Signal Estimation via Selective Harmonic Amplification: MUSIC, Redux

Tryphon T. Georgiou, *Fellow, IEEE*

***Abstract*—The technique known as multiple signal classification (MUSIC) is a semi-empirical way to obtain pseudo-spectra that highlight the spectral-energy distribution of a time series. It is based on a certain canonical decomposition of a Toeplitz matrix formed out of an estimated autocorrelation sequence. The purpose of this paper is to present an analogous canonical decomposition of the state-covariance matrix of a stable linear filter driven by a given time-series. Accordingly, the paper concludes with a modification of MUSIC. The new method starts with filtering the time series and then estimating the covariance of the state of the filter. This step in essence improves the signal-to-noise ratio (SNR) by amplifying the contribution to the actual value of the state-covariance of a selected harmonic interval where spectral lines are expected to reside. Then, the method capitalizes on the canonical decomposition of the filter state-covariance to retrieve information on the location of possible spectral lines. The framework requires uniformly spaced samples of the process.**

***Index Terms*—Canonical decomposition of state covariances, Carathéodory–Fejér–Pisarenko, harmonic decomposition.**

I. Introduction

CONSIDER a discrete-time, wide-sense stationary, zero-mean scalar stochastic process $\mathbf{y}_t$ with $t = \cdots, -1, 0, 1, \cdots$ and denote by $d\sigma_y(\theta)$ its power spectrum. The problem of estimating the spectrum from an observed data record

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

of a finite length N has been the focus of an enormous number of studies over the past 40 years or so. The use of Fourier transform, periodogram analysis, modern nonlinear techniques (e.g., maximum entropy method, maximum likelihood, etc.), and subspace methods that resulted from these studies are, by now, standard textbook material, e.g., see [17].

In this work, we are concerned with the so-called MUSIC method [18], [19]. This method is intended for signals consisting of a number of sinusoidal components in noise (ideally not far from being white). It is based on a canonical decomposition of a positive definite Toeplitz matrix, which goes back to Carathéodory and Fejér [8] (cf. [14]) and, more recently, to Pisarenko [18] (cf. [22]). The Toeplitz matrix is formed out of a finite number of covariance lags, and the decomposition suggests an admissible spectrum that consists of a *minimal* number of spectral lines in white noise. MUSIC builds on the Carathéodory–Fejér–Pisarenko decomposition and uses a singular value decomposition (SVD) of the Toeplitz covariance matrix to generate a "pseudo-spectrum" with poles close to the actual sinusoidal modes. Hence, the premise is that the pseudo-spectrum inherits the peaks of the actual power spectrum of the process. In many cases, this appears to work quite effectively with minimal computation cost, and thus, the method has become quite popular (e.g., see [15], [17], and [22]).

The purpose of the present work is to give an analogous canonical decomposition of the state-covariance matrix of a stable linear filter driven by a stationary input (Theorem 1). More specifically, the state covariance of such a filter can be decomposed in a way consistent with the possibility that the input is the superposition of sinusoids in white noise. We show that there is a canonical such decomposition with a *minimal* number of sinusoidal components, i.e., a minimal number of spectral lines (Remark 3).

We then focus on the problem of estimating spectral lines from time series data. The time series data are used to drive a linear filter. The values of the state process of the filter are then used to estimate the state covariance. We use the aforementioned decomposition of state covariances and a MUSIC-like SVD analysis to construct a suitable pseudo-spectrum that inherits the peaks of the actual power spectrum for similar reasons as in (spectral) MUSIC (cf. [17] and [22] and the references therein). The standard (spectral) MUSIC corresponds to the special case where our *input-to-state filter* (since as output we take the state itself) is simply a cascade of unit time delays.

Filtering time-series data and using the state covariance of the filter allows certain control over the contribution of different harmonic intervals on the derived state covariance. Thus, by emphasizing the contribution of a selected frequency band, we can effectively improve resolution over this segment of the spectrum. The use of input-to-state filters originates in Byrnes *et al.* [3], [4], where a bank of first- and second-order filters were used to filter time-series data. The covariances of the individual outputs were shown to give interpolation conditions for the spectrum of the input process at the location of the filter poles. Given such interpolation data, classical Nevanlinna-Pick theory [24] can be used to characterize admissible spectra. Yet, Byrnes *et al.* [3], [4] went on to explore an important subclass of low-dimension regular spectra and developed a new set of *tunable high resolution estimators*. The relevant mathematical theory for low-dimension interpolants has been developed in [3]–[7] and [9]–[11]. Finally, in [3] and [4], it was argued and supported by case studies that proximity of the filter bank poles to a chosen frequency band leads to improved resolution.

Manuscript received February 22, 1999; revised September 23, 1999. This work was supported in part by the NSF and AFOSR. The associate editor coordinating the review of this paper and approving it for publication was Dr. Lal C. Godara.

The author is with the Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN 55455 USA.

Publisher Item Identifier S 1053-587X(00)01550-6.

The input-to-state filters in the present work are a generalization of the filter banks considered in [3] and [4], whereas the state covariances of these filters provide generalized interpolation conditions for power spectra of the input process (i.e., values as well as higher order derivatives; cf. proof of Lemma 5). The frequency plot of the largest singular value of a *suitably normalized* input-to-state map (see Section VIII) appears to provide a more accurate picture of the harmonic selectivity properties of the filters (as compared with the concept of "proximity" of the filterbank poles to the interval of interest proposed in [3] and [4]). Recent work in [12] and [13] expands on the observation (see proof of Lemma 5) that state covariances provide generalized interpolation conditions on the input spectrum and presents a collection of methods that offer improved resolution as compared with their Toeplitz-based counterparts (e.g., for methods of ESPRIT, Capon, and those in [20]).

Section II gives a brief outline of the main technical results underlying traditional MUSIC. Sections III and IV develop technical background related to the algebraic and analytical structure of the states of a linear system. Sections V through VII deal with the structure of state-covariance matrices. The main result (Theorem 1) on the decomposition of state covariance matrices is given in Section VII. This result suggests, accordingly, a generalization of MUSIC, which is discussed in Section VIII. Section IX gives examples to demonstrate the potential of the new method. The method is suitable in cases where prior information allows the selection of a harmonic range of interest. A suitably defined filtering stage improves the signal-to-noise ratio (SNR), and the decomposition theorem suggest ways to identify spectral lines. The new method compares quite favorably with traditional MUSIC.

II. Traditional MUSIC

The key idea behind the methodology of MUSIC is a certain mathematical fact about Toeplitz matrices. This fact can be traced to the work of Carathéodory–Fejér in the early part of the century (see [14]). In modern times, analogous results were rediscovered, and their relevance in signal processing recognized, by Pisarenko, Schmidt, and others (see [22]).

The result we are referring to states that if a positive semidefinite $n \times n$ Hermitian Toeplitz matrix

$$T_{n-1} = \begin{bmatrix} c_0 & c_1 & c_2 & \cdots & c_{n-1} \\ c_{-1} & c_0 & c_1 & \cdots & c_{n-2} \\ \vdots & \vdots & \vdots & & \vdots \\ c_{-n+1} & c_{-n+2} & c_{-n+3} & \cdots & c_0 \end{bmatrix}$$

happens to be singular and of rank $m\ (< n)$ then it is necessarily of the form

$$T_{n-1} = \sum_{i=1}^{m} \rho_i v_i v_i^*$$

where

$$v_i = \begin{bmatrix} 1 \\ e^{j\theta_i} \\ \vdots \\ e^{j(n-1)\theta_i} \end{bmatrix}, \quad \rho_i > 0, \quad \text{for } i = 1, 2, \cdots, m$$

$\theta_\ell \neq \theta_k$ if $\ell \neq k$, $j := \sqrt{-1}$, and $*$ denotes the "complex conjugate transpose." Further, T_{n-1} has a unique (left) zero-eigenvector of the form

$$\phi = [a_0 \quad \cdots \quad a_m \quad 0 \quad \cdots \quad 0]$$

i.e., $\phi \cdot T_{n-1} = 0$. This eigenvector satisfies

$$\phi \cdot v_i = 0, \quad \text{for } i = 1, 2, \cdots, m$$

i.e., the modes $e^{j\theta_i}$ are precisely the roots of the polynomial

$$a(\lambda) = \sum_{k=0}^{m} a_k \lambda^k.$$

An elegant and useful reformulation of this result makes use of the elements in the singular value decomposition

$$T_{n-1} = U \text{ diag } \{\sigma_1, \sigma_2, \cdots, \sigma_m, 0, \cdots, 0\} U^*$$

of T_{n-1}, where U is a unitary matrix, and $\sigma_i > 0$ $(i = 1, \cdots, m)$. 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 that span the (right) null space of T_{n-1} (which is denoted by $\mathcal{N}_{\text{right}}(T_{n-1})$), then the non-negative trigonometric polynomial

$$d(e^{j\theta}, e^{-j\theta}) := [1 \quad e^{-j\theta} \quad \cdots \quad e^{-j(n-1)\theta}] \cdot U_{m+1:n} \cdot U_{m+1:n}^* \begin{bmatrix} 1 \\ e^{j\theta} \\ \vdots \\ e^{j(n-1)\theta} \end{bmatrix} \quad (1)$$

has m roots precisely at $\{\theta_1, \cdots, \theta_m\}$. 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. A proof of this result can be found in [22].

The significance in the context of signal processing stems from the fact that if c_k $(k = 0, 1, \cdots)$ denotes the autocorrelation function of $\mathbf{y}_t$, that is

$$c_k = \mathbf{E}\{\mathbf{y}_t \mathbf{y}_{t+k}^*\} = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-jk\theta} \, d\sigma_y(\theta)$$

and if $\mathbf{y}_t$ contains sinusoidal components of interest, then the frequencies θ_i's of such components can be identified as above via a singular value decomposition of the corresponding Toeplitz matrix formed out of the autocorrelation function. In case there is background noise (assumed white or close to being white), the columns or $U_{m+1:n}$ are taken to span the subspace of the lowest eigenvalue(s) of T_{n-1} (see [22] as well as Section VII below).

III. Input-to-State Filters

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

$$x_k = Ax_{k-1} + by_k, \quad \text{where } k = \cdots, -1, 0, 1, \cdots.$$

The nontraditional indexing of the input y_k is intended to simplify notation. 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. That is

$$\mathcal{C}_{A,b} := [b, Ab, \cdots, A^{n-1}b]$$

is assumed to have full rank, and the eigenvalues of A to have modulus less than one. Our analysis applies to A, b complex, but it is carried out in a way that requires only real arithmetic when A, b and $\mathbf{y}_k$ are real.

We use λ to designate the transform of the delay operator

$$x_k \to x_{k-1}.$$

The transfer function of the above system, which is referred to as an *input-to-state filter* or a *polyfilter*, is

$$G(\lambda) = \begin{bmatrix} g_1(\lambda) \\ g_2(\lambda) \\ \vdots \\ g_n(\lambda) \end{bmatrix} = (I - \lambda A)^{-1} b. \tag{2}$$

We mention two special cases (see also [3] and [4]). First, we have the case where A is a *diagonal* matrix

$$A = \begin{bmatrix} p_1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & p_2 & 0 & \cdots & 0 & 0 \\ 0 & 0 & p_3 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & p_n \end{bmatrix} \quad \text{and} \quad b = \begin{bmatrix} 1 \\ 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}. \tag{3}$$

In this case

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

are the standard Cauchy kernels, and $G(\lambda)$ a bank of parallel first-order filters. The next interesting special case is when A is the $n \times n$ *companion* matrix

$$A = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \end{bmatrix} \quad \text{and} \quad b = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \tag{4}$$

This case corresponds to

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

The terminology *polyfilter* suggests the general case where A has an arbitrary Jordan structure (subject to the controllability requirement), and thereby, $G(\lambda)$ is a filterbank with higher order subcomponents of possibly different dimension.

The entries $g_i(\lambda)$ of a polyfilter, as we will see below, form a basis for a certain subspace of H_2 (where H_2 is the space of analytic functions in the unit disc with square-integrable radial limits). Further, these functions are generalized Cauchy kernels whose inner product "evaluates" functions in H_2 at the eigenvalues of A^*. More specifically, if $f(\lambda) \in H_2$, then

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

where

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

The expression $b^* f(A^*) e_i^*$ depends either on the value of $f(\lambda)$ at $\overline{p}_i$, as in the case of (3), or on the value of the ith derivative of $f(\lambda)$ at 0, as in the case of (4) or, in general, on the values of f and its derivatives (up to the multiplicity minus one) at the eigenvalues of A^* in a way prescribed by the eigenstructure of A and its relation to the chosen vector b. Conversely, the values of f on the spectrum of A^*, which appear in $f(A^*)$, can be calculated from $b^* f(A^*)$, as shown in the following lemma.

Lemma 1: If A, b is a controllable pair and $f(s), w(s)$ functions defined on the spectrum of A, then

$$f(A) = w(A) \Leftrightarrow f(A) b = w(A) b.$$

Proof: Clearly, $(\Rightarrow)$ is obvious. To show the other way, note that

$$f(A) b = w(A) b \Rightarrow f(A) A^k b = w(A) A^k b$$

since A commutes with $f(A)$ and $w(A)$. Thus

$$f(A) b = w(A) b \Rightarrow f(A) \mathcal{C}_{A,b} = g(A) \mathcal{C}_{A,b}.$$

This completes the proof since $\mathcal{C}_{A,b}$ is invertible. ■

IV. Coinvariant Subspace $\mathcal{K}$

The subspace of H_2-functions $f(\lambda)$ that vanish on the spectrum of A^*, i.e., for which $f(A^*) = 0$, is given by $B(\lambda) H_2$, where

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

is an all-pass function (which is otherwise known as finite inner or Blaschke product). Its orthogonal complement

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

plays a key role in interpolation theory.

The standard (right) shift operator defined on H_2 is

$$S: x(\lambda) \to \lambda x(\lambda).$$

The subspace $\mathcal{K}$ is called "co-invariant" because it is invariant under the adjoint operator

$$S^*: x(\lambda) \to \boldsymbol{\Pi}_{H_2}\, \lambda^{-1} x(\lambda).$$

Note that by $\boldsymbol{\Pi}_{H_2}$, we denote the orthogonal projection from L_2 onto H_2. Similarly, for $\boldsymbol{\Pi}_{\mathcal{K}}$.

Lemma 2: The entries of $G(\lambda)$ form a basis of $\mathcal{K}$.

Proof: The statement of the lemma follows from the following.

i) The entries of $G(\lambda)$ are linearly independent.

ii) The dimension of $\mathcal{K}$ is n and coincides with the number of entries in $G(\lambda)$.

iii) span$\{g_i(\lambda): i = 1, \cdots, n\}$ remains invariant under S^*.

Claim i) follows from the controllability of (A, b). Claim ii) is standard [23]. Claim iii) is easily verified since

$$\begin{aligned} S^*: c \cdot (I - \lambda A)^{-1} b \rightarrow & \boldsymbol{\Pi}_{H_2} \lambda^{-1} (c \cdot (I - \lambda A)^{-1} b) \\ &= cA \cdot (I - \lambda A)^{-1} b. \end{aligned} \tag{5}$$

It is interesting to note from the above proof that A is a matrix representation of the left shift S^* with respect to the basis $G(\lambda)$. Hence, for any functions f, w defined on the spectrum of A, we have that $f(S^*) = w(S^*) \Leftrightarrow f(A) = w(A)$.

V. Power Spectra and State-Covariances

Let the process $\mathbf{y}_t$ pass through $G(\lambda)$, and let the state process be denoted by $\mathbf{x}_t$. The covariance of $\mathbf{x}_t$ is given by

$$P = \frac{1}{2\pi} \int_{-\pi}^{\pi} (G(e^{j\theta})\, d\sigma_y(\theta) G(e^{j\theta})^*). \tag{6}$$

Since $d\sigma_y(\theta)$ is a bounded positive (non-negative) measure, then P is a positive (semi-) definite $n \times n$ matrix.

Two simple cases of special interest are as follows. First, if A is a companion matrix as in (4) and b is the corresponding input vector, then the state process is

$$\mathbf{x}_t = \begin{bmatrix} \mathbf{y}_t \\ \mathbf{y}_{t+1} \\ \vdots \\ \mathbf{y}_{t+n-1} \end{bmatrix}$$

and the covariance coincides with the Toeplitz matrix T_{n-1} given earlier. Second, if A is a diagonal matrix as in (3) and b is the corresponding input vector, then the covariance of the state-space process is in the usual Pick-matrix form

$$\left[\frac{w_i + \overline{w}_k}{1 - p_i \overline{p}_k} \right]_{i,k=1}^{n}. \tag{7}$$

In this case, the values w_i for $i = 1, \cdots, n$ relate to the covariance of the individual entries of the vector-process $\mathbf{x}_t$ (see [4] for details). In general, the state-covariance P is characterized by

$$P - APA^* = [b \quad v] \begin{bmatrix} c_0 & -1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} b^* \\ v^* \end{bmatrix}$$

for some $c_0 > 0$ and a suitable vector v. However, this last fact will not be further exploited.

VI. Singular Spectra

We now consider power spectra of the form

$$d\sigma_y(\theta) = \dot{\sigma}_y \, d\theta + \sum_{k=1}^{m} 2\pi \rho_k \delta(\theta - \theta_k)\, d\theta \tag{8}$$

where $\delta(\cdot)$ denotes the Dirac impulse function. These contain both a regular part as well as a singular part. The singular part, which is the sum containing the Dirac functions, represents spectral lines that we typically want to identify. The residues ρ_k represent the energy density contained in the corresponding sinusoidal components of $\mathbf{y}_t$, whereas $\dot{\sigma}_y(\theta)$ is the spectral density function that is defined a.e. and is non-negative.

We wish to study the eigenstructure of the corresponding state covariance when $\mathbf{y}_t$ is fed as input to a polyfilter $G(\lambda)$. To this end, consider a vector

$$c = [c_1, c_2, \cdots, c_n]$$

and the quadratic form

$$\begin{aligned} cPc^* &= \frac{1}{2\pi} \int_{-\pi}^{\pi} (cG(e^{j\theta})\, d\sigma_y(\theta) G(e^{j\theta})^* c^*) \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} |H(e^{j\theta})|^2 \, d\sigma_y(\theta) \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} |H(e^{j\theta})|^2 \dot{\sigma}(\theta)\, d\theta + \sum_{k=1}^{m} \rho_k |H(e^{j\theta_k})|^2 \end{aligned}$$

where $H(\lambda) = c(I - \lambda A)^{-1} b$. Since $|H(e^{j\theta})|^2$ is positive almost everywhere (provided $c \neq 0$), the quadratic form is singular only if both the regular spectrum $\dot{\sigma}(\theta)\, d\theta$ is identically zero, and

$$H(e^{j\theta_k}) = 0 \quad \text{for } k = 1, \cdots, m.$$

Note that $H(\lambda)$ is a rational function that can be written as

$$H(\lambda) = \frac{n(\lambda)}{\chi(\lambda)}$$

with the denominator $\chi(\lambda) = \det(I - \lambda A)$ of degree $\leq n$ and the numerator $n(\lambda)$ of degree $< n$. It follows that the numerator has to vanish at the points $e^{j\theta_k}$, i.e.,

$$n(e^{j\theta_k}) = 0 \quad \text{for } k = 1, 2, \cdots, m. \tag{9}$$

Hence, assuming the θ_k's are distinct, $m < n$. We summarize these conclusions in the following proposition.

Proposition 1: The $n \times n$ state-covariance matrix defined in (6) is singular of rank $m < n$ if and only if it is of the form

$$P = \sum_{k=1}^{m} \rho_k G(e^{j\theta_k}) G(e^{j\theta_k})^* \tag{10}$$

for a selection of distinct θ_k with $k = 1, \cdots, m$.

Proof: The "if" part is obvious. The "only if" was shown by the arguments leading to the proposition. ■

We now study the correspondence between the elements in the above decomposition and the zero-eigenvectors of P. Therefore, let P be singular of rank $m < n$, and define the (left) null space of P

$$\mathcal{N}_{\text{left}}(P) := \{c: cP = 0\}$$

and the space of polynomials

$$\mathcal{N} := \{n(\lambda): n(\lambda) \text{ polynomial of degree } < n, \text{ such that (9) holds}\}.$$

These two spaces are both of dimension $n-m$, and in fact, they are in bijective correspondence via

$$\phi: c \to n(\lambda) = c(I - \lambda A)^{-1} b \cdot \det(I - \lambda A). \tag{11}$$

Due to controllability of (A, b), ϕ is already a bijection between n vectors and polynomials of degree $< n$. Then, $cP = 0$ is equivalent $cPc^* = 0$, which, in turn, is equivalent to (9) as it follows from the earlier arguments.

A basis for $\mathcal{N}_{\text{left}}(P)$ can be easily identified from a singular value decomposition of P. More specifically, let

$$P = U \text{ diag } \{\sigma_1, \sigma_2, \cdots, \sigma_m, 0, \cdots, 0\} U^* \tag{12}$$

with U unitary and $\sigma_i > 0$ $(i = 1, \cdots, m)$, and let

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

with $U_{m+1:n}$ representing the matrix formed out of the last $n-m$ columns of U. Then, the rows of $U^*_{m+1:n}$ span $\mathcal{N}_{\text{left}}(P)$. Since the range of P is spanned by $G(e^{j\theta_k})$ for $k = 1, \cdots, m$, it follows that the non-negative trigonometric polynomial

$$d(e^{j\theta}, e^{-j\theta}) := G^*(e^{-j\theta}) U_{m+1:n} U^*_{m+1:n} G(e^{j\theta}) \tag{13}$$

has m roots at $\{\theta_1, \cdots, \theta_m\}$. It should be noted that just as in the case of Section II, $d(\lambda, \lambda^{-1})$ may have additional roots off the unit circle, but it has no root on the circle other than the above. This is a direct consequence of the following lemma.

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

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

are linearly independent.

Proof: It suffices to prove the claim for $m = n-1$. Therefore, let $m = n-1$, and consider the determinant of the $n \times n$ matrix function

$$M(\lambda) = [G(\lambda), G(e^{j\theta_1}), \cdots, G(e^{j\theta_m})]. \tag{14}$$

Clearly, $\det\ M(\lambda)$ is an element of $\mathcal{K}$ since all non-constant entries of the first column belong to $\mathcal{K}$. Thus

$$\det\ M(\lambda) = c(I - \lambda A)^{-1} b$$

and it can have at most $n-1$ roots. From (14), these $n-1$ roots are precisely $\{e^{j\theta_1}, \cdots, e^{j\theta_m}\}$. Thus, $M(e^{j\theta_0})$ is nonsingular. ■

We summarize our conclusions in the following key proposition.

Proposition 2: If P given by (6) is singular and is of rank $m < n$ and if $U_{m+1:n}$ is obtained as above from a singular value decomposition of P, then the spectrum of $\mathbf{y}_t$ is discrete and of the form

$$d\sigma_y(\theta) = \sum_{k=1}^{m} 2\pi \rho_k \delta(\theta - \theta_k)\, d\theta$$

where $e^{j\theta_k}$ $(k = 1, \cdots, m)$ are distinct roots of

$$d(\lambda, \lambda^{-1}) := (I - \lambda^{-1} A^*)^{-1} b^* U_{m+1:n} U^*_{m+1:n} b (I - \lambda A)^{-1}$$

on the unit circle. Moreover, $d(\lambda, \lambda^{-1})$ has no root on the circle other than the ones above, i.e., other than $e^{j\theta_k}$ $(k = 1, \cdots, m)$.

Proof: The proof was shown by the argument leading to the proposition. ■

VII. Sinusoids in White Noise

The typical assumption underlying MUSIC is that the signals are corrupted by white noise. The spectrum of $\mathbf{y}_t$ in this case is

$$d\sigma_y(\theta) = \rho_0\, d\theta + \sum_{k=1}^{m} 2\pi \rho_k \delta(\theta - \theta_k)\, d\theta \tag{15}$$

and the corresponding state covariance is of the form

$$P = \rho_0 \frac{1}{2\pi} \int_{-\pi}^{\pi} G(e^{j\theta}) G(e^{j\theta})^*\, d\theta + \sum_{k=1}^{m} \rho_k G(e^{j\theta_k}) G(e^{j\theta_k})^*. \tag{16}$$

Note that the term

$$E := \frac{1}{2\pi} \int_{-\pi}^{\pi} G(e^{j\theta}) G(e^{j\theta})^*\, d\theta$$

is a familiar object. It is precisely a controllability Grammian that satisfies the Lyapunov equation $E - AEA^* = bb^*$. Conversely, given a state-covariance P and the filter parameters A, b, then we compute E as above and then the eigenvalues of the matrix pencil

$$P - \rho E. \tag{17}$$

We denote by ρ_0 the smallest eigenvalue (which is positive by virtue of the fact that both P and E are Hermitian and positive definite). Regardless of the statistics of $\mathbf{y}_t$, this value is a candidate for white noise level in $\mathbf{y}_t$. Indeed, we will show that the decomposition (16) always holds, with no need for any specific requirement on the spectrum of $\mathbf{y}_t$. A key step is to prove that

$$P_0 := P - \rho_0 E \tag{18}$$

is again a state covariance of the form given in (6), with a nonnegative measure $d\sigma(\theta)$, in order to take advantage of the singularity of P_0 and apply our earlier results in Proposition 1. To prove existence of such a $d\sigma(\theta)$, we need to use certain results from interpolation theory. The first lemma below is a version of the standard commutant lifting theorem, which is due to Sarason and, in the most general form, due to Sz.-Nagy and Foias.

Lemma 4 [23, Th. 1]: 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 $\phi(\lambda)$ that is analytic in the open disc and has positive real part there such that

$$T = \phi(S).$$

Proof: The statement is isomorphic to that in [23], except that it is phrased in terms of a positive-real property as opposed to contractiveness. To make the correspondence, it suffices to combine the following well-known facts.

i) T has positive real part if and only if

$$T_o := (I - T)(I + T)^{-1}$$

is strictly contractive, i.e., its norm is < 1.

ii) T commutes with S if and only if T_o does.

iii) A function $\psi(\lambda)$, which is analytic in the disc, has positive real part if and only if the function

$$\psi(\lambda) := (1 - \phi(\lambda))(1 + \phi(\lambda))^{-1}$$

is also analytic and has norm bounded by one, i.e., $|\psi(\lambda)| < 1$ inside the disc.

iv) $\phi(\lambda)$ interpolates T, i.e., $\phi(S) = T$, iff $\psi(\lambda)$ interpolates T_o. ∎

Lemma 5: If $A, b, G(\lambda), P, E, \rho_0$ and P_0 are as defined earlier, then P_0 is a state-covariance matrix for $G(\lambda)$ and a suitable stochastic input, i.e., that

$$P_0 = \frac{1}{2\pi} \int_{-\pi}^{\pi} (G(e^{j\theta})\, d\sigma(\theta) G(e^{j\theta})^*) \tag{19}$$

for some bounded non-negative measure $d\sigma(\theta)$.

Proof: Power spectra, like $d\sigma_y(\theta)$, are characterized by the property that $\sigma_y(\theta)$ is a nondecreasing function of bounded variation or that $d\sigma_y(\theta)$ is a bounded non-negative measure. It is a well-known fact (e.g., see [1]) that if $d\sigma_y(\theta)$ is as above, then

$$\phi(\lambda) := \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{1 + \lambda e^{-j\theta}}{1 - \lambda e^{-j\theta}}\, d\sigma_y(\theta) \tag{20}$$

is analytic in the open disc with positive real part. Conversely, any such *positive-real function* $\phi(\lambda)$ admits a representation as in (20) with a suitable non-negative measure.

To prove the claim in the lemma, begin with $\sigma_y(\theta)$, define $\phi(\lambda)$ as above, and then define the operator

$$T := \phi(S)\colon x(\lambda) \rightarrow \boldsymbol{\Pi}_{\mathcal{K}} \phi(\lambda) x(\lambda)$$

where $x(\lambda) \in \mathcal{K}$. In the present context, it is easier to work with adjoint operators. Thus, with regard to the basis $G(\lambda)$, S^* corresponds to the matrix A, whereas T^* corresponds to a matrix W, i.e.,

$$T^*\colon cG(\lambda) \rightarrow cWG(\lambda).$$

Since T and S commute, W and A also commute. The real part of T is

$$\begin{aligned}&\tfrac{1}{2}(\lambda\langle cG(\lambda), T^* cG(\lambda)\rangle + \lambda\langle T^* cG(\lambda), cG(\lambda)\rangle)\\ &\quad= \tfrac{1}{2}\{c(EW^* + WE)c^*\}.\end{aligned}$$

Since $d\sigma_y(\theta) \sim \text{Re}\ (\phi(e^{j\theta}))\ d\theta$ (see e.g., [1]), the state-covariance matrix P is precisely the matrix that defines the above quadratic form; hence

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

The first part of the argument is fairly standard. To complete the proof, we only need to trace the above steps backwards for P_0. In this way, we will establish that P_0 is given by (19) with a suitable $d\sigma(\theta)$. First, we note that

$$P_0 = \frac{1}{2}(EW_0^* + W_0 E)$$

with $W_0 := W - (\rho_0/2)I$ and that it is non-negative definite. Then, define the operator T_0 on $\mathcal{K}$ through its adjoint by

$$T_0^*\colon cG(\lambda) \rightarrow cW_0 G(\lambda).$$

Since W_0 differs from W only by $(\rho_0/2)I$, W_0 commutes with A as well. Hence, T_0 commutes with S. Since the real part of T_0 is $P_0 \geq 0$, it follows from Lemma 4 that T_0 can be interpolated by a positive-real function $\phi_0(\lambda)$. Invoking the existence of an integral representation for $\phi_0(\lambda)$ as in (20) gives us the required measure $d\sigma(\theta)$ in (19). ∎

We now can apply our earlier analysis of singular state-covariance matrices to P_0 and deduce our main theorem.

Theorem 1: Consider the state-covariance P of a linear filter having parameters A, b, with A, b controllable and input $\mathbf{y}_t$ having power spectrum any general bounded non-negative measure $d\sigma_y(\theta)$. There is a unique decomposition of P in the form

$$\begin{aligned}P = \rho_0 \frac{1}{2\pi} \int_{-\pi}^{\pi} G(e^{j\theta})G(e^{j\theta})^*\, d\theta\\ + \sum_{k=1}^{m} \rho_k G(e^{j\theta_k})G(e^{j\theta_k})^*\end{aligned} \tag{21}$$

where $G(\lambda) = (I - \lambda A)^{-1} b$ and where we have the following.

i) The matrix

$$E := \frac{1}{2\pi} \int_{-\pi}^{\pi} G(e^{j\theta})G(e^{j\theta})^*\, d\theta$$

is the solution of the Lyapunov equation

$$E - AEA^* = bb^*.$$

ii) ρ_0 is the smallest eigenvalue of the matrix pencil

$$P - \rho_0 E.$$

iii) $m = n - \text{rank}(P - \rho_0 E)$.

iv) $e^{j\theta_k}$ for $k = 1, \cdots, m$ are roots of the trigonometric polynomial

$$d(\lambda, \lambda^{-1}) := G^*(\lambda^{-1}) U_{m+1:n} U_{m+1:n}^* G(\lambda) \tag{22}$$

on the unit circle, where $U_{m+1:n}$ is a basis for $\mathcal{N}_{\text{right}}(P - \rho_0 E)$. Further, $d(\lambda, \lambda^{-1})$ has no other root on the unit circle.

Finally, the residues ρ_k above are given by

$$\mathrm{diag}\{\rho_1, \cdots, \rho_m\} = V^{-1}(P - \rho_0 E)(V^{-1})^*$$

where

$$V = \begin{bmatrix} G(e^{j\theta_1})^* \\ \cdots \\ G(e^{j\theta_m})^* \end{bmatrix} \cdot [G(e^{j\theta_1}), \cdots, G(e^{j\theta_m})]. \qquad (23)$$

Proof: The proof follows from Lemma 5, Propositions 1 and 2, and the discussion leading to Lemmas 4 and 5. ■

We wish to emphasize that the theorem makes no particular assumption on the spectrum of the input process. Yet, it states that the state-covariance can always be decomposed in a way consistent with the possibility that the input was the superposition of sinusoids in white noise.

Remark 1: The special case that corresponds to A, b as in (4) reduces to the Carathéodory–Fejér results and is the basis of the method introduced by Pisarenko (cf. [14] and [22]).

Remark 2: Consider the case where $G(\lambda)$ is a general transfer function, i.e., when $G(\lambda) = c(I - \lambda A)^{-1}b$, and P is the output covariance. If the input consists of sinusoids in white noise, then the output covariance is easily seen to be, again, in the form (21). Thus, we might be led to conjecture that a statement analogous to Theorem 1 holds in this case as well. In general, this conjecture is false. This is shown in the following simple counterexample. (Note that c has to have more than one row for the conjecture to be nontrivial and false.)

Counterexample: Let A, b be as in (4) with $n = 3$, and let

$$c = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \end{bmatrix}.$$

Assume that the input process is real with covariance sequence $(c_0, c_1, c_2, \cdots)$. Then, the state-covariance matrix is

$$T_2 = \begin{bmatrix} c_0 & c_1 & c_2 \\ c_1 & c_0 & c_1 \\ c_2 & c_1 & c_0 \end{bmatrix}$$

the output covariance is

$$P = \begin{bmatrix} 3c_0 + 4c_1 + 2c_2 & c_0 + 2c_2 \\ c_0 + 2c_2 & 3c_0 - 4c_1 + 2c_2 \end{bmatrix}$$

and the contribution of unit variance white noise to the output covariance is

$$E = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}.$$

Take $c_0 = 4$, $c_1 = -1$, and $c_2 = 3$. Then, $T_2 > 0$, and hence, $P > 0$. Now, take $\rho_0 = 2$. The difference

$$P - \rho_0 E = \begin{bmatrix} 8 & 8 \\ 8 & 16 \end{bmatrix}$$

is still positive definite, yet no input process could have generated such an output covariance matrix (because $T_2 - \rho_0 I \not> 0$). To see this more explicitly, solve

$$\begin{bmatrix} 3\hat{c}_0 + 4\hat{c}_1 + 2\hat{c}_2 & \hat{c}_0 + 2\hat{c}_2 \\ \hat{c}_0 + 2\hat{c}_2 & 3\hat{c}_0 - 4\hat{c}_1 + 2\hat{c}_2 \end{bmatrix} = \begin{bmatrix} 8 & 8 \\ 8 & 16 \end{bmatrix}$$

for the covariance sequence of such a hypothetical input. This, of course, gives $\hat{c}_0 = 2$, $\hat{c}_1 = -1$, and $\hat{c}_2 = 3$, which is inadmissible.

Thus, Theorem 1 is valid in general only for input-to-state filters.

Remark 3: The decomposition (21) in Theorem 1 contains a *minimal* number of dyads $G(e^{j\theta_k})G(e^{j\theta_k})^*$, i.e., any other decomposition

$$P = \hat{\rho}_0 E + \sum_{k=1}^{\hat{m}} \hat{\rho}_k G(e^{j\hat{\theta}_k})G(e^{j\hat{\theta}_k})^* \qquad (24)$$

with positive coefficients $\hat{\rho}_k$ $(k = 0, 1, \cdots, \hat{m})$ is either identical to (21) or $m < \hat{m}$. The minimality justifies our use of the term *canonical* for the decomposition of Theorem 1 as it is customary in the theory of the moment problem (see [2], [16], and [21]).

To show our claim of minimality, we first note that $m \le n-1$ from Theorem 1 iii), whereas $\rho_0 \le \hat{\rho}_0$ from Theorem 1 ii). We now consider two cases: If $\rho_0 = \hat{\rho}_0$, then

$$\sum_{k=1}^{m} \rho_k G(e^{j\theta_k})G(e^{j\theta_k})^* = \sum_{k=1}^{\hat{m}} \hat{\rho}_k G(e^{j\hat{\theta}_k})G(e^{j\hat{\theta}_k})^*.$$

Hence, we have the matrix

$$[G(e^{j\theta_1}), G(e^{j\theta_2}), \cdots, G(e^{j\theta_m})]$$

and its counterpart with the hats ought to have the same range (which coincides with the range of $P - \rho_0 E$). Therefore, we can deduce from Lemma 3 that the set $\{\theta_1, \theta_2, \cdots, \theta_m\}$ and its counterpart with the hats are identical. It now follows easily that decomposition (24) is identical to the one in Theorem 1, modulo re-ordering the indices. If, on the other hand, $\rho_0 < \hat{\rho}_0$, then by comparing the two decompositions, we get that

$$\sum_{k=1}^{\hat{m}} \hat{\rho}_k G(e^{j\hat{\theta}_k})G(e^{j\hat{\theta}_k})^* > (\rho_0 - \hat{\rho}_0)E > 0.$$

(Here, $A > B$ designates that $A - B$ is positive definite.) It follows that

$$[G(e^{j\hat{\theta}_1}), G(e^{j\hat{\theta}_2}), \cdots, G(e^{j\hat{\theta}_{\hat{m}}})]$$

has full row rank. Hence, $\hat{m} \ge n$, and the decomposition with the hats cannot be minimal.

VIII. MUSIC VIA SELECTIVE HARMONIC AMPLIFICATION

We now come to the point of integrating the above conclusions into the analog of MUSIC for general state-covariance statistics.

A. Choice of A, b

We first observe that there is a natural group action on the parameters A, b, P, E of Theorem 1, namely the similarity transformation on A, b and the corresponding congruence transformation on P, E. These transformations leave the frequencies θ_k and the weights ρ_k in decomposition (21) invariant. Indeed, if $A \to A_o = CAC^{-1}$, $b \to b_o = Cb$, $P \to P_o = CPC^*$, and

$E \to E_o = CEC^*$, with C square and invertible, then (21) holds for the new parameters A_o, b_o, P_o, E_o and the same frequencies θ_k and weights ρ_k as before, i.e.,

$$P_o = \rho_0 E_o + \sum_{k=1}^{m} \rho_k C G(e^{j\theta_k}) G(e^{j\theta_k})^* C^* = \rho_0 E_o + \sum_{k=1}^{m} \rho_k G_o(e^{j\theta_k}) G_o(e^{j\theta_k})^*$$

where $G_o(\lambda) = (I - \lambda A_o)^{-1} b_o$. In order to capture the selectivity properties of input-to-state filters, we need a frequency-response-like function that is invariant under similarity transformation. A suitable such object can be obtained by normalizing the parameters A, b so that the state-covariance due to the input noise is the identity matrix. Hence, we define the *selectivity measure* of an input-to-state filter with parameters A, b (controllable and stable) as the function

$$\mathrm{s}_{A,b}\colon \theta \mapsto \|G_o(e^{j\theta})\| = \|E^{-1/2} G(e^{j\theta})\| \tag{25}$$

where $E^{1/2}$ denotes the Hermitian square-root of the solution E to the Lyapunov equation $E - AEA^* = bb^*$. An alternative justification can be argued based on the formula

$$\mathrm{trace}(P_o) = \int_0^{2\pi} \|G_o(e^{j\theta})\|^2 \, d\sigma_y(\theta) \tag{26}$$

where $\mathrm{s}_{A,b}(\theta)$ weighs the contribution of the spectrum across frequencies on the "size" of the resulting state-covariance.

The filter parameters can now be specified so that the selectivity measure has passband characteristic over a range of interest, i.e., the range of frequencies where the sinusoidal components are expected to reside. A simple approach is to select A in a Jordan form

$$A = \begin{bmatrix} r & 1 & 0 & \cdots & 0 \\ 0 & r & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & r \end{bmatrix} \tag{27}$$

and r in the interval $0 < r < 1$ or, in $-1 < r < 0$, depending on whether a lowpass or a highpass characteristic is desired, respectively, whereas the choice $b = [0, \cdots, 0, 1]'$ can be used. Similarly, a block Jordan structure

$$A = \begin{bmatrix} R & I & 0 & \cdots & 0 \\ 0 & R & I & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & I \\ 0 & 0 & 0 & \cdots & R \end{bmatrix} \tag{28}$$

where

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

with $0 < r < 1$ and $0 < \theta_0 < \pi$ generates bandpass characteristics. Accordingly, the same choice of $b = [0, \cdots, 0, 1]'$ can be used again.

The selectivity measure is shown for two representative cases in Figs. 1 and 2. The first is generated according to (27) with $r = 0.1$ and the second according to (28) with $r = 0.3$ and $\theta_0 = \pi/2$. In both cases, A was chosen 30 × 30. These filters will be used in the examples below.

B. State-Covariance Estimates

Having decided on the filter parameters A, b and having the observation record

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

available, we generate the state vector-process x_k from the equation $x_k = Ax_{k-1} + by_k$ for $k = 0, \cdots, N-1$. (Depending on the dynamics of the filter, we might choose to go even a little beyond $N-1$, but this will not be followed up in here.) Initial conditions can be taken as zero. Then, the time series

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

can be used to estimate the sample state-covariance matrix P via

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

Here, ℓ could either be 0 or equal to a value that gives a level of stationarity for the state time-series after $t \geq \ell$.

C. New MUSIC

1) We need to specify the integer m. It is advisable to chose m slightly larger than the expected number, say m_0, of sinusoidal components in the spectrum of $\mathbf{y}_t$. The method produces m candidate frequencies, and then a subset of m_0 frequencies that are closest to the expected range can be selected.
2) After selecting m, we construct $U_{m+1:n}$ to span the (right) eigensubspace corresponding to the smallest $n - m$ eigenvalues of the matrix pencil $P - \rho E$. This can be done by standard eigenvalue decomposition of the pencil.
3) Finally, the following two options can be followed.

   Root MUSIC: As an estimate for the frequencies θ_k $(k = 1, \cdots, m)$ of potential sinusoidal components take the angle of m roots of $d(\lambda, \lambda^{-1})$ in (22), which are closest to the unit circle and of modulus ≤ 1.

   Spectral MUSIC: As an estimate for the frequencies of potential sinusoidal components take the frequencies where the pseudo-spectrum

$$\frac{1}{d(e^{j\theta}, e^{-j\theta})}$$

   has its peaks.

Remark 4: In case the lowest eigenvalue of the pencil $P - \rho E$ has multiplicity $n - m$, then we can compute an exact candidate spectrum (as opposed to only a pseudo spectrum). This follows from Theorem 1 since, besides ρ_0 and θ_k for $k = 1, \cdots, m$, we can also identify ρ_k for $k = 1, \cdots, m$ from (23). Generically, this is the case only when $m = n - 1$. This corresponds to the method of Pisarenko [cf. (22)]. □

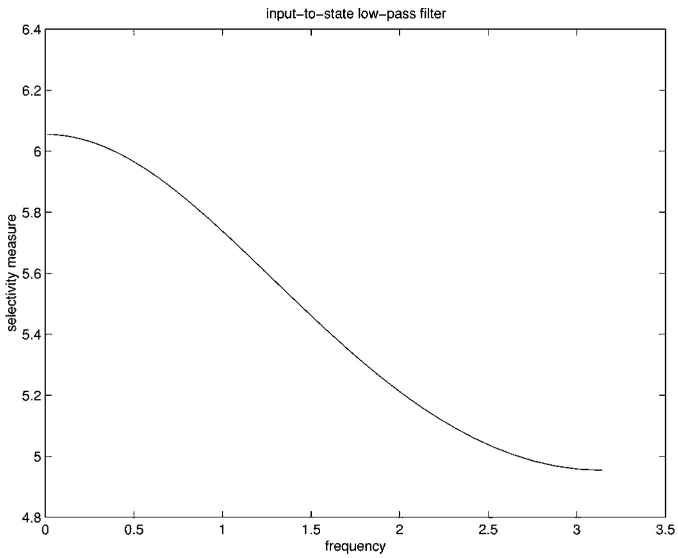


Fig. 1. Selectivity measure for a lowpass input-to-state filter.

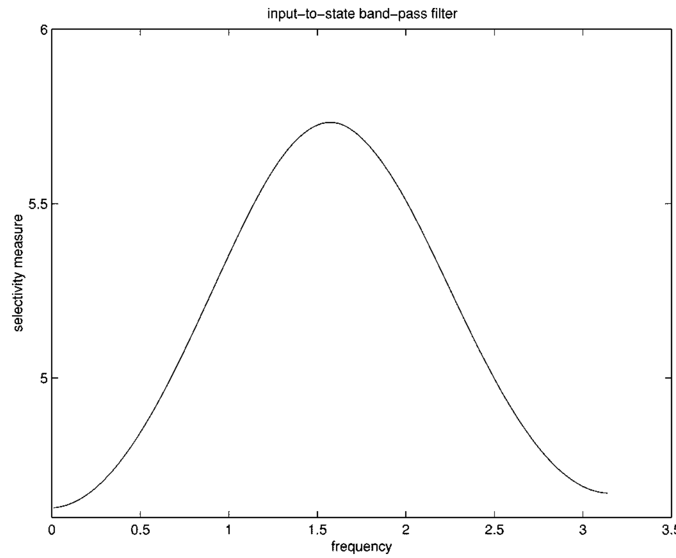


Fig. 2. Selectivity measure for a bandpass input-to-state filter.

IX. Case Studies

We now compare MUSIC in the standard implementation (Method I) versus the implementation suggested in Section VIII (Method II).

A. Example 1: Lowpass Filter

We simulated numerically a random signal

$$y_t = \nu_t + \rho_1 \sin(\omega_1 t + \phi_1) + \rho_2 \sin(\omega_2 t + \phi_2) \quad \text{for } t = 0, 1, \cdots, N-1$$

where $N = 100$, $\omega_1 = 0.4$[rad], $\omega_2 = 0.5$[rad], $\rho_1 = 0.5$, and $\rho_2 = 1$, whereas ν_t $(t = 0, \cdots, N-1)$ were generated as unit variance independent Gaussian random variables and ϕ_1, ϕ_2 as random variables uniformly distributed on $[0, \pi]$.

We used $n = 30$ and $m = 4$. Note that m was taken larger than the known number of sinusoidal components, which in this case is $m_0 = 2$. Thus, among the frequencies $\theta_1, \theta_2, \theta_3, \theta_4$ that MUSIC produces, only the two most "dominant" ones, or the two "closest to the expected frequency range," should be taken as the estimated values for ω_1 and ω_2. We first present one simulation in some detail and then summarize the results of additional runs.

Method I represents the standard implementation of MUSIC, e.g., as described in [22]. This corresponds to the case where all poles are chosen at the origin, and the filter data A, b are chosen according to (4). For this case, the frequency response is not shown since it is constant across frequencies. Hence, the standard implementation shows no preference for any particular frequency band.

Method II represents the implementation of Section VIII. In this case, since the sinusoidal frequencies are (known to be) at the low-frequency part of the spectrum, we make use of a lowpass filter. The parameters for A, b were chosen according to (27) with $r = 0.1$. The normalized frequency response is plotted in Fig. 1.

The time series of a typical simulation is shown in Fig. 3. The (normalized) pseudo-spectra produced by *Spectral-MUSIC* are shown with dashed lines in Fig. 4. The solid line and the solid arrows represent the spectrum of the white noise with the two sinusoids. This is superimposed on the pseudo-spectra in the two graphs of Fig. 4. It is apparent that Method II produces a more accurate pseudo-spectrum than Method I.

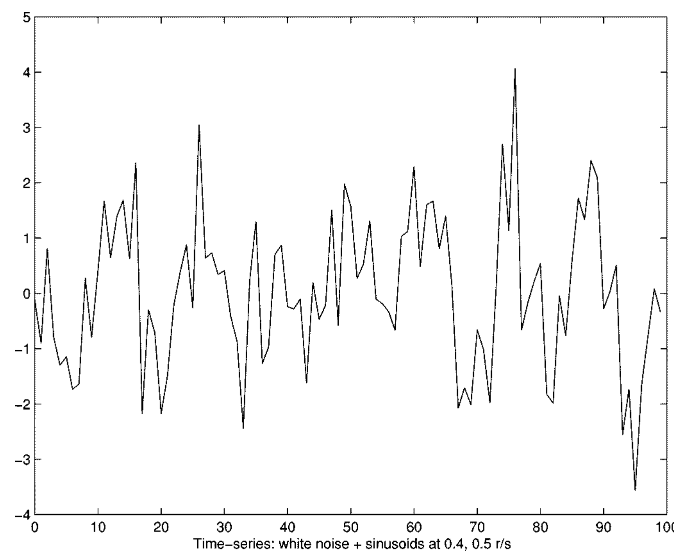


Fig. 3. Time series realization.

On the other hand, *Root-MUSIC* following these two different methods produced the estimates as in the following.

Method I	$(0.4901, 2.4096, 0.3257, 3.1416)$.
Method II	$(0.5020, 0.3763, 2.4223, 1.0701)$.

Thus, the two values that are selected as estimates for ω_1 and ω_2 are as follows.

Method I	$\omega_1 \sim 0.3257$ and $\omega_2 \sim 0.4901$.
Method II	$\omega_1 \sim 0.3763$ and $\omega_2 \sim 0.5020$.

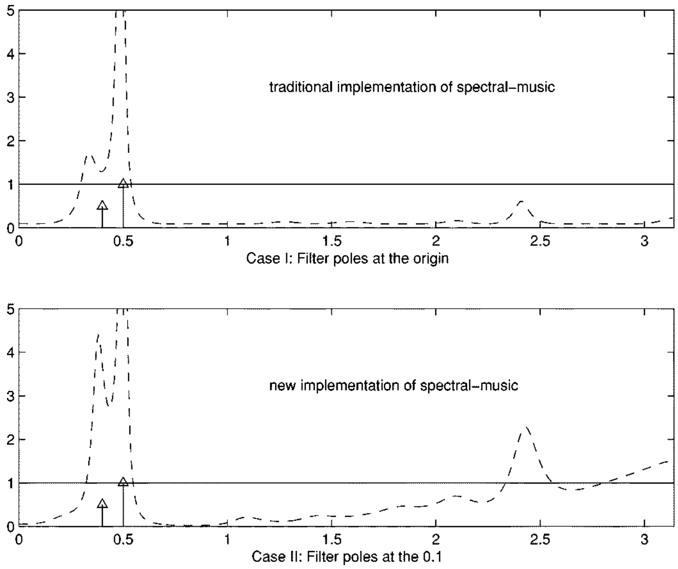


Fig. 4. Estimated pseudo-spectra.

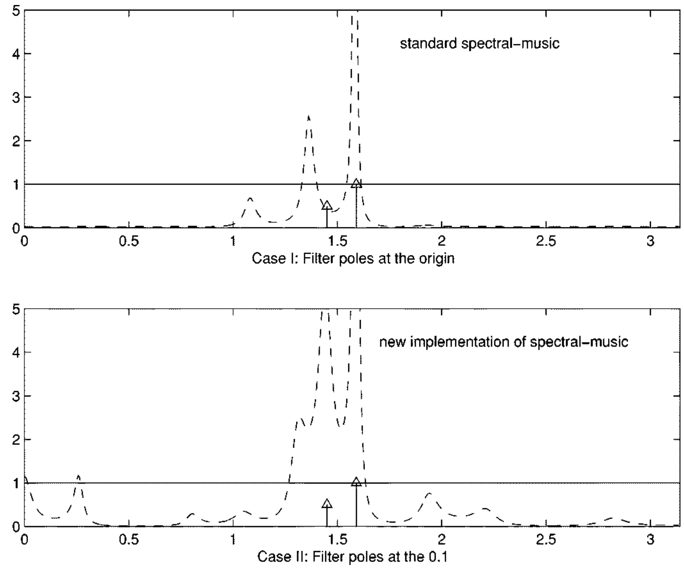


Fig. 5. Estimated pseudo-spectra.

B. Additional Simulation Runs

We tabulate the results produced of six additional runs obtained by *Root-MUSIC* implemented by the two methods. We only list the two frequencies that each method identified closest to the expected frequency range (i.e., closest to the interval [0.4, 0.5]) as follows.

Simulation	Method I	Method II
1	$0.5020, 0.8555$	$0.4042, 0.5124$
2	$0.4670, 0.5680$	$0.4550, 05549$
3	$0.4177, 0.5162$	$0.3863, 0.5100$
4	$0, \quad 0.4844$	$0.3849, 0.4989$
5	$0.0612, 0.4809$	$0.4689, 0.5596$
6	$0.4953, 1.0496$	$0.3890, 0.4994$

It is seen that Method II performs consistently better. In fact, the standard implementation of MUSIC in Method I often missed completely in identifying a second component in the vicinity of [0.4, 0.5].

C. Example 2: Bandpass Filter

Again, with

$$y_t = \nu_t + \rho_1 \sin(\omega_1 t + \phi_1) + \rho_2 \sin(\omega_2 t + \phi_2)$$

for $t = 0, 1, \cdots, N-1$, and all parameters as in Example 1, except for

$$\omega_1 = 1.45[\text{rad}] \quad \text{and} \quad \omega_2 = 1.59[\text{rad}]$$

we simulated and compared the performance of MUSIC in the two different implementations. Method I refers again to the standard implementation, whereas Method II now refers to using a bandpass filter with A in block-Jordan form as described in Section VIII-A. The parameters are chosen $r = 3$ and $\theta_0 = \pi/2$, and the frequency response is the same one shown in Fig. 2.

The results from a typical simulation, when we apply *Spectral-MUSIC* in the two implementations, are shown in Fig. 5. On the other hand, *Root-MUSIC* produced the following estimates.

Method I	$(1.5838, 1.3614, 1.0795, 1.9301)$
Method II	$(1.5953, 0.2612, 1.4458, 0.8007)$

Thus, again, the two values selected as estimates for ω_1 and ω_2 are

Method I	$\omega_1 \sim 1.3614$ and $\omega_2 \sim 1.5838.$
Method II	$\omega_1 \sim 1.4458$ and $\omega_2 \sim 1.5953.$

It is seen that the method proposed in this paper produced consistently more accurate estimates and that it has the potential to give better resolution and accuracy than the standard implementation of MUSIC. It is our expectation that more detailed analysis will reveal a quantitative assessment of the improvement in performance.

X. Recap

The effectiveness of MUSIC in locating sinusoids in noise can be enhanced by using state-covariance statistics of an *input-to-state* filter driven by the given time series. The filter can be chosen to amplify the effect of a specific harmonic interval where sinusoids are expected to reside. A general canonical decomposition theorem is presented for state-covariance matrices (Theorem 1). The method takes advantage of the canonical decomposition and applies it to the estimated covariance matrices to retrieve information on the frequency of embedded sinusoids. Simulation studies show that the method is capable of significantly higher resolution than traditional MUSIC. The current framework requires uniformly spaced samples of the process.

Acknowledgment

Motivation for the present work was gained by recent collaboration of the author with C. Byrnes and A. Lindquist.

REFERENCES

[1] N. I. Akhiezer, *The Classical Moment Problem*. New York: Oliver and Boyd, 1965.

[2] N. I. Akhiezer and M. G. Krein, *Some Questions in the Theory of Moments*. New York: Amer. Math. Soc., 1962.

[3] C. I. Byrnes, T. Georgiou, and A. Lindquist, "A generalized entropy criterion for Nevanlinna-Pick interpolation with degree constraint," *IEEE Trans. Automat. Contr.*, to be published.

[4] ——, "A new approach to spectral estimation: A tunable high-resolution spectral estimator," *IEEE Trans. Signal Processing*, to be published.

[5] C. I. Byrnes, A. Lindquist, S. V. Gusev, and A. S. Matveev, "A complete parametrization of all positive rational extensions of a covariance sequence," *IEEE Trans. Automat. Contr.*, vol. 40, pp. 1841–1857, 1995.

[6] C. I. Byrnes, S. V. Gusev, and A. Lindquist, "A convex optimization approach to the rational covariance extension problem," *SIAM J. Contr. Optimiz.*, to be published.

[7] C. I. Byrnes, H. J. Landau, and A. Lindquist, "On the well-posedness of the rational covariance extension problem," in *Current and Future Directions in Applied Mathematics*, M. Alber, B. Hu, and J. Rosenthal, Eds. Boston, MA: Birkhäuser, 1997, pp. 83–106.

[8] C. Carathéodory and L. Fejér, *Über den Zusammenghang der Extemen von harmonischen Funktionen mit ihren Koeffizienten und über den Picard-Landauschen Satz*. Palermo, Sicily: Rendiconti del Circolo Matematico di Palermo, 1911, vol. 32, pp. 218–239.

[9] T. T. Georgiou, "Realization of power spectra from partial covariance sequences," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-35, pp. 438–449, Apr. 1987.

[10] ——, "A topological approach to Nevanlinna-Pick interpolation," *SIAM J. Math. Anal.*, vol. 18, no. 5, pp. 1248–1260, Sept. 1987.

[11] ——, "The interpolation problem with a degree constraint," *IEEE Trans. Automat. Contr.*, vol. 44, pp. 631–635, Mar. 1999.

[12] ——, "Spectral estimation via selective harmonic amplification," *IEEE Trans. Automat. Contr.*, to be published.

[13] ——, "Spectral estimation via selective harmonic amplification: A canonical decomposition for state-covariances," in *Proc. IEEE Conf. Decision Contr.*, Phoenix, AZ, Dec. 1999.

[14] U. Grenander and G. Szegö, *Topelitz forms and Their Applications*. New York: Chelsea, 1958.

[15] S. Haykin, *Nonlinear Methods of Spectral Analysis*. Englewood Cliffs, NJ: Springer-Verlag, 1979.

[16] M. G. Krein and A. A. Nudel'man, *The Markov Moment Problem and Extremal Problems*. New York: Amer. Math. Soc., 1977.

[17] V. K. Madisetti and D. B. Williams, *The Digital Signal Processing Handbook*. Boca Raton, FL: CRC, 1998.

[18] V. F. Pisarenko, "The retrieval of harmonics from a covariance function," *Geophys. J. R. Astron. Soc.*, vol. 33, pp. 347–366, 1973.

[19] R. Schmidt, "Multiple emitter location and signal parameter estimation," in *Proc. RADC Spectrum Estimation Workshop*, 1979, pp. 243–258.

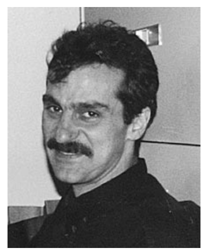
[20] C. Shankwitz and T. T. Georgiou, "On the envelope of spectral power of stochastic processes," *IEEE Trans. Inform. Theory*, vol. 40, pp. 584–588, Apr. 1994.

[21] J. A. Shohat and J. D. Tamarkin, *The Problem of Moments*, New York: Amer. Math. Soc., 1943.

[22] P. Stoica and R. Moses, *Introduction to Spectral Analysis*. Englewood Cliffs, NJ: Prentice-Hall, 1997.

[23] D. Sarason, "Generalized interpolation in $\mathcal{H}^\infty$," *Trans. Amer. Math. Soc.*, vol. 127, pp. 179–203, 1967.

[24] J. L. Walsh, *Interpolation and Approximation by Rational Functions in the Complex Domain*. New York: Amer. Math. Soc., 1956, vol. 20.



Tryphon T. Georgiou (SM'99–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 in 1979 and the Ph.D. degree from the University of Florida, Gainesville, in 1983.

He served on the Faculty of Florida Atlantic University, Boca Raton, from 1983 to 1986 and Iowa State University, Ames, from 1986 to 1989. Since August 1989, he has been with the University of Minnesota, Minneapolis, where he is currently a Professor of electrical and computer engineering. His research interests lie in the areas of control theory, signal processing, and applied mathematics.

Dr. Georgiou is a co-recipient, with Dr. M. C. Smith, of both the 1992 and the 1999 George Axelby Outstanding Paper Awards from the IEEE Control Systems Society. He has served as an Associate Editor for the IEEE TRANSACTIONS ON AUTOMATIC CONTROL, the *SIAM Journal on Control and Optimization*, and *Systems and Control Letters*.