Spectral Estimation via Selective Harmonic Amplification

Tryphon T. Georgiou, *Fellow, IEEE*

Abstract— **The state-covariance of a linear filter is characterized by a certain algebraic commutativity property with the state matrix of the filter and also imposes a generalized interpolation constraint on the power spectrum of the input process. This algebraic property and the relationship between state-covariance and the power spectrum of the input allow the use of matrix pencils and of analytic interpolation theory for spectral analysis. Several algorithms for spectral estimation will be developed with resolution higher than state of the art.**

I. Introduction

In this paper we consider the following two basic questions. Suppose we are given a known linear filter driven by some unknown stochastic process, *what can we say apriori about the structure of the state-covariance of the filter?* Then, suppose we measure the state-covariance, *what can we say about the power spectrum of the input process?* Both questions turn out to have interesting ramifications in the context of spectral estimation. Several algorithms for high resolution spectral estimation will be based on the answers to these two questions, summarized as follows:

(i) the state-covariance of a linear filter is characterized by a certain algebraic commutativity relation with the state matrix of the filter (Theorem 1).

(ii) the state-covariance imposes an interpolation condition on the power spectrum of the input (Theorem 2).

The first result allows the use of matrix pencils in order to decompose the state-covariance into components corresponding to particular input signals (Corollary 3 and Theorem 5). Such analysis can be used to identify dominant components of the input. The second result is a generalization of an observation in [4, 5], where it was shown that the output covariance of a first-order filter relates to the value of the input power spectrum at the pole of the filter. Knowing the state-covariance, analytic interpolation [7, 20, 21] provides a parametrization of all possible input power spectra (Section III). Particular interpolants are considered which correspond to emission and absorption spectra, and the envelope of the input power spectrum can be constructed based on the same theory (Sections IV, VI, VII). Several alternative methods for generating particular spectra will be presented.

Our general methodology as it pertains to spectral estimation requires a filtering stage where an *input-to-state* (IS-) filter is used and the state-covariance is estimated from time-series data. The filter may be chosen to weigh more heavily the spectral content of a selected harmonic interval. As a consequence spectral analysis based on state-covariance statistics allows for higher resolution over selected intervals. The analysis makes use of the algebraic structure and of a canonical decomposition of state-covariances. The particular decomposition is consistent with the hypothesis that the input is made up of (a minimal number of) sinusoids in background noise and is a generalization of a classical result of Carathéodory and Féjer for Toeplitz matrices. As far as computation is concerned, it requires the singular value decomposition (SVD) of the state-covariance while the specifics follow analogous decompositions of Toeplitz matrices which underly several modern nonlinear spectral estimation methods [13].

In Section II we briefly survey the framework underlying modern nonlinear spectral estimation techniques. This requires a finite number of estimated covariance lags and the relevant theory has connections to the moment problem and orthogonal polynomials (e.g., [2,10,12,15,24]). We outline the techniques which, in subsequent sections, will be generalized to the context of state-covariance data so as to highlight the analogies. In particular, we discuss subspace methods which rely on a suitable eigen-decomposition of a Toeplitz matrix ([18]; see [25]) and then introduce a dual solution, again based on a decomposition of a Toeplitz matrix, for a canonical absorption spectrum. This material is novel since the current literature contains virtually no method for identifying absorption spectra. Finally, we discuss the Capon method, point out that the relevant "spectral estimate" truly represents envelope of spectral power, and mention an enhancement for real-valued processes given in [23].

Section III deals with the algebraic structure of state-covariances and their interpolation theoretic significance (also noted in [11]). Theorems 1 and 2 show how covariances relate to generalized interpolation constraints for the input spectrum and the section concludes with an explanation of how to obtain such constraints numerically from estimated state-covariances.

In Section IV we present a canonical decomposition theorem. This is a generalization of an analogous result in [11] and forms the basis of our computational approach. It is also a generalization of a classical result by Carathéodory and Féjer going back to the beginning of the 20th century [6]. The Carathéodory and Féjer result was rediscovered in more recent years by Pisarenko [18] and forms the basis of subspace methods for the estimation of spectral lines.

Subsequently, in Sections V, VI, and VII, we develop analogs of techniques surveyed in Section II to the new state-covariance framework. More specifically, in Section V

Department of Electrical and Computer Engineering, University of Minnesota, MN 55455. E-mail: tryphon@ece.umn.edu

This research was supported in part by the NSF and AFOSR.

we develop subspace methods which are analogous to the widely used MUSIC (MUltiple SIgnal Classification), and ESPRIT (EStimation of Parameters by Rotational Invariant Techniques), in Section VI we discuss how to generate canonical absorption spectra based on state-covariance data, and in Section VII we present an analog of the Capon Method as well as an analog of [23] to the state-covariance setting.

In Section VIII we discuss two alternatives for quantifying the selectivity of IS-filters and suggest simple realizations for IS-filters with desired properties and finally, in Section IX, we present simulation results that demonstrate the performance of the new techniques and point to a substantial improvement over current state of the art.

II. The Toeplitz-Carathéodory framework

The purpose of this section is to give a brief survey of important concepts and techniques which will subsequently be generalized to the context interpolation with state-covariance data. At the same time we introduce notation and review certain basic facts.

Consider a discrete-time, wide-sense stationary, zero-mean, scalar stochastic process $\mathbf{y}_t$ with $t \in \mathbb{Z}$, and denote its power spectrum by $d\sigma_y(\theta)$. In general, the power spectrum is a positive bounded measure on $[-\pi, \pi]$. The regular part of $d\sigma_y(\theta)$, which is defined almost everywhere, is $d\sigma_{y,\text{regular}} = \dot{\sigma}_y(\theta)d\theta$ (where $\dot{\sigma}_y(\theta)$ is the *spectral density function*), while the remaining singular part contains any spectral lines which may be present (e.g., Dirac functions $\delta(\cdot)$). In the simplest case, when it is of finite-type, then it is precisely of the form

$$d\sigma_{y,\text{singular}}(\theta) = \sum_{k=1}^{m} 2\pi\rho_k \delta(\theta - \theta_k)d\theta, \tag{1}$$

with θ_k's representing the frequencies of sinusoids embedded in $\mathbf{y}_t$.

The covariance lags of $\mathbf{y}_t$

$$c_k := E\left\{\mathbf{y}_\ell \overline{\mathbf{y}}_{\ell+k}\right\},\ k = 0, 1, \ldots, n-1,$$

are simply trigonometric *moments*

$$c_k = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-jk\theta} d\sigma_y(\theta), \text{ for } k = 0, 1, \ldots, \tag{2}$$

of the power spectrum. The theory of moments [12, 24] provides tools for characterizing admissible power spectra $d\sigma_y$ from a finite number of such covariances c_k ($k = 0, 1, \ldots, n-1$). This is usually referred to as the Carathéodory problem.

We will explore a connection with analytic interpolation which is based on the following fact (e.g., see [12]): $d\sigma_y(\theta)$ *is a bounded positive measure if and only if the function*

$$f_y(\lambda) := \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{1 + \lambda e^{-j\theta}}{1 - \lambda e^{-j\theta}} d\sigma_y(\theta) \tag{3}$$

$$= c_0 + 2c_1\lambda + \ldots + 2c_{n-1}\lambda^{n-1} + \ldots \tag{4}$$

is analytic in the open unit disc with positive real part. Such a function will be referred to as *positive-real.* Starting from $f_y(\lambda)$ the power spectrum can be recovered via

$$d\sigma_y(\theta) \sim \lim_{r\to 1} \Re(f_y(re^{j\theta}))d\theta, \tag{5}$$

where $\Re$ denotes "the real part of". Thus, (3-5) establish a one-to-one correspondence between covariance sequences

$$\mathbf{c} := (c_0, c_1, \ldots, c_{n-1}, c_n, \ldots),$$

positive real functions $f_y(\lambda)$, and bounded positive measures $d\sigma_y(\theta)$.

The "half-spectrum" function $f_y(\lambda)$ is central to the analysis in this paper. Spectral estimation based on a partial covariance sequence

$$\mathbf{c}_{n-1} := (c_0, c_1, \ldots, c_{n-1}),$$

amounts to an interpolation problem with a positive real function $f(\lambda)$ satisfying n interpolation conditions at the origin for its 0th to $(n-1)$th first derivatives. All solutions to such a problem can be described by a linear fractional transformation, e.g., see [1]. Historically, a particular solution, corresponding to a singular σ_y of the form (1) plus a constant, was used in the context of the Carathéodory problem [6] to show existence of solutions under suitable conditions (i.e., non-negativity of a Toeplitz form made up of the c_k's). The same solution was rediscovered in [18] and forms the basis of high resolution subspace methods. This is discussed next.

A. The Carathéodory-Fejér/Pisarenko solution

The Carathéodory-Fejér-Pisarenko (CFP) result states that if a positive semi-definite $n \times n$ (Hermitian) Toeplitz matrix

$$T_{n-1} = \begin{bmatrix} c_0 & c_1 & \cdots & c_{n-1} \\ c_{-1} & c_0 & \cdots & c_{n-2} \\ \vdots & \vdots & & \vdots \\ c_{-n+1} & c_{-n+2} & \cdots & c_0 \end{bmatrix}$$

is singular and of rank m $(< n)$, then it is of the form

$$T_{n-1} = \sum_{i=1}^{m} \rho_i G(e^{j\theta_i})G(e^{j\theta_i})^*$$

where $\rho_i > 0$, for $i = 1, 2, \ldots, m$, $\theta_\ell \neq \theta_k$ if $\ell \neq k$,

$$G(e^{j\theta_i}) = \begin{bmatrix} 1 \\ e^{j\theta_i} \\ \vdots \\ e^{j(n-1)\theta_i} \end{bmatrix}, \tag{6}$$

and "*" denotes "complex-conjugate transpose". Moreover, T_{n-1} has a unique (left) zero-eigenvector of the form

$$\phi = \left[\begin{array}{cccccc} a_0 & \cdots & a_m & 0 & \ldots & 0 \end{array}\right].$$

Since

$$0 \;=\; \phi T_{n-1}\phi^* = \sum_{i=1}^{m} \rho_i |\phi \cdot G(e^{j\theta_i})|^2,$$

it follows that the modes $e^{j\theta_i}$ are precisely the roots of the polynomial

$$\phi(\lambda) = a_0 + a_1\lambda + \ldots + a_m\lambda^m.$$

If we now define a polynomial $\psi(\lambda)$ to contain the first n powers of λ in the product

$$\phi(\lambda)(c_0 + 2c_1\lambda + \ldots + 2c_{n-1}\lambda^{n-1}),$$

then it turns out that $\psi(\lambda)$ is also of degree m, i.e.,

$$\psi(\lambda) = b_0 + b_1\lambda + \ldots + b_m\lambda^m,$$

and that

$$f_0(\lambda) := \frac{\psi(\lambda)}{\phi(\lambda)} = c_0 + 2c_1\lambda + \ldots + 2c_{n-1}\lambda^{n-1} + o(\lambda^n)$$

is a positive-real function. In fact,

$$f_0(\lambda) = \sum_{i=1}^{m} \rho_i \frac{1+\lambda e^{j\theta_i}}{1-\lambda e^{j\theta_i}}.$$

The real part of $f_0(\lambda)$ is zero at all points on the unit circle except the poles of $f_0(\lambda)$ (cf. [12, page 148]). The polynomials $\phi(\lambda)$ and $\psi(\lambda)$ (or rather, $\lambda^m\phi(\lambda^{-1})$ in the notation of Geronimus [10, page 10] and [12, page 40]) are precisely the m-th "orthogonal Szegö" polynomials of the first and second kind, respectively, associated with the *singular* quadratic form defined by T_{n-1} and they both have all their roots *on* the unit circle (cf. [10], [12]).

These facts directly yield a solution $f_y(\lambda)$ to the Carathéodory interpolation problem: If $T_{n-1} > 0$ and $\tau \geq 0$ denotes its lowest eigenvalue, then apply the above analysis to the *singular* positive semi-definite Toeplitz matrix

$$T_{n-1} - \tau I,$$

(with I the identity matrix of appropriate size,) to obtain ϕ, ψ for this new set of data $(c_0 - \tau, c_1, \ldots, c_n)$. Then the function

$$f_p(\lambda) := \tau + \frac{\psi(\lambda)}{\phi(\lambda)}$$

is positive-real, satisfies the interpolation constraints in (4), and the corresponding spectrum/measure is singular and given by

$$d\sigma(\theta) = \tau d\theta + \sum_{k=1}^{m} 2\pi\rho_k \delta(\theta - \theta_k) d\theta.$$

Thus, it consists entirely of spectral lines on a white noise "flat" background.

The detailed theory and derivation of the above can be found in [12, page 151]; see also [25, page 155] for an independent derivation.

B. Subspace Methods

An elegant and useful reformulation of the above [22], to which we will return in due time, makes use of the elements in the singular value decomposition

$$T_{n-1} - \tau I = U \mathrm{diag}\{\sigma_1, \sigma_2, \ldots, \sigma_m, 0, \ldots, 0\} U^*,$$

where U is a unitary matrix and $\sigma_i > 0$ $(i = 1, \ldots, m)$. If

$$U = [U_{1:m}, U_{m+1:n}]$$

and $U_{m+1:n}$ represents the matrix formed out of the last $n-m$ columns of U which span the (right) null-space of T_{n-1}, then the non-negative trigonometric polynomial

$$d(e^{j\theta}, e^{-j\theta}) := G(e^{j\theta})^* U_{m+1:n} U_{m+1:n}^* G(e^{j\theta}) \qquad (7)$$

has m roots precisely at $\{\theta_1, \ldots, \theta_m\}$. Note that $d(\lambda, \lambda^{-1})$ may have additional roots off the unit circle, but has no root on the circle other than the above. This result underlies the methodology of MUSIC and a proof can be found in [25, page 157].

Focusing on the signal subspace, the range of $T_{n-1} - \tau I$ is spanned by the columns of both

$$\mathbf{G}_s := \begin{bmatrix} 1 & 1 & \ldots & 1 \\ e^{j\theta_1} & e^{j\theta_2} & \ldots & e^{j\theta_m} \\ \vdots & \vdots & & \vdots \\ e^{j(n-1)\theta_1} & e^{j(n-1)\theta_2} & \ldots & e^{j(n-1)\theta_m} \end{bmatrix}$$

as well as the columns of $U_{1:m}$. Thus,

$$\mathbf{G}_s = U_{1:m}\Theta$$

where Θ is an invertible matrix. The method of ESPRIT [19] is based on the following observation: if

$$R_{\mathrm{top}} := [I_{n-1}, 0], \text{ and } R_{\mathrm{bottom}} := [0, I_{n-1}],$$

where 0 denotes a zero column of appropriate size, and I_{n-1} the $(n-1)\times(n-1)$ identity matrix, then

$$R_{\mathrm{top}}\mathbf{G}_s \begin{bmatrix} e^{j\theta_1} & 0 & \ldots & 0 \\ 0 & e^{j\theta_2} & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & e^{j\theta_m} \end{bmatrix} = R_{\mathrm{bottom}}\mathbf{G}_s.$$

Moreover, the range of $\mathbf{G}_s$ is the same as the range of $U_{m+1:n}$. It follows that the diagonal matrix $\mathrm{diag}\left(e^{j\theta_1}, \ldots, e^{j\theta_m}\right)$ is similar to the matrix

$$(R_{\mathrm{top}}U_{1:m})^\sharp (R_{\mathrm{bottom}}U_{1:m})$$

where $\sharp$ denotes "left inverse". Hence the frequencies θ_k, $k = 1, \ldots, m$, can be readily computed. In practice these can be computed as the eigenvalues of the "least squares " solution X of $(R_{\mathrm{top}}U_{1:m})X = (R_{\mathrm{bottom}}U_{1:m})$, while U is obtained by an SVD of the sampled Toeplitz matrix. This is the method of ESPRIT (cf. [25, p. 164]).

C. The dual absorption spectrum

Very much as the Carathéodory-Fejér-Pisarenko solution corresponds to a singular spectrum, which is made up of spectral lines, the solution we present below gives rise to an alternative canonical spectrum which has absorption lines instead of emission lines.

The key idea relies on the fact that *a function $f(\lambda)$ is positive-real if and only if $(f(\lambda))^{-1}$ is positive-real.* The proof of this fact is straightforward. Simply observe that if a function is positive-real it cannot have any roots inside the open unit disc. Thus, its inverse is analytic as well. Then observe that the inverse of any complex number $a+jb$, with $a>0$, is $\frac{a-jb}{a^2+b^2}$ which again has positive real part.

Using the above observation we are led to the following construction. Given the partial covariance sequence $\mathbf{c}_{n-1}$, consider the "analytic" upper triangular portion of $2T_{n-1}$:

$$W=\begin{bmatrix} c_0 & 2c_1 & 2c_2 & \dots & 2c_{n-1}\\ 0 & c_0 & 2c_1 & \dots & 2c_{n-2}\\ \vdots & \vdots & \vdots & & \vdots\\ 0 & 0 & 0 & \dots & c_0\end{bmatrix}.$$

Then define by means of its inverse

$$W^{-1}=\begin{bmatrix} \hat{c}_0 & 2\hat{c}_1 & 2\hat{c}_2 & \dots & 2\hat{c}_{n-1}\\ 0 & \hat{c}_0 & 2\hat{c}_1 & \dots & 2\hat{c}_{n-2}\\ \vdots & \vdots & \vdots & & \vdots\\ 0 & 0 & 0 & \dots & \hat{c}_0\end{bmatrix},$$

the *inverse sequence*

$$\hat{\mathbf{c}}_{n-1}=(\hat{c}_0,\hat{c}_1,\dots,\hat{c}_{n-1}).$$

I.e., $\hat{c}_0=c_0^{-1}$, $\hat{c}_1=-c_1/c_0^2$, etc. This is also a partial covariance sequence. To see this note that the corresponding Toeplitz matrix, which is given by

$$\begin{aligned}\hat{T}_{n-1} &= \frac{1}{2}\left(W^{-1}+(W^{-1})^*\right)\\ &= W^{-1}\frac{(W+W^*)}{2}(W^{-1})^*\\ &= W^{-1}T_{n-1}(W^{-1})^*, \qquad (8)\end{aligned}$$

is positive semi-definite.

We now apply the CFP solution to this new set of data and obtain a positive-real function $\hat{f}(\lambda)$ which interpolates the values $\hat{\mathbf{c}}_{n-1}$:

$$\hat{f}(\lambda)=\hat{\tau}+\frac{\hat{\psi}(\lambda)}{\hat{\phi}(\lambda)}.$$

Here, $\hat{\tau}$ is the smallest eigenvalue of $\hat{T}_{n-1}$ and $\hat{\phi}(\lambda)$, $\hat{\psi}(\lambda)$ are the m-th orthogonal polynomials (of first and second kind, respectively) of the sequence $(\hat{c}_0-\hat{\tau},\hat{c}_1,\dots)$. We finally arrive at a canonical spectrum, consistent with the original data $\mathbf{c}_{n-1}$, which is dominated by absorption lines. This corresponds to the positive real function

$$f_z(\lambda):=\hat{f}(\lambda)^{-1}=\frac{\hat{\phi}(\lambda)}{\hat{\tau}\hat{\phi}(\lambda)+\hat{\psi}(\lambda)}. \qquad (9)$$

The corresponding spectrum is regular, i.e., without any spectral lines, and is given by

$$d\sigma(\theta) = \Re\{f_z(e^{j\theta})\}d\theta=\hat{\tau}\frac{|\hat{\phi}(e^{j\theta})|^2}{|\hat{\tau}\hat{\phi}(e^{j\theta})+\hat{\psi}(e^{j\theta})|^2}d\theta \qquad (10)$$

It contains absorption lines at the roots of $\hat{\phi}(\lambda)$ which are all on the unit circle.

D. Spectral Envelopes

The so-called *Capon Method* is usually introduced as follows. Given the Toeplitz covariance T_{n-1} of a stochastic process $\mathbf{y}_k$, and any frequency θ, a finite-impulse response (FIR) filter

$$u_k=\sum_{\ell=0}^{n-1}h_\ell y_{k-\ell}$$

is sought, which allows exponential signals $e^{jk\theta}$ at the given frequency to pass unscathed, i.e.,

$$\sum_{\ell=0}^{n-1}h_\ell e^{j\ell\theta}=1,$$

while it minimizes the "total output energy"

$$\begin{aligned}\rho(\theta) &= E\{|u_k|^2\}\\ &= \begin{bmatrix} h_0 & \dots & h_{n-1}\end{bmatrix} T_{n-1}\begin{bmatrix}\bar{h}_0\\ \vdots\\ \bar{h}_{n-1}\end{bmatrix}.\end{aligned}$$

The envelope $\rho(\theta)$, plotted over all θ's, is then taken as a spectral estimate [25, p. 197].

Our terminology *spectral envelope* originates in the *little known fact* that

$$\begin{aligned}\rho(\theta) = \sup\Big\{&\lim_{\epsilon\to 0}\left(\sigma(\theta+\epsilon)-\sigma(\theta-\epsilon)\right)\ :\ \text{where } d\sigma\geq 0\\ &\text{and satisfies (2) for } 0\leq k\leq n-1.\Big\}. \qquad (11)\end{aligned}$$

In other words, $\rho(\theta)$ represents the maximal spectral "mass" located at θ which is consistent with the data, or equivalently, it represents the amplitude of the largest possible spectral line at θ amongst admissible spectra.

Either way, the solution can be obtained by least squares by solving

$$\rho(\theta)=\min_h\left\{hT_{n-1}h^*: h=(h_0,\dots,h_{n-1}), hG(e^{j\theta})=1\right\},$$

or

$$\rho(\theta)=\max\left\{\rho\ :\ T_{n-1}-\rho G(e^{j\theta})G(e^{j\theta})^*\geq 0\right\}, \qquad (12)$$

respectively. Of course, in both cases the solution is the same, given by

$$\rho(\theta)=\frac{1}{G(e^{j\theta})^*T_{n-1}^{-1}G(e^{j\theta})},$$

and coincides with the (unique) eigenvalue of the matrix pencil

$$T_{n-1}-\rho G(e^{j\theta})G(e^{j\theta})^*.$$

Several alternative expressions can be found in [10], e.g., in terms of orthogonal polynomials or a suitable reproducing kernel. Also, in the classical work of Geronimus [10], one can find a detailed study of the asymptotic properties of such a quantity when the length of the partial sequence $\mathbf{c}_{n-1}$ increases towards ∞.

For our current purposes we need the formalism in (11). In case $\mathbf{y}_k$ is known to be real, then only spectra $d\sigma(\cdot)$ with an even symmetry need to be considered. The maximal spectral mass at θ is given by

$$\max\left\{\rho \,:\, T_{n-1} - \rho\Im\left\{G(e^{j\theta})G(e^{j\theta})^*\right\} \geq 0\right\},$$

where

$$\Im\left\{G(e^{j\theta})G(e^{j\theta})^*\right\} := \frac{1}{2}\left(G(e^{j\theta})G(e^{j\theta})^* + G(e^{-j\theta})G(e^{-j\theta})^*\right). \quad (13)$$

This notation is non-standard and obviously differs from the "Hermitian part of" since $G(e^{j\theta})G(e^{j\theta})^*$ is already Hermitian. Equivalently, $\rho(\theta)$ is given by the smallest of the two eigenvalues of the matrix pencil

$$T_{n-1} - \rho\Im\left\{G(e^{j\theta})G(e^{j\theta})^*\right\}.$$

These "real" envelopes were introduced in [23].

We now have concluded our sketchy overview of the classical Toeplitz-Carathéodory framework and we are ready to introduce the key new elements of our approach.

III. Structure of State-covariances and Generalized Interpolation

Consider a single-input/n-output dynamical system, where the output is the n-state vector itself:

$$x_k = Ax_{k-1} + by_k, \text{ where } k \in \mathbb{Z}. \quad (14)$$

The pair A, b, where A is an $n\times n$ matrix and b an $n\times 1$ vector, is assumed to be controllable and A is assumed to be a stable matrix (i.e., with eigenvalues in the interior of the unit disc). Such a system will be referred to as an *input-to-state filter (IS-filter)*. We use λ to designate the delay operator $x_k \to x_{k-1}$, and hence, the transfer function of (14) is

$$G(\lambda) = (I - \lambda A)^{-1} b = \begin{bmatrix} g_1(\lambda) \\ \vdots \\ g_n(\lambda) \end{bmatrix}. \quad (15)$$

From this point on, analogies with the Toeplitz-Carathéodory framework can be sought by specializing to the case where A is the *companion* matrix

$$A = \begin{bmatrix} 0 & 0 & \dots & 0 & 0 \\ 1 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \end{bmatrix}, \text{ and } b = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \quad (16)$$

Obviously, in this case

$$g_i(\lambda) = \lambda^{i-1} \text{ for } i = 1, 2, \dots, n.$$

Thus, the notation $G(e^{j\theta})$ in, e.g., equation (6), is consistent with the above and intended to highlight analogies with the forthcoming material.

Another interesting case is when

$$g_i(\lambda) = \frac{1}{1 - p_i\lambda} \text{ for } i = 1, \dots, n.$$

Then $G(\lambda)$ is a bank of first order filters while A can be taken to be diagonal. Such a filter bank was used in [4, 5]. In general, A may have an arbitrary Jordan structure (subject to the controllability requirement).

Now let the IS-filter in (14) be driven by $\mathbf{y}_t$. The covariance of the state-process

$$P = \frac{1}{2\pi}\int_{-\pi}^{\pi} (G(e^{j\theta})d\sigma_y(\theta)G(e^{j\theta})^*) \quad (17)$$

is of course a positive semi-definite $n\times n$ matrix. But, besides being positive semi-definite, state-covariances are characterized by a certain algebraic condition which depends on the filter parameters and is given below.

Theorem 1: A positive semi-definite matrix P is a state-covariance matrix of (14) for a suitable input process if and only if it is of the form

$$P = \frac{1}{2}(WE + EW^*) \quad (18)$$

for a matrix W which commutes with A, with E being the unique (positive definite) solution to the Lyapunov equation $E - AEA^* = bb^*$. Furthermore, any such matrix W is uniquely defined from (18) modulo an imaginary constant αI with $\alpha \in j\mathbb{R}$.

Proof: Deferred to the appendix. ■

The $n\times n$ matrix W specifies generalized interpolation conditions for admissible input power spectra as claimed in the following theorem.

Theorem 2: Let P, W be as in Theorem 1. If a power spectrum $d\sigma_y(\theta)$ satisfies (17), then the positive real function (3) satisfies

$$f_y(A^*) = W^*. \quad (19)$$

Conversely, if $f_y(\lambda)$ is positive real and satisfies (19), then (5) defines a power spectrum such that (17) holds.

Proof: Deferred to the appendix. ■

Since W commutes with A, then

$$W = w(A) = w_0 I + w_1 A + \ldots + w_{n-1}A^{n-1} \quad (20)$$

with $w(\lambda) = w_0 + \ldots + w_{n-1}\lambda^{n-1}$ a polynomial of degree $n-1$. Thus, according to Theorem 2, the property which characterizes "half-spectra" consistent with P, is that they agree with $w(\lambda)$ on the eigen-spectrum of A^*. An alternative way of writing (19) is either

$$f(\lambda) - w(\lambda) = B(\lambda)q(\lambda), \quad (21)$$

with $B(\lambda)$ the Blaschke product

$$B(\lambda) = \frac{\det(\lambda I - A^*)}{\det(I - \lambda A)} \quad (22)$$

and $q(\lambda)$ a function which is analytic in the open unit disc, or simply, to say that $f(\lambda)-w(\lambda)$ vanishes at the eigenvalues of A^*, taking multiplicities into account. Either way, this is a standard generalized interpolation problem and all solutions can be expressed via linear fractional transformations on a "free" parameter [21, 26]. In particular, the formulae in [8, Section 8.3] can be readily modified to parametrize positive-real interpolants analytic on the disc. (The formulae as they appear in [8, Section 8.3] apply to contractive interpolants which are analytic in the RHP, hence the Möbius map $z \to w = \frac{1-z}{1+z}$ has to be used to transform both their range and domain.)

The following is an immediate corollary to Theorem 1 which is central to the proof of several subsequent results.

Corollary 3: If P, Q are state-covariance matrices for the IS filter (14) such that $P - Q \geq 0$, then $P - Q$ is also a state-covariance for the same filter.

Proof: According to Theorem 1, there exist matrices W_P and W_Q which commute with A and satisfy

$$\begin{aligned} P &= \frac{1}{2}(W_P E + E W_P^*), \text{ and} \\ Q &= \frac{1}{2}(W_Q E + E W_Q^*), \end{aligned}$$

respectively. Hence, their difference is of the same form, i.e., $P-Q = \frac{1}{2}(W_\Delta E + E W_\Delta^*)$, for $W_\Delta = W_P - W_Q$. Since $P - Q \geq 0$ the conclusion of the corollary follows from Theorem 1. ■

Finally, we conclude the section by indicating that equation (18) can be used to obtain the interpolation datum W. Consider

$$\begin{aligned} 2P &= \overline{w}_{n-1}(A^*)^{n-1}E + \ldots + \overline{w}_1 A^* E + \\ & (\overline{w}_0 + w_0)I + w_1 EA + \ldots + w_{n-1}EA^{n-1} \quad (23) \end{aligned}$$

as a set of n^2 linear equations in $2n-1$ parameters – the real and imaginary parts of w_k's for $k = 1, \ldots, n-1$ and $\Re\{w_0\}$. Provided P is a state-covariance matrix as hypothesized, according to Theorem 1 equation (23) has a solution which is unique modulo the imaginary part of w_0. The imaginary part of w_0 cannot be determined and is irrelevant anyhow. If P has been obtained from experimental data and does not exactly satisfy (23), then a least-squares solution for the coefficients $\Re\{w_0\}, w_1, \ldots, w_{n-1}$ can be sought, from which one can determine an estimate for a suitable $W = w(A)$ via (20).

IV. Canonical decomposition

A complex sinusoidal input $e^{jk\theta}$, $k \in \mathbb{Z}$, to the IS-filter in (14), gives rise to a state-covariance equal to $G(e^{j\theta})G(e^{j\theta})^*$ and of rank one. Conversely, a singular state-covariance can only originate from $m < n$ such distinct sinusoidal input components. We highlight this fact as a separate lemma since it is of independent interest.

Lemma 4: [11, Proposition 1]The $n \times n$ state-covariance matrix defined in (17) is singular of rank $m < n$ if and only if it is of the form

$$P = \sum_{k=1}^{m} \rho_k G(e^{j\theta_k})G(e^{j\theta_k})^*, \quad (24)$$

for a selection of distinct θ_k, and $\rho_k > 0$, for $k = 1, \ldots, m$.

Building on this lemma and on Corollary 3 the following canonical decomposition of state-covariance matrices can be shown.

Theorem 5: Let P, Q be state-covariances of the IS-filter (14). Then there exists a unique minimal set of $m < n$ distinct values θ_k along with corresponding values $\rho_k > 0$ $(k = 1, 2, \ldots, m)$, such that

$$P = \rho_0 Q + \sum_{k=1}^{m} \rho_k G(e^{j\theta_k})G(e^{j\theta_k})^*. \quad (25)$$

Moreover, ρ_0 is the smallest eigenvalue of the matrix pencil $P - \rho Q$, and $m = \text{rank}(P - \rho_0 Q)$.

Proof: Deferred to the Appendix. ■

The decomposition provides a representation of $\mathbf{y}_t$ (corresponding to P) as a superposition of a minimal number of sinusoids on top of a suitably scaled background (corresponding to Q). The background typically represents a known process, e.g. white or colored noise, or a sinusoidal component at a known frequency. The theorem is a (slight) generalization of [11, Theorem 1] which is cast for $Q = E$. It is also generalization of the Carathéodory-Féjer/Pisarenko result discussed in Section II-A which corresponds to A, b as in (16).

We conclude the present section with a lemma from [11] which shows that "directions" in P excited by sinusoids at different frequencies are linearly independent.

Lemma 6: [11, Lemma 3] If $m < n$ and θ_k for $k = 0, 1, \ldots, m$, are pairwise distinct, then the columns of

$$\left[G(e^{j\theta_0}), G(e^{j\theta_1}), \ldots, G(e^{j\theta_m})\right]$$

are linearly independent.

V. Subspace Methods

In order to exploit Theorem 5 for the purpose of spectral-line estimation we need to assess a value for m as well as compute the θ_k's from an experimentally obtained P.

With exact knowledge of the state-covariance P and the covariance Q for the "background noise" signal, the decomposition in Theorem 5 becomes

$$P - \rho_0 Q = \mathbf{G}_s \text{diag}(\rho_1, \ldots, \rho_m)\mathbf{G}_s{}^*$$

where

$$\mathbf{G}_s := \left[G(e^{j\theta_1}), \ldots, G(e^{j\theta_m})\right]. \quad (26)$$

Consider the singular value decomposition

$$P - \rho_0 Q = U \Sigma U^*$$

with U a unitary matrix and $\Sigma = \text{diag}(\sigma_1, \ldots, \sigma_m, 0, \ldots, 0)$ the diagonal matrix of singular values. Partition U into the submatrices containing columns $1:m$ and $(m+1):n$, i.e.,

$$U = \left[\begin{array}{cc} U_{1:m} & U_{m+1:n} \end{array}\right]. \quad (27)$$

It follows that:

I. The range spaces (column spans) of $P - \rho_0 Q$, $\mathbf{G}_s$, and

$U_{1:m}$ coincide. This is the *signal subspace*.
II. The null spaces (right kernel) of $P-\rho_0 Q$, ${\bf G}_s{}^*$, and $U_{1:m}^*$ coincide and is spanned by the columns of $U_{m+1:n}$. This is the *noise subspace*.

When P is only an estimate of the state-covariance, a value for m can be assessed by identifying a "break" point in the singular values σ_k's. The next task is to identify the frequencies θ_k's and the corresponding residues ρ_k's. To this end we rewrite I and II as follows:

$$\begin{aligned} U_{1:m} &= {\bf G}_s M, \text{ and} && (28)\\ 0 &= {\bf G}_s{}^* U_{m+1:n}, && (29)\end{aligned}$$

respectively, where M is an invertible matrix. Condition (28) forms the basis of the *signal-subspace method*, whereas (29) the basis of a *noise-subspace method*.

A. Analysis based on the noise-subspace

Condition (29) in combination with Lemma 6 gives rise to the following fact.

Proposition 7: [11, Theorem 1(iv)] With $G(\lambda)$ and $U_{m+1:n}$ as above and θ_k as in Theorem 5, the function

$$d(\lambda,\lambda^{-1}) := G^*(\lambda^{-1}) U_{m+1:n} U_{m+1:n}^* G(\lambda) \tag{30}$$

has precisely m roots on the unit circle at $e^{j\theta_k}$ for $k = 1,\ldots,m$.

The proposition suggests generalizations of MUSIC as follows. Starting from an estimated sampled covariance P, and given information on the background noise, either assumed white or colored of a known signature, the trigonometric polynomial d can be estimated after we perform an SVD on $P-\rho_0 Q$. The frequencies θ_k for $k = 1,\ldots,m$ can be identified in a variety of ways. In particular, generalization of *Spectral* and *Root MUSIC* (cf. [25]) have been discussed in [11] for the case where the input noise is assumed white, i.e., when $Q = E$. Below, we first mention these two methods from [11], since they also apply to general Q, and then list a third possibility which appears more reliable in practice.

A. SPECTRAL MUSIC:
Identify θ_k's for $k = 1,\ldots,m$ as the values on $[-\pi,\pi]$ where $1/(d(e^{j\theta}, e^{-j\theta}))$ achieves the m-highest local maxima.

B. ROOT MUSIC:
Identify θ_k's with the angles of the m-roots of $d(\lambda,\lambda^{-1})$ which have amplitude < 1 and are closest to the unit circle.

C. RESIDUE MUSIC:
Identify θ_k's with the angles of the m roots of $d(\lambda,\lambda^{-1})$ which have the largest corresponding residues in the decomposition of main theorem. More specifically:

C-1: Select an ℓ in $m < \ell \le n-1$ allowing some margin beyond the known, or estimated, value m. Typically ℓ can be selected as either equal to $n-1$, or as a slightly conservative estimate of a "break" in the singular values of P. If ℓ is selected to be less than $n-1$ then a pre-selection of a set of candidate θ_k's for $k = 1,\ldots,\ell$ is necessary following methods (A) or (B) above.

C-2: Compute

$$V = \begin{bmatrix} G(e^{j\theta_1})^* \\ \ldots \\ G(e^{j\theta_\ell})^* \end{bmatrix} \cdot \left[G(e^{j\theta_1}),\ldots,G(e^{j\theta_\ell})\right] \text{ and} \tag{31}$$

$$\operatorname{diag}\{\rho_1,\ldots,\rho_\ell\} = V^{-1}\left(P-\rho_0 Q\right)\left(V^{-1}\right)^*$$

C-3: Select and re-order θ_k's for $k = 1,\ldots,m$ as the frequencies corresponding to the largest residues among $\rho_1,\ldots,\rho_\ell$.

We point out that traditional *Root MUSIC* as well as *Spectral MUSIC* (e.g., see [25]) generate estimates for the frequency of possible spectral lines. The generalization of either method to the new state-covariance provides similar information but not necessarily a power spectrum which is consistent with the data. *Residue MUSIC* on the other hand produces the parameters in the decomposition (25) and corresponds directly to a power spectrum which is consistent with the covariance data. This is of the form

$$d\sigma(\theta) = \rho_0 d\sigma_Q(\theta) + \sum_{k=1}^{m} 2\pi\rho_k \delta(\theta-\theta_k) d\theta.$$

where $d\sigma_Q(\theta)$ represents the assumed shape of power for the "background noise" which corresponds to a state-covariance Q.

B. Analysis based on the signal subspace

Condition (28) suggests an alternative way to identify the frequencies in the decomposition of Theorem 5. This is summarized in the following proposition.

Proposition 8: Under the notation and assumptions of Theorem 5 and $U_{1:m}$ defined as in (27), there is a unique solution to the following matrix equation

$$U_{1:m} = \begin{bmatrix} b & AU_{1:m} \end{bmatrix} \begin{bmatrix} \mu \\ \Phi \end{bmatrix},$$

where μ is a $1\times m$ row vector and Φ a $m\times m$ matrix. The eigenvalues of Φ are precisely $e^{j\theta_k}$ for $k = 1,\ldots,m$.

Proof: Since

$$\begin{aligned} G(\lambda) &= (I-\lambda A)^{-1} b \\ &= b + \lambda A (I-\lambda A)^{-1} b, \end{aligned}$$

it follows that

$${\bf G}_s = \begin{bmatrix} b & b & \ldots & b \end{bmatrix} + A\cdot{\bf G}_s\cdot\Theta, \tag{32}$$

with ${\bf G}_s$ as in (26) and

$$\Theta = \operatorname{diag}(e^{j\theta_1}, e^{j\theta_2},\ldots,e^{j\theta_m}).$$

Using (28) we have that

$$\begin{aligned} U_{1:m} &= b\begin{bmatrix} 1 & 1 & \ldots & 1 \end{bmatrix} M + A\cdot{\bf G}_s\Theta M \\ &= b\mu + A\cdot{\bf G}_s M M^{-1}\Theta M \\ &= b\mu + A\cdot U_{1:m}\Phi \\ &= \begin{bmatrix} b & AU_{1:m} \end{bmatrix} \begin{bmatrix} \mu \\ \Phi \end{bmatrix}, \end{aligned} \tag{33}$$

where
$$\mu := \begin{bmatrix} 1 & 1 & \ldots & 1 \end{bmatrix} M$$
and
$$\Phi := M^{-1}\Theta M.$$
Equation (33) is linear with $m \times n$ constraints and $m \times (m+1)$ entries in the variables μ, Φ. A solution exists by virtue of the above analysis.

Further, the solution is unique. To see this we need to prove that the (right) null space of $\begin{bmatrix} b & AU_{1:m} \end{bmatrix}$ is trivial, i.e., it contains only the zero vector. First note that this null space is trivial if and only if the null space of $\begin{bmatrix} b & A\mathbf{G}_s \end{bmatrix}$ is trivial. By simple algebra it can be shown that this is equivalent to the null space of $\begin{bmatrix} b & \mathbf{G}_s \end{bmatrix}$ being trivial as well. Thus, in order to prove our claim we need to prove that the columns of $\begin{bmatrix} b & \mathbf{G}_s \end{bmatrix}$ are linearly independent. It suffices to show that b is linearly independent of the columns of $\mathbf{G}_s$ when $m = n-1$. Thus, assuming $m = n-1$, we consider the square matrix
$$M(\lambda) = [G(\lambda),\ \mathbf{G}_s].$$
Since the numerator of
$$\det M(\lambda) = c(I - \lambda A)^{-1} b,$$
is of degree $n-1$, it cannot have any root other than $\{e^{j\theta_1}, \ldots, e^{j\theta_m}\}$ (which are obviously roots). Hence $M(0) = [G(0),\ \mathbf{G}_s]$ is nonsingular, and since $b = G(0)$, this proves our claim. Therefore (33) has a unique solution.

Finally, the sought roots $e^{j\theta_k}$ $(k = 1, \ldots, m)$ are eigenvalues of Φ because Φ relates to Θ via a similarity transformation. ∎

In practice, when m is estimated from a break point in the singular values of $P - \rho_0 Q$, better accuracy is often obtained if we begin with a conservative estimate ℓ for the number of frequencies to be considered, with $n-1 \geq \ell > m$, and then identify m of them with the largest residue, as we did in the method (C) in Section V-A.

VI. Absorption Spectra

The construction in Section II-C carries over to the more general framework of IS-filters and their state-covariances.

Theorem 9: Let P be a state-covariance of the IS-filter (14), and W as in Theorem 1. Define
$$\hat{W} := W^{-1},$$
and
$$\hat{P} := \frac{1}{2}(\hat{W}E + E\hat{W}^*).$$
The following hold:
(i) if $\hat{f}$ is positive real and $\hat{f}(A^*) = \hat{W}^*$, then
$$f(\lambda) := \hat{f}(\lambda)^{-1}$$
is positive real and $f(A^*) = W$.
(ii) $\hat{P}$ is a state-covariance of the IS-filter (14).

Proof: As shown in Section II-C, $\hat{f}$ is positive real if and only if $f := \hat{f}^{-1}$ is positive real. Since $f(\lambda)\hat{f}(\lambda) = 1$ it follows that $f(A^*)\hat{f}(A^*) = I$. This proves (i).

Since $\hat{P}$ satisfies
$$\begin{aligned} \hat{P} &= \frac{1}{2}(\hat{W}E + E\hat{W}^*) \\ &= \frac{1}{2}(W^{-1}E + E(W^*)^{-1}) \qquad (34) \\ &= W^{-1}\frac{1}{2}(EW^* + WE)(W^*)^{-1} \\ &= W^{-1}P(W^*)^{-1} \geq 0, \qquad (35) \end{aligned}$$
as well as, it is of the form required by Theorem 1, it follows that it is a state-covariance of the IS-filter. Note that W^{-1} commutes with A since W does. This shows (ii). ∎

Equation (35) is important due to the following practical consideration. If P is obtained from experimental data as a sample covariance and W is computed from it via least squares and (23) (as explained in Section III), then estimating $\hat{P}$ from the right hand side of (34) may not produce a positive definite quantity, whereas (35) always does!

The above proposition can be used to produce spectra from a sample covariance matrix P. When such spectra are selected, based on an $\hat{f}$ which contains spectral lines (e.g., using the earlier canonical decomposition), then the spectra corresponding to $f := \hat{f}^{-1}$ will have absorption lines at the frequencies of the spectral lines of $\hat{f}$. We summarize below an option based on the earlier canonical decomposition.

Canonical absorption spectra: Begin with a state-covariance P. Determine W as in Theorem 1 and $\hat{P}$ as in Theorem 9. Determine ρ_k and θ_k as in Theorem 5, when the theorem is applied to $\hat{P}$ instead of P. The computation of ρ_k, θ_k can be done using, e.g., residue MUSIC as described in Section V with $Q = E$ as in Theorem 1. Then θ_k will correspond to absorption lines for a power spectrum consistent with the covariance data P.

To verify the last statement, note that the positive real function
$$\hat{f}(\lambda) = \rho_0 + \sum_{k=1}^{m} \rho_k \frac{1 + e^{j\theta_k}}{1 - e^{j\theta_k}}$$
has poles at these frequencies. Hence $f(\lambda) = \hat{f}(\lambda)^{-1}$ will have zeros at the same points, and so will its real part $\Re\{f(e^{j\theta})\}$. Of course these will be shared by the relevant spectrum which is defined from the radial limits of the $\Re\{f(\lambda)\}$.

VII. Spectral Envelopes

In this section we present a generalization of the methods in Section II-D to state-covariance data. Hence, we seek to compute the envelope of maximal spectral power
$$\rho(\theta) := \sup\left\{\lim_{\epsilon \to 0} (\sigma(\theta + \epsilon) - \theta(\theta - \epsilon)) \ :\ d\sigma \in \Sigma_P\right\},$$
where
$$\Sigma_P := \{d\sigma \text{ nonnegative measure satisfying (17)}\}.$$

As in Section II-D we also define $\rho_{\rm real}(\theta)$ where $\sigma(\theta)$'s are further restricted to have an odd symmetry about $\theta = 0$ so that they correspond to a real time-series $\mathbf{y}_k$ (if that is known to be the case). Below, the notation $\Im$ applied to products $G(e^{j\theta})G(e^{j\theta})$, follows our earlier convention (13).

Proposition 10: Let P be a state-covariance of the IS-filter (14), then

$$\begin{aligned}\rho(\theta) &= \max\left\{\rho \;:\; P - \rho G(e^{j\theta})G(e^{j\theta})^* \geq 0\right\}\\ \rho_{\rm real}(\theta) &= \max\left\{\rho \;:\; P - \rho\Im\left(G(e^{j\theta})G(e^{j\theta})^*\right) \geq 0\right\}\end{aligned}$$

Proof: First note that

$$\begin{aligned}Q_\theta &:= G(e^{j\theta})G(e^{j\theta})^*, \text{ and} &(36)\\ Q_{\theta,\rm real} &:= \Im\left(G(e^{j\theta})G(e^{j\theta})^*\right), &(37)\end{aligned}$$

respectively, represent the state-covariance of the (complex and real, resp.) sinusoidal signals $e^{jk\theta}$, $\cos(k\theta)$.

Corollary 3 implies that, for any given θ_0, $\Delta := P - \rho(\theta_0)Q_{\theta_0}$ is an admissible state-covariance. Hence, the decomposition $P = \Delta + \rho(\theta_0)Q_{\theta_0}$ allows a power spectrum of the form $d\sigma(\theta) = d\sigma_\Delta(\theta) + \rho(\theta_0)\delta(\theta-\theta_0)d\theta$. Therefore, a spectral line at $\theta = \theta_0$ of amplitude $\rho(\theta_0)$ is consistent with P. It is easy to see that $\rho(\theta_0)$ is the maximal such amplitude consistent with the data. Analogous argument proves the case of $Q_{\theta,\rm real}$. ■

Evidently, the ρ's can be obtained by computing the smallest eigenvalue of the corresponding matrix pencils $P - \rho Q_\theta$ and $P - \rho Q_{\theta,\rm real}$, respectively, in complete analogy with Section II-D. In particular, since $\mathrm{rank}(Q_\theta) = 1$ we have

$$\rho(\theta) = \frac{1}{G(e^{j\theta})^* P^{-1} G(e^{j\theta})} \qquad (38)$$

while

$$\rho_{\rm real}(\theta) = \min\left\{\rho \;:\; \det\left(P - \rho\Im\left(G(e^{j\theta})G(e^{j\theta})^*\right)\right) = 0\right\}.$$

VIII. Input-to-state filter characteristics

Given state-covariance data P, Q, and IS-filter parameters $G(\lambda) = (I - \lambda A)^{-1}b$, then the m, ρ's, and θ's in the decomposition of Theorem 5 remain the same when the data P, Q are replaced with $P_o = MPM^*$, $Q_o = MQM^*$ and the filter parameters by $G_o(\lambda) = (I - \lambda A_o)^{-1}b_o$, with $A_o = MAM^{-1}$, $b_o = Mb$. Here, M is any invertible matrix. To see this, simply note that if (25) holds, then

$$P_o = \rho_0 Q_o + \sum_{k=1}^{m} \rho_k MG(e^{j\theta_k})G(e^{j\theta_k})^* M^*.$$

Hence, it is convenient to normalize the filter parameters via similarity transformation so that $Q = I$, and this is what will be followed in here. The frequency characteristics of IS-filters, for the purposes of the decomposition in Theorem 5 can be quantified by the frequency function

$$s(\theta) := \|(I - e^{j\theta}A_o)^{-1}b_o\| = \|Q^{-\frac{1}{2}}G(e^{j\theta})\|. \qquad (39)$$

In case no information on the "color" of the input noise is available we can use $Q = E$, E being the state-covariance of white noise input (i.e., a controllability grammian) which satisfies $E = bb^* + AEA^*$. Such a normalization was used in [11]. Guidelines to obtain suitable characteristics are summarized below:

If A is taken in a Jordan form

$$A = \begin{bmatrix} r & 1 & 0 & \dots & 0\\ 0 & r & 1 & \dots & 0\\ \vdots & \vdots & \vdots & & \vdots\\ 0 & 0 & 0 & \dots & 1\\ 0 & 0 & 0 & \dots & r\end{bmatrix}, \qquad (40)$$

$b = [0, \dots, 0, 1]'$ and r in the interval $0 < r < 1$ (resp. $-1 < r < 0$), then a low-pass (resp. high-pass) characteristic $s(\theta)$ is obtained. Similarly, if r is complex, the pass-band of $s(\theta)$ is centered around the phase of r. To obtain a bandpass characteristic with a real filter, A can be chosen in block-Jordan form

$$A = \begin{bmatrix} R & I & 0 & \dots & 0\\ 0 & R & I & \dots & 0\\ \vdots & \vdots & \vdots & & \vdots\\ 0 & 0 & 0 & \dots & I\\ 0 & 0 & 0 & \dots & R\end{bmatrix} \qquad (41)$$

with

$$R = \begin{bmatrix} r\cos(\theta_0) & r\sin\theta_0\\ -r\sin(\theta_0) & r\cos\theta_0\end{bmatrix},\; I = \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix},$$

$0 < r < 1$ and $0 < \theta_0 < \pi$. In this case, again the choice $b = [0, \dots, 0, 1]'$ can be used.

Other characteristics might also be useful in quantifying the resolution of IS-filters, e.g., the sensitivity $\|\frac{d}{d\theta}G(e^{j\theta})G(e^{j\theta})^*\|$ of the state-covariance on the frequency of sinusoidal inputs.

IX. Case studies

We first present two simulation studies where the goal is to resolve two closely spaced sinusoids. We compare traditional MUSIC and ESPRIT with the subspace methods of Section V. Next we present an example of a typical absorption spectrum for which the approach in Section VI is appropriate. Finally we compare the Capon method and the method in [23] (both outlined in Section II-D), with those presented in VII. Invariably, spectral estimation based on suitably selected IS-filters, shows remarkably better resolution.

In each example, the IS-filter parameters are chosen following the guidelines of the previous section. The time series

$$\mathcal{Y} := \{y_0, y_1, y_2, \dots, y_{N-1}\}$$

is used to generate the state process x_k via $x_k = Ax_{k-1} + by_k$, for $k = 0, \dots, N-1$, with $x_{-1} = 0$. The time-series

$$\{x_\ell, \dots, x_{N-1}\}$$

is used to obtain the sample state-covariance matrix

$$P = \frac{1}{N}\sum_{t=\ell}^{N-1} x_t x_t^*.$$

Here, we take $\ell = 0$. However, in general, if the time-constant of the IS-filter is significant, then $\ell > 0$ can be chosen to ensure a level of stationarity for x_k.

A. Subspace methods

Case 1: Closely spaced sinusoids

We generated time-series data according to

$$\mathbf{y}_k = \nu_k + \frac{1}{2}\sin(\theta_1 k + \phi_1) + \sin(\theta_2 k + \phi_2),$$

with $\theta_1 = 0.1$, $\theta_2 = 0.15$, ν_k a Gaussian zero-mean white noise with variance ρ_0, and ϕ_1, ϕ_2 uniformly distributed on $[0, 2\pi]$. We compared the following 6 different methods: (i) standard Root MUSIC and (ii) ESPRIT, (iii) standard root MUSIC after a pre-filtering of the data with a lowpass scalar filter, (iv) ditto for ESPRIT, and finally, (v) the noise-subspace method (residue MUSIC) of Section V-A and (vi) the signal-subspace method of Section V-B.

We tested whether the methods are capable of resolving the two sinusoids based on $N = 100$ data points, i.e., when the frequencies θ_1 and θ_2 are closer than the theoretical resolution limit $2\pi/N$ of periodogram-based methods. We report on 1000 simulation runs. In each, we considered the outcome successful if the method identified one, and only one spectral line, in each of the two intervals

$$(\theta_1 - 2\delta, \theta_1 + \delta) \text{ and } [\theta_2 - \delta, \theta_2 + 2\delta)$$

where $\delta = (\theta_1 - \theta_2)/2$ (and hence $\theta_1 + \delta = \theta_2 - \delta$ is the mid point). For values of noise variance ρ_0 above 0.35, methods (i-iv) almost always fail. Thus, we used $\rho_0 = 0.35$.

For (v) and (vi) we used a low-pass IS filter of order 30 as in (40) with $r = 0.3$ For (iii) and (iv) we used a standard Butterworth filter of order 10 and -3dB frequency at 0.5. Encoding for standard MUSIC and ESPRIT was taken from our favorite reference [25] with parameters `music(y,4,30); esprit(y,4,30)`.

Below we tabulate the number of successful hits and, in Figure 1, we display the scatter diagram of values for θ_1 vs. θ_2 identified by each method.

Method	success / 1000 runs
(i) Root MUSIC	4
(ii) ESPRIT	36
(iii) prefiltered Root MUSIC	15
(iv) prefiltered ESPRIT	160
(v) Section V-A	527
(vi) Section V-B	963

Methods (iii-iv) were suggested for the purpose of a "fair comparison" by an anonymous referee. The fact, that methods (iii-iv) are only marginaly better than (i-ii), suggests that filtering alone is not the main factor behind the higher resolution of methods (v) and (vi).

Case 2: Variance of frequency estimates

This is similar to Case 1 except for the fact that the time-series has only one sinusoidal component

$$\mathbf{y}_k = \nu_k + \sin(\theta_1 k + \phi_1)$$

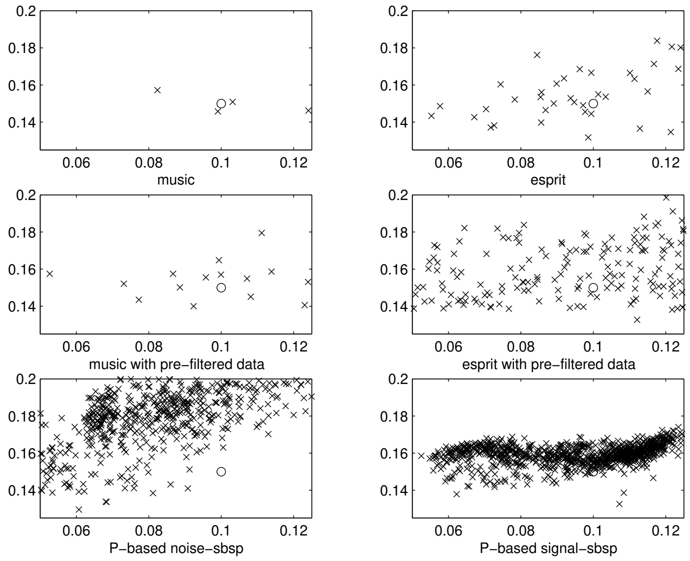


Fig. 1. Scatter diagram

with $\theta_1 = 0.1$, ϕ_1 uniformly distributed on $[0, 2\pi]$, and ν_k Gaussian white-noise with variance $\rho_0 = 1$. A single sinusoid was chosen so as to compare the variance of estimates on an equal footing, because in this case, we do not differentiate failures as in Case 1.

The results we report are based on 1,000 runs, each time generating a 100 samples of the process. We used a low-pass IS filter as in (40) with $r = 0.3$ and dimension 20. The figures we obtained when comparing standard ESPRIT (coded as in [25]; using `esprit(y,2,20)`), with the signal subspace method of Section V-B are as follows:

$\theta_1 = 0.1$	ESPRIT	Section V-B
Mean estimate of θ_1	0.09728	0.09957
Standard deviation	0.09992	0.00681

We note that the standard deviation of the estimate using Section V-B shows a dramatic improvement when compared to the estimate via standard ESPRIT.

B. Absorption spectra

In current literature there are no standard methods for the detection of the "absence of spectral energy". Indeed, the identification of absorption lines is not as reliable as the detection of spectral lines. With this caveat we present one example of the type of spectra that may be suitable for applying the approach.

We consider a notch-filter with transfer function

$$g(\lambda) = \frac{1 - 1.9754\lambda + \lambda^2}{3 - 2 \cdot 1.9754\lambda - \lambda^2}$$

with absorption frequency at $\frac{\pi}{20} = 0.1572$. Unit-variance, zero-mean, Gaussian white noise is used to drive the notch filter and a 1,000 data points of the output time-series $\mathbf{y}_k$ were used. The time-series is shown along with the corresponding periodogram in Figure 2. It is quite evident that, although the power at low frequencies is small, any

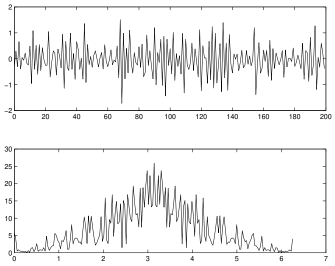


Fig. 2. Time-series and FFT of output of the notch filter

attempt to gauge the location of an absorption line would be futile.

For the purposes of applying Section VI a lowpass was selected as in (41) having 3 blocks R with eigenvalue $.5+.1j$. Following the steps of Section VI, a frequency $\theta_1 = 0.1149$ was identified.

C. Spectral envelopes

We conclude by comparing the methods of Section VII with the Capon method of maximum likelihood, e.g., as encoded in [25, page 196], and with the one in [23]. These methods generate spectral envelopes as explained in Sections II-D and VII.

Once again we generated a time-series of 300 data points according to

$$\mathbf{y}_k = \nu_k + 0.5\sin(0.1k + \phi_1) + \sin(0.5k + \phi_2)$$

with ν_k Gaussian, zero-mean, white noise and ϕ_1, ϕ_2 in $[0, 2\pi]$ (independently distributed as usual). The

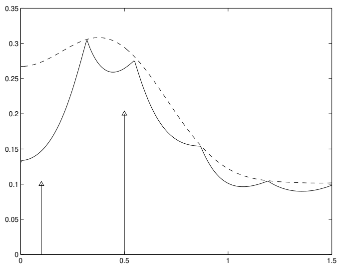


Fig. 3. Capon (dashed) and ref. [23] (solid) envelopes

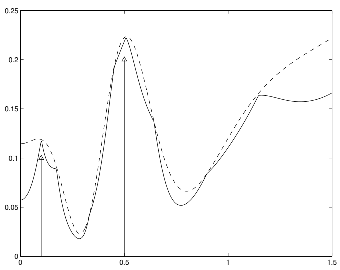


Fig. 4. "Complex" (dashed), "real" (solid) envelopes via Sect. VII

Capon and Shankwitz-Georgiou [23] methods used a sample Toeplitz matrix of size 10×10, and accordingly, for the methods of Section VII we estimated a 10×10 state-covariance matrix of an IS-filter designed as in (41) with 5 jordan blocks R of size 2×2 with eigenvalues $0.6 \pm 0.2j$. This is low pass with a cutoff of about 1.

Figure 3 shows with dashed line the Capon estimate and with solid line the estimate according to [23], as explained in Section II-D. On the same figure vertical arrows are used to indicate the frequency and, scaled by a factor of 0.2, the amplitude of the two sinusoids.

Similarly Figure 4 shows with dashed and solid lines the spectral envelopes computed using the theory of Section VII. The dashed line indicates $\rho(\theta)$ computed according to (38), while the solid line represents $\rho_{\rm real}(\theta)$ which takes into account the fact that the signal and sinusoids are real. It is apparent that over the passband of the IS filter, the resolution is significantly increased as compared with the earlier methods. It is instructive to observe that improvement in resolution over a frequency band is traded off with a degradation elsewhere.

X. Concluding comments

This work was motivated by Byrnes, Georgiou and Lindquist [4,5], where it was noted that output covariances of first-order filters provide Nevanlinna-Pick interpolation conditions for the input power spectrum. It also builds on [11] which explored the use of state-covariances in a generalization of MUSIC. However, the scope of the present work has been to expose an underlying theory for spectral analysis via Theorems 1, 2, and Corollary 3, and thereby, present a state-covariance framework for nonlinear spectral estimation techniques. It is interesting to speculate whether an analogous viewpoint of characterizing covariances for non-uniform arrays of varying spacial geometries and of spectra which are consistent with such statistics, will be beneficial in sensor array processing.

XI. Appendix: proofs

We give a brief exposition of needed concepts and facts from generalized interpolation. These allow a rather compact derivation of the results presented earlier in the paper. We restrict our attention to a finite-dimensional framework since this is all we need, and accordingly, we follow our earlier notation regarding $G(\lambda), A, b, n$, etc. Also, "$*$" denotes the "complex conjugate transpose" or the "adjoint operator" depending on the context, as is customary. When we want to emphasize "complex conjugate" of scalar quantities we occasionally use "$^{-}$".

Let H_2 denote the usual Hardy space of functions which are analytic in the unit disc with square-integrable radial limits, and define

$$\mathcal{K} := H_2 \ominus B(\lambda)H_2,$$

where $B(\lambda) = \det(\lambda I - A^*)/\det(I - \lambda A)$ as in (22). I.e., $\mathcal{K}$ contains all functions in H_2 which are orthogonal to those that vanish on the spectrum of A^*. (See [3,17,21] for the role of $\mathcal{K}$ in operator theory; cf. [9,14,16] for its use in systems theory.)

With $G(z)$ as in (15), its entries $g_i(\lambda)$ form a basis for $\mathcal{K}$. Indeed, these are generalized Cauchy kernels whose inner product "evaluates" at the spectrum of A^*; e.g., if $f(\lambda) \in H_2$ then

$$\begin{aligned}\langle f(\lambda), g_i(\lambda)\rangle &:= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(e^{j\theta})g_i(e^{j\theta})^* d\theta \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(e^{j\theta})\overline{e_i(I-e^{j\theta}A)^{-1}b} d\theta \\ &= b^*\left(\frac{1}{2\pi}\int_{-\pi}^{\pi} f(e^{j\theta})(I-e^{-j\theta}A^*)^{-1}d\theta\right)e_i^* \\ &= b^* f(A^*)e_i^*,\end{aligned}$$

where

$$e_i := [0,\ldots,0,1,0,\ldots,0] \text{ with a 1 in the } i\text{-th spot.}$$

Consider the ordinary right shift $x(\lambda) \mapsto \lambda x(\lambda)$ in H_2, and denote by S the *compressed shift* onto $\mathcal{K}$:

$$S : \mathcal{K} \to \mathcal{K} : x(\lambda) \mapsto \mathbf{\Pi}_{\mathcal{K}}\lambda x(\lambda),$$

and by S^* the adjoint operator. It is instructive to express both S and S^* with respect to the basis $G(\lambda)$. Starting with S^*, this maps the general element $c \cdot G(\lambda)$ of $\mathcal{K}$ into

$$\begin{aligned}\mathbf{\Pi}_{H_2}\lambda^{-1}c(I-\lambda A)^{-1}b &= \mathbf{\Pi}_{H_2}\lambda^{-1}cb + cAb + \lambda cA^2b + \ldots \\ &= \mathbf{\Pi}_{H_2}cAb + \lambda cA^2 b + \ldots \\ &= cA\cdot(I-\lambda A)^{-1}b.\end{aligned}$$

Turning to S we first need the Grammian

$$E = \langle G(\lambda), G(\lambda)\rangle = \int_{-\pi}^{\pi}\left(G(e^{j\theta})\frac{d\theta}{2\pi}G(e^{j\theta})^*\right),$$

which coincides with the solution of $E = bb^* + AEA^*$ and noted earlier on. If M is the sought matrix representation for S, then

$$\langle c_1 G(\lambda), c_2 AG(\lambda)\rangle = \langle c_1 MG(\lambda), c_2 G(\lambda)\rangle$$

must hold for all row n-vectors $c_1, c_2 \in \mathbb{C}^n$. Hence $c_1 EAc_2^* = c_1 MEc_2^*$ and therefore $M = EA^*E^{-1}$. We summarize for future reference:

$$\begin{aligned} S^* :&\quad \mathcal{K}\to\mathcal{K} : \quad c\cdot G(\lambda) \mapsto cA\cdot G(\lambda), && (42)\\ S :&\quad \mathcal{K}\to\mathcal{K} : \quad c\cdot G(\lambda)\mapsto cEA^*E^{-1}\cdot G(\lambda). && (43)\end{aligned}$$

A key result which will be repeatedly needed in the sequel is the celebrated Commutant Lifting Theorem [21, Theorem 1]. A special version which is sufficient for our purposes is summarized below (cf. [11, Lemma 4], [26], [7,20]).

Lemma 11: . If T is an operator on $\mathcal{K}$ that commutes with S and the real part of T is positive semidefinite, then there exists a function $f(\lambda)$ which is analytic in the open disc and has positive real part there, such that $T = f(S)$. If the real part of T is singular and $\mathcal{K}$ finite-dimensional, then the interpolant f is unique.

A. Proofs of Theorem 1 and Theorem 2

We recall for convenience the notation for the set of power spectra consistent with a given state-covariance P of (14):

$$\Sigma_P = \{d\sigma(\theta)\geq 0 \,:\, P = \int_{-\pi}^{\pi}(G(e^{j\theta})\frac{d\sigma(\theta)}{2\pi}G(e^{j\theta})^*)\}.$$

We first assume that P is a state-covariance of (14) and show that it is of the form claimed in Theorem (1): Since P is a state-covariance, Σ_P is not empty. So we consider $d\sigma \in \Sigma_P$ and the corresponding "positive-real half spectrum"

$$f(\lambda) = \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{1+\lambda e^{-j\theta}}{1-\lambda e^{-j\theta}}d\sigma(\theta).$$

It follows that

$$\begin{aligned} P &= \frac{1}{2\pi}\int_{-\pi}^{\pi}(G(e^{j\theta})d\sigma(\theta)G(e^{j\theta})^*) \\ &= \lim_{r\nearrow 1}\langle G(\lambda), \Re(f(r\lambda))G(\lambda)\rangle \qquad (44)\\ &= \lim_{r\nearrow 1}\frac{1}{2}\left(\langle G(\lambda), f(r\lambda)G(\lambda)\rangle + \langle G(\lambda), \overline{f(r\lambda)}G(\lambda)\rangle\right) \\ &= \lim_{r\nearrow 1}\frac{1}{2}\left(\langle \overline{f(r\lambda)}G(\lambda), G(\lambda)\rangle + \langle G(\lambda), \overline{f(r\lambda)}G(\lambda)\rangle\right) \\ &= \frac{1}{2}\left(\langle \bar f(\lambda^{-1})G(\lambda), G(\lambda)\rangle + \langle G(\lambda), \bar f(\lambda^{-1})G(\lambda)\rangle\right) \\ &= \frac{1}{2}\left(\langle WG(\lambda), G(\lambda)\rangle + \langle G(\lambda), WG(\lambda)\rangle\right) \\ &= \frac{1}{2}\left(WE+EW^*\right),\end{aligned}$$

where W is a representation of the mapping

$$x(\lambda)\to \mathbf{\Pi}_{\mathcal{K}}\bar f(\lambda^{-1})x(\lambda)$$

with respect to the basis $G(\lambda)$ for $\mathcal{K}$. Despite the fact that $\bar f(\lambda^{-1})$ may not be bounded, the above mapping exists and can be identified e.g., with the weak limit of $\mathbf{\Pi}_{\mathcal{K}}(\bar f_0 + \bar f_1\lambda^{-1}+\ldots)$ since $\mathcal{K}$ is finite-dimensional. Note that here and throughout, $\bar f(\lambda)$ denotes conjugating only the coefficients of $f(\lambda)$, as opposed to $\overline{f(\lambda)}$ which would include the argument. Recall that A is a matrix representation of

$$S^* : x(\lambda)\to \mathbf{\Pi}_{\mathcal{K}}\lambda^{-1}x(\lambda).$$

Hence, $W = \bar f(A)$ and therefore commutes with A.

The same arguments verify the first part of Theorem (2). Namely, if P, W are as in Theorem 1 and $d\sigma_y(\theta)$ satisfies (17), then $f_y(\lambda)$ obtained from (3) satisfies the noted interpolation condition $\bar f(A) = W$.

We now show the converse to Theorem 1, namely that, if a nonnegative matrix P is of the form $P = WE + EW^*$ for a W which commutes with A, then it is a state-covariance of (14): Since W commutes with A, it can be expressed as a function of A, and in particular, a polynomial function

$$W := w(A) = w_0 I + w_1 A + \ldots + w_{n-1}A^{n-1}$$

with $w(\lambda) = w_0 + \ldots + w_{n-1}\lambda^{n-1}$ of degree $n-1$. Via $cG(\lambda) \mapsto cWG(\lambda)$, the matrix W defines the operator $T^* = w(S^*)$ which commutes with S^*. According to (42-43) the real part of T is

$$\begin{aligned}\frac{1}{2}\langle (T+T^*)G, G\rangle &= \frac{1}{2}(E\bar{w}(A^*) + w(A)E)\\ &= \frac{1}{2}(EW^* + WE) \;=\; P.\end{aligned}$$

Hence the real part of T is nonnegative. Therefore, according to Lemma 11, T can be selected in the form $T = f(S)$ with $f(\lambda)$ a positive real function. It is now easy to see that $d\sigma(\theta)$ defined as in (5) via radial limit of $\Re\{f(\lambda)\}$ serves as the required power spectrum yielding P.

We now argue the second part of Theorem 2, namely that if $f_y(\lambda)$ is positive real which satisfies $f_y(A^*) = W^*$, then it defines a power spectrum consistent with P. Clearly a nonnegative $d\sigma_y(\theta)$ defined via radial limit of $\Re\{f_y(\lambda)\}$ as above satisfies the equations following (44) on, until $\frac{1}{2}(EW^* + WE)$, which is precisely P by assumption.

We finally show that any such W as in Theorem 1 is unique modulo an imaginary constant αI with $\alpha \in j\mathbb{R}$. Begin with equation (23),

$$\begin{aligned}2P &= w(A)E + E\bar{w}(A^*)\\ &= w_{n-1}A^{n-1}E + \ldots + w_1 AE + (w_0 + \overline{w}_0)I\\ &\quad + \bar{w}_1 E(A^*) + \ldots + \bar{w}_{n-1}E(A^*)^{n-1}.\end{aligned}$$

Since P is a state-covariance and, by the earlier arguments, of the form $\frac{1}{2}(EW^* + WE)$, with W commuting with A, we can certainly determine a polynomial $w(\lambda)$ for which $w(A) = W$ and therefore the above equation has a solution. Clearly, the imaginary part of w_0 cannot be determined from this equation. Other than that we need to show that the values for $\Re\{w_0\}, w_1, \ldots, w_{n-1}$ are uniquely defined.

It suffices to show that if

$$w(A)E + E\bar{w}(A^*) = 0, \tag{45}$$

then $w(\lambda)$ is identically equal to an imaginary constant. Setting $T^* = w(S^*)$ with $T = \bar{w}(S)$ as usual, the real part of T is singular and therefore the positive real f claimed by Lemma (11) is unique. If f is this unique interpolant of Lemma 11 and $d\sigma$ the corresponding nonnegative measure, since $P = 0$, both f and $d\sigma$ are obviously zero. If now $w(\lambda)$ be any polynomial satisfying (45), assumed real at $\lambda = 0$ to avoid the obvious indeterminacy of w_0 by an additive imaginary constant, then

$$w(A) = \bar{f}(A) = 0.$$

In order for $w(A) = 0$, $w(\lambda)$ needs to have the minimal polynomial of A as a factor. Since (A, b) is a controllable pair, A is cyclic, and the minimal polynomial has degree n. But the degree of w is $n-1$, therefore $w(\lambda)$ is the zero polynomial. This completes the proof of uniqueness.

B. Proof of Theorem 5

If ρ_0 is the smallest eigenvalue of the pencil $P - \rho Q$, then according to Corollary 3, $P - \rho_0 Q$ is a state-covariance. Lemma 4 then shows that it is of the form required in (25) and that $m < n$. We now show that this m is minimal.

Assume an alternative decomposition

$$P = \rho_0' Q + \sum_{k=1}^{m'} \rho_k' G(e^{j\theta_k'})G(e^{j\theta_k'})^*.$$

Clearly, $\rho_0 \geq \rho_0'$, otherwise $P - \rho_0' Q \not\geq 0$. Then

$$(\rho_0' - \rho_0)Q + \sum_{k=1}^{m'} \rho_k' G(e^{j\theta_k'})G(e^{j\theta_k'})^* = \sum_{k=1}^{m'} \rho_k' G(e^{j\theta_k'})G(e^{j\theta_k'})^*,$$

where both sides represent state-covariances. If $\rho_0' = \rho_0$ then the two decompositions are identical by virtue of Lemma 6. When $\rho_0' > \rho_0$ then we distinguish two cases. First, $\mathrm{rank}(Q) = n$: then $m' \geq n$ since the left hand side has full rank. Second, $\mathrm{rank}(Q) < n$: then by Proposition 4 $(\rho_0' - \rho_0)Q = \sum_{k=1}^{m^\dagger} \rho_k^\dagger G(e^{j\theta_k^\dagger})G(e^{j\theta_k^\dagger})^*$. But $\theta_k^\dagger \notin \{\theta_1, \ldots, \theta_m\}$, for otherwise ρ_0 will not be the smallest eigenvalue of the pencil. Hence, again by virtue of Lemma 6, $m' > m$.

References

[1] N. I. Akhiezer, **The classical moment problem,** Hafner Publishing, Translation 1965, Oliver and Boyd.

[2] N. I. Akhiezer and M. G. Krein, **Some Questions in the Theory of Moments,** American Mathematical Society, Providence, RI, 265 pages, 1962.

[3] J. A. Ball and N. Cohen, De Branges-Rovnyak operator models and systems theory: A survey, preprint.

[4] C. I. Byrnes, T.T. Georgiou, and A. Lindquist, A generalized entropy criterion for rational Nevanlinna-Pick interpolation: A convex optimization approach to certain problems in systems and control, *IEEE Trans. on Aut. Contr.*, to appear.

[5] C. I. Byrnes, T.T. Georgiou, and A. Lindquist, A new approach to spectral estimation: A tunable high-resolution spectral estimator, *IEEE Trans. on Signal Proc.*, to appear.

[6] C. Carathéodory and L. Fejér, Über den Zusammenghang der Extemen von harmonischen Funktionen mit ihren Koeffizienten und über den Picard-Landauschen Satz, Rendiconti del Circolo Matematico di Palermo, 32: 218-39, 1911.

[7] C. Foias and A.E. Frazho, **The commutant lifting approach to interpolation problems**, **OT 44**, Birkhäuser Verlag, Basel, 1990.

[8] B.A. Francis, **A course in H_∞ control theory**, Springer-Verlag, 1987.

[9] P.A. Fuhrmann, Algebraic system theory: An analyst's point of view, *J. Franklin Inst.* **301**, pp. 521-540, 1976.

[10] Ya. L. Geronimus, **Orthogonal Polynomials**, English translation from Russian by Consultants Bureau, New York, 570 pages, 1961.

[11] T.T. Georgiou, Signal Estimation via Selective Harmonic Amplification: MUSIC, Redux, *IEEE Trans. on Signal Processing*, March 2000, **48(3)**: 780-790.

[12] U. Grenander and G. Szegö, **Toeplitz Forms and their Applications**, Chelsea, 1958.

[13] S. Haykin, **Nonlinear Methods of Spectral Analysis,** Springer-Verlag, New York, 247 pages, 1979.

[14] P.P. Khargonekar, T.T. Georgiou, and A.B. Özgüler, Skew prime polynomial matrices: the polynomial model approach, *Linear Algebra and Applications*, **50:** 403-435, 1983.

[15] M.G. Krein and A.A. Nudel'man, **The Markov Moment Problem and Extremal Problems,** American Mathematical Society, Providence, RI, 417 pages, 1977.

[16] N. Levan, *The Nagy-Foias operator models, networks and systems,* IEEE Transactions on Circuits and Systems, vol. CAS-23(6), 1976.

[17] N.K. Nikol'skiĭ and V.I. Vasyunin, Notes on two function models, in **The Bieberbach conjecture: proceedings of the symposium on the occasion of the proof**, A. Bernstein II *etal.*,

(eds.) Mathematical Surveys and Monographs, No. 21, American Mathematical Society, Providence, RI, pp. 113-141, 1986.

[18] V.F. Pisarenko, The retrieval of harmonics from a covariance function, Geophys. J. Roy. Astron. Soc., 33: 347-66, 1973.

[19] A. Paulraj, R. Roy and T. Kailath, A subspace rotation approach to signal parameter estimation, *Proc. of the IEEE*, **74**: 1044:5, 1986.

[20] M. Rosenblum and J. Rovnyak, **Hardy Classes and Operator Theory**, Oxford University Press, 1985.

[21] D. Sarason, Generalized Interpolation in $\mathcal{H}^\infty$, *Transactions of the American Math. Society*, **127:** 179-203, 1967.

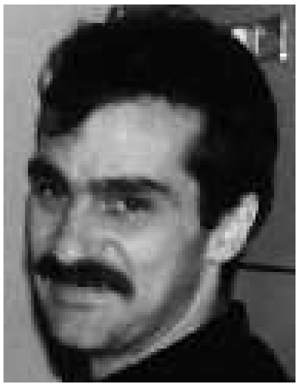
[22] R. Schmidt, Multiple emitter location and signal parameter estimation, Proc. RADC Spectrum Estimation Workshop, pp. 243-58, 1979.

[23] C. Shankwitz and T.T. Georgiou, On the envelope of spectral power of stochastic processes, *IEEE Trans. on Information Theory*, **40(2)**: 584-588, 1994.

[24] J.A. Shohat and J.D. Tamarkin, **The problem of moments,** New York city, American mathematical society, 1943, 140 pages, 1943.

[25] P. Stoica and R. Moses, **Introduction to Spectral Analysis**, Prentice Hall, 1997; `http://www.prenhall.com/stoica`

[26] J.L. Walsh, **Interpolation and Approximation by Rational Functions in the Complex Domain,** American Mathematical Society, Providence, RI, Colloquium Publications, vol. 20, 382 pages, 1956.



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, and Iowa State Universities. Since August 1989 he has been with the University of Minnesota where he is a Professor of Electrical and Computer Engineering. He has been an associate editor for IEEE Transactions on Automatic Control, *SIAM Journal on Control and Optimization*, and *Systems and Control Letters*.

He is a recipient of two George Axelby Outstanding Paper awards, for the years 1992 and 1999, both times for joint work with Dr. Malcolm C. Smith.