Lambda}_s \right) \hat{\Lambda}_s - \Lambda_s \right] \left(\hat{U}_s^* U_n \right) \\
 &\quad + \left(U_s^* \hat{U}_n \right) \left(\hat{\Lambda}_n - \sigma^2 I \right) \left(\hat{U}_n^* U_n \right) \\
 &= \left\{ \left[\left(U_s^* \hat{\Lambda}_s \right) \hat{\Lambda}_s - \Lambda_s \right] - \sigma^2 \left[U_s^* \hat{U}_s - I \right] \right\} \left(\hat{U}_s^* U_n \right) \\
 &\quad + \Lambda_\sigma \left(\hat{U}_s^* U_n \right) + \left(U_s^* \hat{U}_n \right) \left(\hat{\Lambda}_n - \sigma^2 I \right) \left(\hat{U}_n^* U_n \right).
 \end{aligned}$$

Neglecting the higher order terms, which is reasonable for large-samples, we get

$$U_s^* \hat{P} U_n = \Lambda_\sigma \hat{U}_s^* U_n$$

and finally

$$U_n \hat{U}_s = U_n^* \hat{P} U_s \Lambda_\sigma^{-1}.$$

D. Proof of Proposition 3

Note that $U_{s\perp}$ is equal to U_n and that

$$\begin{aligned}
 C &= [c_{ij}]_{i,j=1}^{n-m} := E\{SS^*\} \\
 &= E \left\{ \Lambda_\sigma^{-1} U_s^* \hat{P} U_n U_n^* \hat{P} U_s \Lambda_\sigma^{-1} \right\} \\
 &= \Lambda_\sigma^{-1} U_s^* E \left\{ \hat{P} U_n U_n^* \hat{P} \right\} U_s \Lambda_\sigma^{-1}.
 \end{aligned}$$

Since $U_n U_n^* = \sum_{i=1}^{n-m} g_i g_i^*$, then

$$E \left\{ \hat{P} U_n U_n^* \hat{P} \right\} = \sum_{i=1}^{n-m} E \left\{ \hat{P} g_i g_i^* \hat{P} \right\}.$$

Hence, the problem reduces to estimating $E\{\hat{P} g_i g_i^* \hat{P}\}$. In subsequent calculations we use the fact that $s(k)$ and $\xi(k)$ are independent random variable with zero mean, and also (see [13]) that, for any four complex-valued Gaussian random variables $\{z_1, z_2, z_3, z_4\}$ at least one of which has zero mean

$$\begin{aligned}
 E\{z_1 z_2 z_3 z_4\} &= E\{z_1 z_2\} E\{z_3 z_4\} + E\{z_1 z_3\} E\{z_2 z_4\} \\
 &\quad + E\{z_1 z_4\} E\{z_2 z_3\}.
 \end{aligned}$$

Let g be any column of U_n , i.e., $g \in \{g_1, \dots, g_{n-m}\}$, then $g^* x(k) = g^* \xi(k)$, which follows from the fact that $U_n^* G_s = 0$.

Hence

$$\begin{aligned}
 &E\{\hat{P} g g^* \hat{P}\} \\
 &= E \left\{ \sum_{k=1}^{N_x} \frac{x(k) x^*(k) g}{N_x} \times \sum_{\ell=1}^{N_x} \frac{g^* x(\ell) x^*(\ell)}{N_x} \right\} \\
 &= \frac{1}{N_x^2} \sum_{k=1}^{N_x} \sum_{\ell=1}^{N_x} E \{ x(k) \xi^*(k) g g^* \xi(\ell) x^*(\ell) \} \\
 &= \frac{1}{N_x^2} \sum_{k=1}^{N_x} \sum_{\ell=1}^{N_x} E \{ (G_s(k) + \xi(k)) \xi^*(k) g g^* \xi(\ell) \\
 &\quad \times (G_s(\ell) + \xi(\ell))^* \} \\
 &= \frac{1}{N_x^2} \sum_{k=1}^{N_x} \sum_{\ell=1}^{N_x} E \\
 &\quad \times \{ g^* \xi(\ell) \xi^*(k) g (G_s(k) s^*(\ell) G^* + G_s(k) \xi^*(\ell) \\
 &\quad \quad + \xi(k) s^*(\ell) G^* + \xi(k) \xi^*(\ell)) \} \\
 &= \frac{1}{N_x^2} \sum_{k=1}^{N_x} \sum_{\ell=1}^{N_x} E \{ g^* \xi(\ell) \xi^*(k) g \xi(k) \xi^*(\ell) \} \\
 &\quad + (g^* E \{ \xi(\ell) \xi^*(k) \} g) (G E \{ s(k) s(\ell) \} G^*).
 \end{aligned}$$

It follows that

$$\begin{aligned}
 c_{ij} &= u_i^* \left(\frac{1}{N_x^2} \sum_{k=1}^{N_x} \sum_{\ell=1}^{N_x} g^* Q_{(\ell-k)} g G R_{(\ell-k)} G^* \right) u_j \\
 &\quad + \frac{1}{N_x^2} \sum_{k=1}^{N_x} \sum_{\ell=1}^{N_x} E \{ g^* \xi(\ell) \xi^*(k) g u_i^* \xi(k) \xi^*(\ell) u_j \}.
 \end{aligned}$$

□

Now $E\{g^* \xi(\ell) \xi^*(k) g u_i^* \xi(k) \xi^*(\ell) u_j\}$ can be expanded using the aforementioned result for four complex Gaussian random variables. The other fact that allows further simplification is that $E\{\xi(k) \xi(\ell)^T\} = 0$. Thus, c_{ij} can be written as

$$\begin{aligned}
 c_{ij} &= \frac{1}{N_x^2} \sum_{k=1}^{N_x} \sum_{\ell=1}^{N_x} [g^* Q_{(\ell-k)} g u_i^* G R_{(\ell-k)} G^* u_j \\
 &\quad + g^* Q_{(\ell-k)} g u_i^* Q_{(k-\ell)} u_j + g^* Q_0 u_j u_i^* Q_0 g] \\
 &= \frac{1}{N_x^2} \sum_{\ell=-N_x}^{N_x} (N_x - |\ell|) (g^* Q_\ell g) \\
 &\quad \times (u_i^* G R_\ell^* G^* u_j + u_i^* Q_\ell^* u_j) \\
 &= \frac{1}{N_x^2} \sum_{\ell=-N_x}^{N_x} (N_x - |\ell|) (g^* Q_\ell g) u_i^* (G R_\ell^* G^* + Q_\ell^*) u_j \\
 &= \frac{1}{N_x^2} \sum_{\ell=-N_x}^{N_x} (N_x - |\ell|) (g^* Q_\ell g) u_i^* P_\ell^* u_j.
 \end{aligned}$$

This gives the expression claimed in the proposition. □

E. Derivation of $\gamma(\omega)$

We derive $\gamma(\omega)$ in terms of the IS filter parameters, i.e., A and B . To this end, we define

$$\begin{aligned}
 T_1 &:= (I - e^{-j\omega} A)^{-1}, \\
 T_2 &:= -T_1 (j e^{-j\omega} A) T_1.
 \end{aligned}$$

Using this notation, we have

$$\frac{dG(\omega)}{d\omega} = T_2 B$$

and

$$\frac{d\|G(\omega)\|}{d\omega} = \frac{B^*(T_2^*T_1 + T_1^*T_2)B}{2\|G(\omega)\|}.$$

By inserting the above in

$$\gamma(\omega) = \frac{\frac{dG(\omega)}{d\omega}\|G(\omega)\| - \frac{d\|G(\omega)\|}{d\omega}G(\omega)}{\|G(\omega)\|^2} \quad (27)$$

we get the expression for $\gamma(\omega)$ in terms of the IS filter parameters. \square

F. Determining (A, B)

We want to specify a pair (A, B) corresponding to an IS filter with n modes at prescribed locations. Let a_k for $k = 1, 2, \dots, l$ and p_k denotes the multiplicity of the corresponding mode. Consider the square matrix $J_k^{p_k \times p_k}$ with the following Jordan form

$$J_k = \begin{bmatrix} a_k & 0 & \dots & 0 \\ 1 & a_k & 0 & \dots \\ 0 & 1 & a_k & \dots \\ \vdots & & \ddots & \ddots \\ 0 & \dots & 1 & a_k \end{bmatrix} \quad b_k = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$J = \text{diag}(J_1, J_2, \dots, J_l)$, and $b = [b_1; b_2; \dots; b_k]$, where “;” denotes vertical concatenation. The pair (J, b) is controllable with eigenvalues of J at the prescribed locations but it does not satisfy the normalization condition in (8). A suitable similarity transformation can generate the desired (A, B) from (J, b) . Let E be the solution of the following discrete-time Lyapunov equation $JEJ^* + bb^* = E$ then $(E^{-(1/2)}JE^{1/2}, E^{-(1/2)}b)$ satisfies (8) and gives the desired (A, B) . \square

APPENDIX I

STATISTICAL PERFORMANCE ANALYSIS FOR ARRAY PROCESSING

The statistical properties of the subspace methods, when they apply for direction of arrival (DOA) estimation, have received considerable attention in the literature. Examples of the reported works are Kaveh and Barabell [14], Wang and Kaveh [9], Stoica and Nehorai [15]. Here we consider DOA estimation by means of a uniform linear array (ULA) of sensors. ULA assumption corresponds to the Toeplitz state-covariance case, and the dimension of the covariance matrix, i.e., n , is number of the sensors.

The statistical analysis for array processing involves an assumption that simplifies the analysis as compared to harmonic retrieval. It is assumed that the output of the array is sampled properly so as *snapshots*, which is denoted by $x(k)$ in this paper, are independent and identically distributed. Thus, the autocorrelation function of $x(k)$, i.e., P_ℓ , is zero for $\ell \neq 0$ as well as Q_ℓ .

This simplifies Proposition 3 to the following (see [6, ch. 3]).³ Let the canonical angles between U_s and \hat{U}_s are denoted by $\{\alpha_1, \alpha_2, \dots, \alpha_m\}$ and N_x denotes the number of snapshots

$$\sum_{i=1}^m E\{\sin^2 \alpha_i\} = \frac{\sigma^2(n-m)}{N_x} \sum_{i=1}^m \frac{\lambda_i}{(\lambda_i - \sigma^2)^2}. \quad (28)$$

It is interesting to compare this result to the earlier work on the performance of the subspace methods. In [9, pp. 1204, 1205], the largest canonical angle between true signal subspace and its estimate is used to describe the distance between them. The mean-square error of the cosine of this angle from the perfect value (i.e., 1.0) is then used to evaluate the quality of the estimated signal subspace. For the case of two targets in spatially white noise, approximate closed-form expression is derived in terms of the source and the array parameters. It is shown in [9, p. 1205] that if $n \gg 2$ and $N_x/n \gg 1$

$$E\{(1 - \cos \alpha_{max})^2\}^{\frac{1}{2}} = \frac{\sigma^2(n-2)}{N_x} \times \frac{\lambda_2}{(\lambda_2 - \sigma^2)^2}.$$

If we set $m = 2$ in (28) to compare the results, we get

$$\frac{\sigma^2(n-2)}{N_x} \left(\frac{\lambda_1}{(\lambda_1 - \sigma^2)^2} + \frac{\lambda_2}{(\lambda_2 - \sigma^2)^2} \right).$$

There is a meaningful similarity between these two results. However, (28) considers all of the canonical angles not only the largest one, and also it is applicable for the case of more than two sources.

ACKNOWLEDGMENT

The authors would like to thank Prof. M. Kaveh, Prof. G. Giannakis, and anonymous reviewers for their valuable comments and suggestions.

REFERENCES

- [1] S. Haykin, *Nonlinear Methods of Spectral Analysis*. New York: Springer-Verlag, 1983.
- [2] P. Stoica and R. Moses, *Introduction to Spectral Analysis*. Upper Saddle River, NJ: Prentice-Hall, 1997.
- [3] T. T. Georgiou, “Signal estimation via selective harmonic amplification: music, redux,” *IEEE Trans. Signal Process.*, vol. 48, no. 11, pp. 780–790, Nov. 2000.
- [4] —, “Spectral estimation via selective harmonic amplification,” *IEEE Trans. Autom. Control*, vol. 46, no. 1, pp. 29–42, Jan. 2001.
- [5] H. L. V. Trees, *Optimum Array Processing*. New York: Wiley, 2002.
- [6] A. N. Amini, “State covariance spectral estimation,” Master’s thesis, Elect. and Computer Eng. Dept., Univ. Minnesota, Minneapolis, 2003.
- [7] G. Stewart and J. Sun, *Matrix Perturbation Theory*. New York: Academic, 1990.
- [8] P. Stoica and T. Söderström, “Statistical analysis of MUSIC and subspace rotation estimates of sinusoidal frequencies,” *IEEE Trans. Signal Process.*, vol. 39, no. 8, pp. 1836–1847, Aug. 1991.
- [9] M. Kaveh and H. Wang, “On the performance of signal-subspace processing—Part I: Narrow-band systems,” *IEEE Trans. Signal Process.*, vol. 34, no. 10, pp. 1201–1209, Oct. 1986.
- [10] P. Stoica and A. Nehorai, “MUSIC, maximum likelihood, and Cramer–Rao bound,” *IEEE Trans. Signal Process.*, vol. 37, no. 5, pp. 720–741, May 1989.

³For $m > 1$ since $E\{S^*S\}$ is a matricial object, $\text{tr}(E\{S^*S\})$ can be applied as a scalar statistical measure of the error in the estimated signal subspace. This object has a nice geometric interpretation; it is the sum of the sines of the canonical angles. To see this, let $\{\alpha_1, \alpha_2, \dots, \alpha_m\}$ be the nonzero canonical angles between U_s and \hat{U}_s . Then we have

$$\text{tr}(E\{S^*S\}) = E\{\text{tr}(S^*S)\} = E\left\{\sum_{i=1}^m \sin^2 \alpha_i\right\} = \sum_{i=1}^m E\{\sin^2 \alpha_i\}.$$

- [11] A. Nasiri Amini, E. Ebbini, and T. Georgiou, "Noninvasive estimation of tissue temperature via high-resolution spectral analysis techniques," *IEEE Trans. Biomed. Eng.*, vol. 52, no. 2, pp. 221–228, Feb. 2005.
- [12] S. Kay, *Modern Spectral Estimation*. Englewood Cliffs, NJ: Prentice-Hall, 1988.
- [13] P. Janssen and P. Stoica, "On the expectation of the product of four matrix-valued Gaussian random variables," *IEEE Trans. Autom. Control*, vol. 33, no. 9, pp. 867–870, Sep. 1998.
- [14] M. Kaveh and A. Barabell, "The statistical performance of the MUSIC and minimum-norm algorithms in resolving plane waves in noise," *IEEE Trans. Signal Process.*, vol. 34, no. 4, pp. 331–341, Apr. 1986.
- [15] P. Stoica and A. Nehorai, "Performance comparison of subspace rotation and MUSIC methods for direction estimation," *IEEE Trans. Signal Process.*, vol. 39, no. 2, pp. 446–453, Feb. 1991.



Ali Nasiri Amini (S'01–M'05) was born in Tehran, Iran, in 1979. He received the B.Sc. degree in electrical engineering from the Sharif University of Technology, Tehran, in 1999 and the M.Sc. and Ph.D. degrees in electrical engineering from the University of Minnesota, Minneapolis, in 2002 and 2005, respectively.

He is currently working toward the Ph.D. degree with the Department of Electrical and Computer Engineering, University of Minnesota. His research interests include statistical signal processing, statistical data analysis, biomedical engineering, and control theory.



Tryphon T. Georgiou (F'00) was born in Athens, Greece, on October 18, 1956. He received the Diploma in mechanical and electrical engineering from the National Technical University of Athens, Greece, in 1979, and the Ph.D. degree from the University of Florida, Gainesville, in 1983.

He has been a member of the faculty at Florida Atlantic (1983–1986) and Iowa State (1986–1989) Universities. Since August 1989, he has been with the University of Minnesota, Minneapolis, where he is a Professor of Electrical and Computer Engineering.

Dr. Georgiou has served as an Associate Editor for *IEEE TRANSACTIONS ON AUTOMATIC CONTROL*, the *SIAM Journal on Control and Optimization*, *Systems and Control Letters* and on the Board of Governors of the Control Systems Society of the IEEE. He has been a corecipient of three George Axelby Outstanding Paper awards for the years 1992, 1999, and 2003. The first two times were for joint work with Prof. M. C. Smith (Cambridge University, U.K.), and the third time for joint work with Prof. C. Byrnes (Washington University, St. Louis) and Prof. A. Lindquist (KTH, Stockholm).