Signal Estimation via Selective Harmonic Amplification: MUSIC, redux

Tryphon T. Georgiou

February 1999

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 a general state-covariance matrix. This is the state covariance of a stable linear filter driven by any process of unknown spectral content. Accordingly, the paper concludes with a modification of the MUSIC method. The new method allows filtering the time-series prior to computing covariance estimates. This step can be used to improve the S/N ratio by amplifying a selected frequency band where signals are expected to reside. Then the method capitalizes on the canonical decomposition to retrieve information on spectral lines.

Keywords

Canonical decomposition of a state-covariance matrix
Carathéodory-Fejér-Pisarenko harmonic decomposition
Carathéodory-decomposition of a covariance matrix

---

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

Dept. of Electrical and Computer Engineering, Univ. of Minnesota, MN 55455.

Tel: +1-612-625-3303, **Fax:** +1-612-625-4583, **E-mail:** georgiou@ece.umn.edu

1 Introduction

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

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

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

In this work we are concerned with the so-called MUSIC method. This method is intended for signals consisting of a number of sinusoidal periodic components in noise (ideally not far from being white). It is based on a canonical decomposition of a hermitian positive-definite Toeplitz matrix formed out of the partial covariance sequence of the process. The premise of MUSIC is that a certain suitably defined quantity (pseudo-spectrum) will inherit the peaks of the actual power spectrum and its poles will be close to the actual sinusoidal modes. In many cases this appears to work quite effectively with minimal computation cost and thus the method has become quite popular (e.g., see [6]).

The purpose of the present note is to give an analogous canonical decomposition of a general state-covariance matrix based on knowledge of the filter parameters (Theorem 1) and no particular assumption on the input statistics. The filter is required to be stable and its output can be thought of as the complete state-vector. More specifically, it is shown that the state-covariance can always be represented uniquely in a canonical form consistent with the possibility that the input was the superposition of sinusoids in white noise. Then we focus on the problem of estimating lines in the power spectrum of a time series. We subscribe to the guidelines and premises of MUSIC and seek an approximate location of the sinusoids by suitable decomposition of a sampled state-covariance matrix. We utilize data obtained after passing the time-series through the aforementioned input-to-state filter in order to amplify a pre-selected frequency band, and we then capitalize on the structure-theorem for state-covariance matrices in order to identify potential frequencies for such sinusoidal components.

Filtering an observed time-series prior to computing covariance estimates, has the effect of amplifying selected bands of the spectrum, and thereby, has the potential to improve resolution. This fact was first noted in Byrnes, Georgiou, and Lindquist [2, 3] where suitably defined filter banks were introduced in order to obtain interpolation conditions for the sought spectrum of the time-series. Byrnes, Georgiou, and Lindquist [2, 3]) went on to present a complete theory for parametrizing and constructing all regular spectra of *bounded complexity* which are consistent with the observed data. The theory in [2, 3] encompasses the Maximum Entropy method and Nevanlinna-Pick interpolation as special cases, while it allows for a general senario where the design of the processing apparatus (i.e., of the filter bank) and the choice of a certain free parameter can be tailored to yield high resolution and reliability over a specific harmonic aperture. While the input-to-state filtering used in the present work relates to, and is in fact a generalization of, the filter banks considered in [2, 3], the signal retrieval rests on the aforementioned canonical decomposition theorem.

Section 2 gives a brief outline of the main technical results underlying traditional MUSIC . Sections 3 and 4, develop technical background related to the algebraic and analytical structure of the states of a linear system. Sections 5 through 7 deal with the structure of state-covariance matrices. The main result (Theorem 1) on the decomposition of state-covariance matrices is given in Section 7. This result suggests accordingly a generalization of MUSIC which is discussed in Section 8. Section 9 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 S/N ratio, and the decomposition theorem suggest ways to identify spectral lines. The new method compares quite favorably to traditional MUSIC .

2 Traditional MUSIC

The key idea behind the methodology of MUSIC is a certain mathematical fact about Toeplitz matrices. This fact can be traced in the work of Carathéodory-Fejér in the early part of the century (see [5, page 151]. In modern times analogous results were re-discovered, and their relevance in signal processing recognized, by Pisarenko, Schmidt, and others (see [6, page 155]).

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

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

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

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

$\theta_\ell \neq \theta_k$ 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

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

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

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

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

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

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

$$d(e^{j\theta}, e^{-j\theta}) := \begin{bmatrix} 1 & e^{-j\theta} & \dots & e^{-j(n-1)\theta} \end{bmatrix} U_{m+1:n} U_{m+1:n}^* \begin{bmatrix} 1 \\ e^{j\theta} \\ \vdots \\ e^{j(n-1)\theta} \end{bmatrix} \tag{1}$$

has m roots precisely at $\{\theta_1, \dots, \theta_m\}$. 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 [6, page 157].

The significance in the context of signal processing stems from the fact that if c_k ($k = 0, 1, \dots$) 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 coresponding 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 [6] and also Section 7 below).

3 Input-to-state filters

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

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

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,

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

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 is

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

Such a system will be refered to as a *polyfilter*, or a *filter bank*, or an *input-to-state filter.*

The use of such filter-banks was initiated in [2, 3]. They were used to process a stochastic process $\mathbf{y}_t$ and the output made available for statistical measurements (i.e., estimation of covariances). The output covariances impose interpolation constraints for the sought power spectrum of $\mathbf{y}_t$. It was further noted that suitable selection of such interpolation conditions improves resolution. This was a starting point for a general theory for tunable high resolution estimators, based on new results in analytic interpolation theory (see [2, 3]). Here we use a similar idea to generate statistical estimates.

To motivate and clarify the terminology *filter bank* we recall that (as in [2, 3]) for the special case where A is a *diagonal* matrix

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

it follows that

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

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

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

This case coresponds to

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

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 filter-bank 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

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

The expression $b^* f(A^*) e_i^*$ depends either on the value of $f(\lambda)$ at $\bar{p}_i$ as in the case of (3), or on the value of the i-th 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. Q.E.D.

4 The co-invariant subspace $\mathcal{K}$

The subspace of H_2-functions $f(\lambda)$ which 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 (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

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

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

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

Proof. The statement of the lemma follows from:

(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 $G(\lambda)$, and

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

Claim (i) follows from the controllability of (A, b). Claim (ii) is standard [7]. Claim (iii) is easily verified since

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

Q.E.D.

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 which are defined on the spectrum of A, we have that $f(S^*) = w(S^*) \Leftrightarrow f(A) = w(A)$.

5 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 (nonnegative) 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 the coresponding 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 the coresponding input vector, then the covariance of the state-space process is in the usual Pick-matrix form

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

In this case, the values w_i for $i = 1, \ldots, n$ relate to the covariance of the individual entries of the vector-process $\mathbf{x}_t$ (see [3] for details). In general, the state-covariance P is characterized by

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

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

6 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 typically one wants to identify. The residues ρ_k represent the energy density contained in the coresponding sinusoidal components of $\mathbf{y}_t$ while $\dot{\sigma}_y(\theta)$ is the spectral density function which is defined a.e. and is non-negative.

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

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

and the quadratic form

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

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 $H(e^{j\theta_k}) = 0$ for $k = 1, \ldots, m$. Note that $H(\lambda)$ is a rational function which 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.,

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

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 $k = 1, \ldots, m$.

Proof. The "if" part is obvious. The "only if" was shown by the arguments leading to the proposition. Q.E.D.

We now study the corespondence between the elements in the above decomposition and the zero-eigenvectors of P. So 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 corespondence 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

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

with U unitary and $\sigma_i > 0$ ($i = 1, \ldots, m$), 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 $k = 1, \ldots, m$, 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 $\{\theta_1, \ldots, \theta_m\}$. It should be noted that, just as in the case of Section 2, $d(\lambda, \lambda^{-1})$ 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 $k = 0, 1, \ldots, m$, 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.

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

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

Clearly det $M(\lambda)$ is an element of $\mathcal{K}$ since all nonconstant 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 $\{e^{j\theta_1}, \ldots, e^{j\theta_m}\}$. Thus $M(e^{j\theta_0})$ is nonsingular. Q.E.D.

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 $e^{j\theta_k}$ $(k = 1, \ldots, m)$ 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 $e^{j\theta_k}$ $(k = 1, \ldots, m)$.

Proof. It was shown by the argument leading to the proposition. Q.E.D.

7 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 coresponding 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 which 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) holds always. even if $\mathbf{y}_t$ has an arbitrary power spectrum.

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 $d\sigma(\theta)$, so as 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 (Commutant Lifting Theorem [7, Theorem 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 $\phi(\lambda)$ which 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 [7, Theorem 1] except that it is phrased in terms of a positive-real property as opposed to contractiveness. To make the corespondence 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 $\phi(\lambda)$ 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. Q.E.D.

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 $G(\lambda)$ with 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 nonnegative measure $d\sigma(\theta)$.

Proof. Power spectra, like $d\sigma_y(\theta)$, are characterized by the property that $\sigma_y(\theta)$ is a non-decreasing function of bounded variation, or that $d\sigma_y(\theta)$ is a bounded nonnegative measure. It is a well known fact (e.g., see [1, page 91]) 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 nonnegative measure.

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

$$T := \phi(S) \; : \; x(\lambda) \to \mathbf{\Pi}_{\mathcal{K}} \phi(\lambda) x(\lambda),$$

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^* coresponds to the matrix A while T^* coresponds to a matrix W, i.e.,

$$T^* \; : \; cG(\lambda) \to cWG(\lambda).$$

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

$$\frac{1}{2} \left(\langle cG(\lambda), T^* cG(\lambda) \rangle + \langle T^* cG(\lambda), cG(\lambda) \rangle \right) = \frac{1}{2} \left\{ c \left(EW^* + WE \right) c^* \right\}.$$

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

$$P = (EW^* + WE).$$

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

$$P_0 = (EW_0^* + W_0 E)$$

with $W_0 := W - \frac{\rho_0}{2} I$ and that it is nonnegative definite. Then define the operator T_0 on $\mathcal{K}$ through its adjoint by

$$T_0^* \; : \; cG(\lambda) \to cW_0 G(\lambda).$$

Since W_0 differs from W only by $\frac{\rho_0}{2} I$, 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). Q.E.D.

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 nonnegative measure $d\sigma_y(\theta)$. There is a unique decomposition of P in the form*

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

where $G(\lambda) = (I - \lambda A)^{-1} b$, and

(i) $E := \frac{1}{2\pi} \int_{-\pi}^{\pi} G(e^{j\theta}) G(e^{j\theta})^* d\theta$ *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) $m = n - \mathrm{rank}(P - \rho_0 E)$, *and*

(iv) $e^{j\theta_k}$ *for* $k = 1, \ldots, m$ *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*

$$\text{diag}\{\rho_1, \ldots, \rho_m\} = V^{-1} \left(P - \rho_0 E\right) \left(V^{-1}\right)^*$$

where

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

Proof. Follows from Lemma 5, Propositions 1 and 2, and the discussion leading to Lemmas 4 and 5. Q.E.D.

We wish to emphasize that the theorem makes no particular assumption on the spectrum of the input process. Yet the theorem 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 which coresponds A, b as in (4) reduces to the Carathéodory-Fejér results and is the basis of the method introduced by Pisarenko (cf. [5, page 151] and [6, pages 155, 170]).

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 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, one 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 non-trivial 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 $(c_0, c_1, c_2, \ldots)$. 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

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

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 explicitely 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 valied in general only for input-to-state filters of the form $G(\lambda) = (I - \lambda A)^{-1} b$. $\square$

8 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.

8.1 Choice of A, b

The filter parameters can be specified so that the frequency response

$$f_{A,b}(\theta) := \|(I - e^{j\theta}A)^{-1}b\|$$

has pass-band over the range of frequencies of interest. I.e., the range of frequencies where the sinusoidal components are expected to reside. This can be done, e.g., by taking A is a Jordan form

$$A = \begin{bmatrix} r & 1 & 0 & \dots & 0 \\ 0 & r & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & \dots & r \end{bmatrix} \tag{24}$$

and r in the interval $0 < r < 1$ or, in $-1 < r < 0$, depending on whether a low-pass or a high-pass characteristic is desired, respectively. The only constraint on the b vector is the one of controllability. Thus, $b = [0, \dots, 0, 1]'$ can be used.

Similarly, a block Jordan structure

$$A = \begin{bmatrix} R & I & 0 & \dots & 0 \\ 0 & R & I & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & I \\ 0 & 0 & 0 & \dots & R \end{bmatrix} \tag{25}$$

where

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

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

The normalized frequency response

$$\frac{f_{A,b}(\theta)}{\max_\theta\{f_{A,b}(\theta)\}}$$

is shown for two representative cases along with their pole pattern in Figures 1 and 2. The first is generated according to (24) with $r = 0.1$ and the second according to (25) with $r = .3$ and $\theta_0 = \pi/2$. In both cases A was chosen 30×30. These will be used and referred to in the examples below.

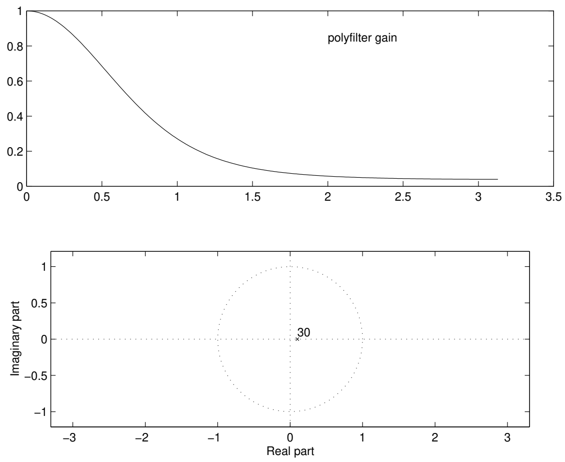

Figure 1: Frequency response and filter poles for a lowpass polyfilter

8.2 State-covariance Estimates

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

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

available, we generate the state vector-process x_k from the equation $x_k = Ax_{k-1} + by_k$ for $k = 0, \ldots, N-1$. (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

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

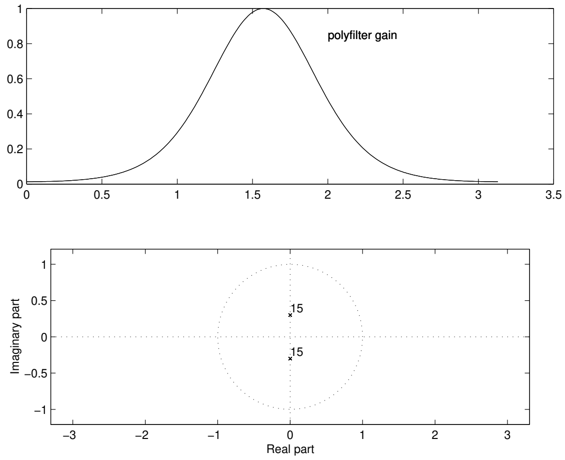

Figure 2: Frequency response and filter poles for a bandpass polyfilter

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 which gives a level of stationarity of state time-series after $t \geq \ell$.

8.3 New MUSIC

(8.3.1) We need to specify m. This is typically taken slightly larger than the expected number of sinusoidal components in the spectrum of $\mathbf{y}_t$, say m_0. The method produces m candidate frequencies, and then a subset of m_0 frequencies which are closest to the expected range are selected.

(8.3.2) To this end, after selecting m we construct $U_{m+1:n}$ to span the (right) eigensubspace coresponding to the smallest $n-m$ eigenvalues of the matrix pencil $P - \rho E$. This can be done by standard eigenvalue decomposition of the pencil.

(8.3.3) Finally, the following two option can be followed:

ROOT MUSIC: *As an estimate for the frequencies θ_k $(k = 1, \ldots, m)$ 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 pseudospectrum*

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

has its peaks.

Remark 3. 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 pseudospectrum). This follows from Theorem 1, since, besides ρ_0 and θ_k for $k = 1, \ldots, m$, we can also identify ρ_k for $k = 1, \ldots, m$ from (23). Generically, this is the case only when $m = n - 1$. This coresponds to the method of Pisarenko (cf. [6, page 155]).

9 Case studies

We now compare MUSIC in the standard implementation (METHOD I) vs. the implementation suggested in Section 8 (METHOD II).

Example 1: low pass filter

We simulated numerically a random signal

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

where $N = 100$, $\omega_1 = .4$[rad], $\omega_2 = .5$[rad], $\rho_1 = .5$, and $\rho_2 = 1$, while ν_t $(t = 0, \ldots, N-1)$ 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 [6, pages 157-158]. This coresponds 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 8. In this case, since the sinusoidal frequencies are (known to be) at the low frequency part of the spectrum, we make use of a low pass filter. The parameters for A, b were chosen according to (24) with $r = .1$. The normalized frequency response is plotted in Figure 1. The filter gives a $-6dB$ attenuation at about $\omega \sim 0.7$. As a consequence, estimates outside this bandwidth should be considered less reliable, while improved accuracy may be expected within the bandwidth.

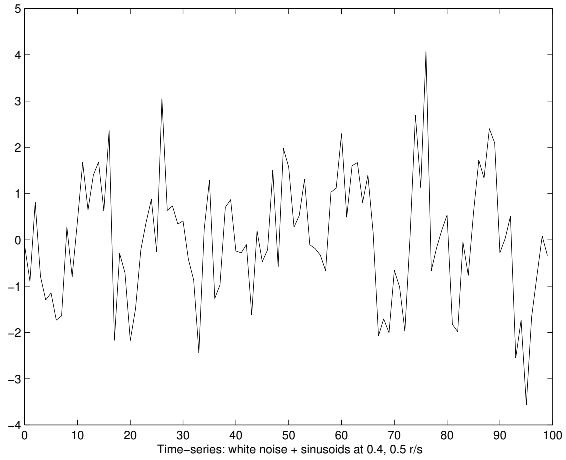

Figure 3: Time series realization

The time-series of one simulation is shown in Figure 3. The (normalized) pseudospectra produced by *Spectral-MUSIC* are shown with dashed lines in Figure 4. Solid line and solid arrows represent the spectrum of the white noise and of the two sinusoids respectively, which is superimposed on the two pseudospectra. It is apparent that METHOD II produces a more accurate pseudospectrum than METHOD II.

On the other hand, *Root-MUSIC* following these two different methods produced the estimates given below:

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

For METHOD I: $\omega_1 \sim 0.3257$ and $\omega_2 \sim 0.4901$.

For METHOD II: $\omega_1 \sim 0.3763$ and $\omega_2 \sim 0.5020$.

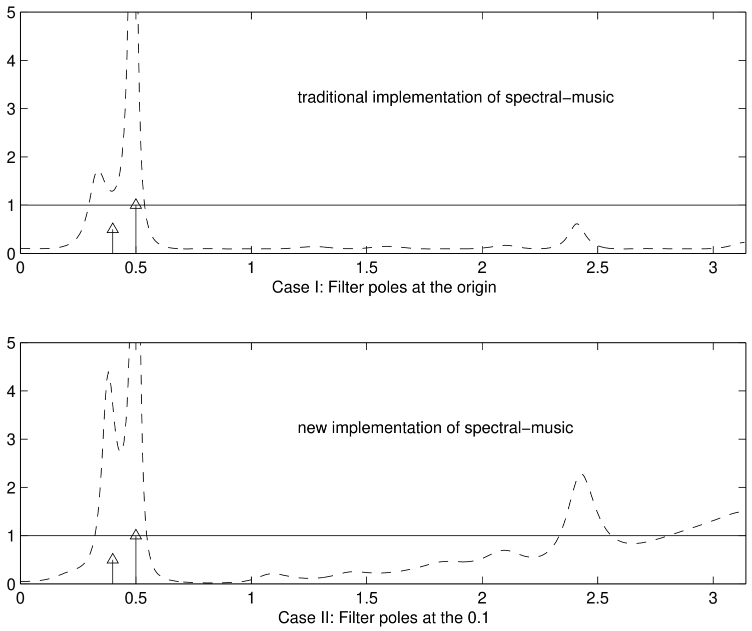

Figure 4: Estimated pseudo-spectra

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 $[.4, .5]$).

Simulation	METHOD I		METHOD II	
1	0.5020,	0.8555	0.4042,	0.5124
2	0.4670,	0.5680	0.4550,	0.5549
3	0.4177,	0.5162	0.3863,	0.5100
4	0,	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 does consistently better. In fact, the standard implementation of MUSIC in METHOD I, often missed completely in identifying a second component in the vicinity of $[.4, .5]$.

Example 2: bandpass filter

With

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

and all parameters as in Example 1, except for

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

we simulated again 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 8.1. The parameters are chosen $r = 3$ and $\theta_0 = \pi/2$ and the frequency response is the same one shown in Figure 2.

The results from a typical simulation, when we apply *Spectral-MUSIC* in the two implementations, are shown in Figure 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

For METHOD I: $\omega_1 \sim 1.3614$ and $\omega_2 \sim 1.5838$.

For METHOD II: $\omega_1 \sim 1.4458$ and $\omega_2 \sim 1.5953$.

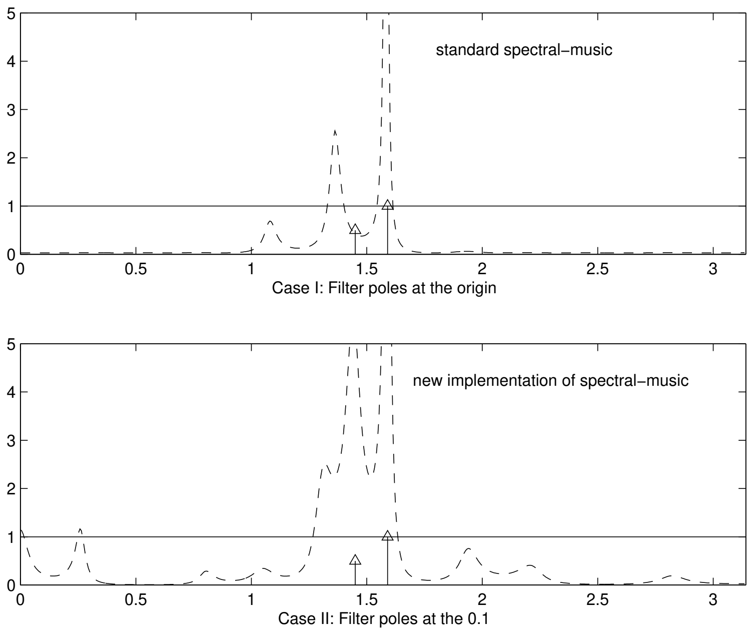

Figure 5: Estimated pseudo-spectra

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 detail analysis will reveal a quantitative assessment of the improvement in performance.

10 Re-cap

The effectiveness of MUSIC in locating sinusoids in noise can be enhanced by filtering the time-series prior to computing covariance statistics. This step effectively reduces S/N ratio by amplifying the harmonic aperture where the 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 so as to retrieve information on the frequency of enbedded sinusoids. Simulation studies show that the method has significantly higher resolution than traditional MUSIC .

The idea of filtering prior to computing statistics originates in Byrnes, Georgiou and Lindquist [2, 3], where the use of similar polyfilter banks was introduced. The resolution of MUSIC , with or without harmonic selectivity, is theoretically less than what is possible with the general tunable high resolution estimators in [2, 3]. This is due to the fact that MUSIC is intended for rather special conditions (i.e., for spectra of sinusoids in white noise). However, the construction of pseudospectra and the strategy prescribed in Root-MUSIC comes with a minimal computational cost and is often quite effective.

Acknowledgment. This work was motivated by the recent collaboration with Chris Byrnes and Anders Linquist [2, 3].

References

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

[2] C. I. Byrnes, 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*, preprint, submitted to *IEEE Trans. on Aut. Control.*

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

[4] C. I. Byrnes, T. Georgiou, and A. Lindquist, *Generalized interpolation in H^∞: Solutions of bounded complexity*, in preparation.

[5] U. Grenander and G. Szegö, **Toeplitz forms and their Applications**, Chelsea, 1958.

[6] P. Stoica and R. Moses, **Introduction to Spectral Analysis**, Prentice Hall, 1997.

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