The structure of state covariances and its relation to the power spectrum of the input

Tryphon T. Georgiou

Abstract—
We study the relationship between power spectra of stationary stochastic inputs to a linear filter and the corresponding state covariances. We identify the structure of positive semidefinite matrices that qualify as state covariances of the filter. This structure is best revealed by a rank condition pertaining to the solvability of a linear equation involving the state covariance and the system matrices. We then characterize all input power spectra consistent with any specific state covariance. The parametrization of input spectra is achieved through a relation to solutions of an analytic interpolation problem which is analogous, but not equivalent, to a matricial Nehari problem.

Keywords: Linear filters, spectral analysis.

I. Introduction

CONSIDER a finite-dimensional linear filter which is driven by a multivariable stationary stochastic process, and suppose that nothing else is known about the input process. In this context, we are interested in the following two basic questions. Firstly, is there anything we can say about the structure of the covariance of the state vector which is independent of the specific input? In particular, how can tell whether a given positive semidefinite matrix qualifies as the state covariance of the filter for a suitable input process? Second, assuming knowledge of the state-covariance matrix, what are all admissible input power spectra which are consistent with the particular state-covariance?

The material presented herein generalizes the work in [7] which deals with scalar inputs. The required theory is significantly more complex since we deal with a multivariable framework, yet certain of the results are derived in a simpler and more definitive form. In particular, the answer to our first question is given by a rank condition involving the state-covariance and the state matrices of the filter (Theorem 1 below), as opposed to [7, Theorem 1] which requires solvability of a linear equation by a suitable commutator of the filter state-matrix. The underlying analytic interpolation problem is substantially more complex and in this case, as it appears, there is no direct translation into standard H_∞-Sarason-type interplation, as the case was in the description given in [7].

The exposition in this paper focuses on the theory needed for answering the above two basic questions. The premise of these questions is to view spectral analysis and estimation as an inverse problem, and characterize the totality of spectra which are consistent with measured statistics. Exposition and discussion on specific techniques for analysis of scalar time-series, which spring out of this viewpoint, can be found in [1,6–8]. Numerical algorithms and applications for the multivariable case will be presented in a subsequent publication.

Department of Electrical and Computer Engineering, University of Minnesota, MN 55455

E-mail: tryphon@ece.umn.edu

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

II. Problem formulation

Consider the linear discrete-time state equations

$$x_k = Ax_{k-1} + Bu_k, \text{ for } k \in \mathbb{Z}, \tag{1}$$

where $x_k \in \mathbb{C}^n$, $u_k \in \mathbb{C}^m$, $A \in \mathbb{C}^{n\times n}$, $B \in \mathbb{C}^{n\times m}$, (A,B) is a controllable pair, and the eigenvalues of A lie in the open unit disk of the complex plane. Let $\{u_k : k \in \mathbb{Z}\}$ be a zero-mean stationary stochastic process with power spectrum $d\mu(\theta)$; i.e., $d\mu(\theta)$, for $\theta \in (-\pi, \pi]$, is a non-negative matrix-valued measure having as Fourier-Stieljes coefficients the sequence of covariance lags of u_k:

$$R_\ell = \mathcal{E}\{u_{k+\ell}u_k^*\} = \int_{-\pi}^{\pi} e^{-j\ell\theta}\frac{d\mu(\theta)}{2\pi} \tag{2}$$

where $\mathcal{E}\{\cdot\}$ denotes the expectation operator. Then, *under stationarity conditions,* the state covariance

$$\Sigma := \mathcal{E}\{x_k x_k^*\}$$

can be expressed in the form of the integral (cf. [10, Ch. 6])

$$\Sigma = \int_{-\pi}^{\pi}\left(G(e^{j\theta})\frac{d\mu(\theta)}{2\pi}G(e^{j\theta})^*\right) \tag{3}$$

where

$$G(\lambda) := (I - \lambda A)^{-1}B$$

is the transfer function of system (1). Note that we use λ to denote the transform of the delay operator and therefore $G(\lambda)$ is analytic in the unit disc of the complex plane. Our goal is to study the correspondence in equation (3) between Σ and μ. More specifically, given knowledge of the system parameters (A,B), we pose and answer the following two questions:

Question 1: How can we tell that a given non-negative definite matrix qualifies as a state covariance of (1)?
Equivalently, given $G(\lambda)$ as before and a non-negative definite matrix Σ, how can we tell that (3) holds for a suitable non-negative definite measure $d\mu(\theta)$?

Question 2: Given a state covariance Σ of system (1), parametrize the set of power spectra for the input process which are consistent with Σ.
Equivalently, given Σ for which there exists a non-negative measure $d\mu(\theta)$ satisfying (3), it is required to parametrize all such non-negative $d\mu$'s.

Section III provides answers to both questions, while Sections V and IV are devoted to the theory needed for the derivation of the results.

Notation and preliminaries

Throughout the paper we use the notation $X \geq Y$ and $X > Y$ for Hermitian matrices X, Y to indicate that the difference $X - Y$ is nonnegative definite and positive definite, respectively. We denote by $0, I$ the zero matrix and the identity matrix, respectively. Their size will be implicit from the context. We use $\mathbb{D}$ to denote the open unit disk in the complex plane.

We begin by recalling certain facts concerning matrix-valued measures and integration (e.g., see [2, 10]). Denote by $\mathcal{M}$ the set of finite Hermitian non-negative matrix-valued measures on $[-\pi, \pi]$. Such measures are defined by bounded and nondecreasing Hermitian matrix-valued functions $\mu(\theta)$ with $\theta \in [-\pi, \pi]$, i.e., $\mu(\beta) \geq \mu(\alpha)$ whenever $\beta \geq \alpha$. Note that $d\mu$ corresponds to μ in a one-to-one fashion, except for an arbitrary additive constant. It is known that $\mu(\theta)$ has at most countably many discontinuities, that the limits $\mu(\theta^-)$ and $\mu(\theta^+)$ exist everywhere, and that the derivative $d\mu(\theta)/d\theta$ exists almost everywhere. Further, the measure decomposes into $d\mu_a + d\mu_s$ where μ_a is the integral of the derivative and is *absolutely continuous* while μ_s is a *singular measure* with derivative 0 almost everywhere. If $d\mu$ represents the power spectrum of a multivariable stochastic process, then $d\mu_s$ corresponds to a deterministic component. Whether $d\mu_a$ corresponds to a nondeterministic component as well depends on the integrability of $\log(\det(d\mu_a(\theta)/d\theta))$; this is the content of the celebrated Szegö-Kolmogorov-Whittle theorem (e.g., see [2, Theorem 18], [10, Ch. 10]).

There is a natural correspondence between elements in $\mathcal{M}$ and "positive-real" matrix-valued functions (modulo a skew-Hermitian constant); this is the content of Riesz-Herglotz's theorem (see [2, p. 150]). More specifically, if $d\mu \in \mathcal{M}$ and jc an arbitrary skew-Hermitian constant matrix, then the matrix-valued function

$$F(\lambda) = \int_{-\pi}^{\pi} \left(\frac{1 + \lambda e^{j\ell\theta}}{1 - \lambda e^{j\ell\theta}}\right) \frac{d\mu(\theta)}{2\pi} + jc, \tag{4}$$

is analytic in $\mathbb{D}$ with non-negative definite real part. The class of functions $F(\lambda)$ with the above property will be denoted by $\mathcal{F}$. Conversely, if $F(\lambda) \in \mathcal{F}$ then the Cesaro means of the truncated Fourier series of the $\Re\{F(\lambda)\}$ converge to a measure $\mu \in \mathcal{M}$ in the weak-star topology [9, Chapter 2]. One can also use

$$d\mu(\theta)/d\theta = \lim_{r \nearrow 1} \Re\{F(re^{j\theta})\} \tag{5}$$

which holds almost everywhere whereas at points of discontinuity

$$\frac{1}{2}\left(\mu(\theta^+) + \mu(\theta^-)\right) = \lim_{r \nearrow 1} \int_{\theta^-}^{\theta^+} \Re\{F(re^{j\theta})\} d\theta + \mu_0 \tag{6}$$

with μ_0 an arbitrary constant matrix. These facts are classical in the scalar case [9, Chapter 2] and easy to extend to the matrix case by considering scalar functions $x^* F(\lambda) x$ for appropriate constant vectors x [2, 4].

Finally, we use the standard notation $\mathcal{H}_2$ for the Hardy space of analytic functions in $\mathbb{D}$ with square integrable boundary values, and $(\mathcal{H}_2)^\perp$ for its orthogonal complement. These are precisely the spaces of Fourier transforms of square integrable signals with support on non-negative and negative integers, respectively. Finally, for any function $f(\lambda)$ with Fourier series $\sum_k f_k \lambda^k$, we use the notation $f(\lambda)^*$ to also denote the "para-Hermitian conjugate" $f^*(\lambda^{-1}) = \sum_k f_k^* \lambda^{-k}$ where in the last expressions the "conjugate-traspose" operation applies to each of the coefficients of the series.

III. Main results

We now provide answers to the two questions posed in the introduction. In Theorem 1 we characterize the propery of a matrix Σ being a state covariance as a rank condition in the problem data Σ, A, B. The necessity of such a condition is not surprising because, aside from positivity requirements, Σ must lie in the range of an operator, defined by the right hand side of (3), which is linear in μ. However, the sufficiency is deeper and amounts to an existence result in analytic interpolation. The proof of the sufficiency part is constructive and is tightly connected to Theorem 2. This last theorem provides a parametrization of all admissible input power spectra which are consistent with Σ in terms of $\mathcal{F}$-solutions to an associated analytic interpolation problem.

Theorem 1: Let $B \in \mathbb{C}^{n \times m}$, $A \in \mathbb{C}^{n \times n}$ be such that (A, B) is a controllable pair with A having its eigenvalues in $\mathbb{D}$, and let $\Sigma \in \mathbb{C}^{n \times n}$ be a Hermitian positive semidefinite matrix. The matrix Σ is the stationary state covariance of system (1) for a suitable input process if and only if

$$\text{rank}\begin{bmatrix} \Sigma - A\Sigma A^* & B \\ B^* & 0 \end{bmatrix} = \text{rank}\begin{bmatrix} 0 & B \\ B^* & 0 \end{bmatrix}. \tag{7}$$

If B is full column rank then the right hand side of (7) is simply $2m$. In fact, the general case can be easily reduced to this case where B is full column rank—if $B = [B_0,\ 0]V_0$ with V_0 unitary and B_0 full column rank, then simply replace B with B_0 for an equivalent problem statement. Thus, whenever helpful, we will use the simplifying assumption that B is full column rank.

Condition (7) of the theorem is equivalent to the solvability of a certain linear matrix equation. This is given below as equation (9) and its solution specifies the data of the underlying analytic interpolation problem which will be introduced in Theorem 2.

Proposition 1: Let $B \in \mathbb{C}^{n\times m}$ and $\Delta = \Delta^* \in \mathbb{C}^{n\times n}$. Then,

$$\operatorname{rank}\begin{bmatrix} \Delta & B \\ B^* & 0 \end{bmatrix} = \operatorname{rank}\begin{bmatrix} 0 & B \\ B^* & 0 \end{bmatrix}, \tag{8}$$

if and only if the matrix equation

$$\Delta = BH + H^*B^* \tag{9}$$

has a solution $H \in \mathbb{C}^{m\times n}$.

Remark 1: If the solution set of (9) is not empty and H_0 is a particular solution, then the set of solutions includes $H = H_0 + jcB^*$ with jc a skew-Hermitian matrix. For instance, a solution H can always be selected so that e.g., HB is Hermitian.

In Theorem 1, the case where $m = 1$, i.e., where u_k is a scalar input and B is a column vector, has been analyzed in [6, 7]. It is interesting to recall the relevant result from [7, Theorem 1] and derive it as corollary to Theorem 1. This is stated below.

Corollary 1: (*Case $m = 1$ — [7, Theorem 1].*) Let $B \in \mathbb{C}^{n\times 1}$, $A \in \mathbb{C}^{n\times n}$ be such that (A, B) is a controllable pair with A having its eigenvalues in $\mathbb{D}$, and let $\Sigma \in \mathbb{C}^{n\times n}$ be a Hermitian positive semidefinite matrix. The matrix Σ is the stationary state covariance of system (1) for a suitable input process if and only if it is of the form

$$\Sigma = WE + EW^* \tag{10}$$

with E the unique solution to the equation $E - AEA^* = BB^*$ and W a matrix which commutes with A.

The following theorem establishes the correspondence between input power spectra which are consistent with a given state covariance Σ and solutions to an analytic interpolation problem. This problem is analogous, but *not equivalent*, to a matrix-valued Nehari problem (cf. [5]).

Theorem 2: Let $B \in \mathbb{C}^{n\times m}$, $A, \Sigma \in \mathbb{C}^{n\times n}$ be such that (A, B) is a controllable pair, the eigenvalues of A are all in $\mathbb{D}$, the rank of B is equal to m, and Σ is a Hermitian positive semidefinite matrix which satisfies (7). The following hold:
(i) The equation

$$\Sigma - A\Sigma A^* = BH + H^*B^* \tag{11}$$

has a solution $H \in \mathbb{C}^{m\times n}$.
(ii) The set of nonnegative matrix-valued measures $d\mu$ which satisfy (3) is non-empty and is in bijective correspondence via (4) to the set of positive real matrix-valued functions $F(\lambda) \in \mathcal{F}$ of the following form:

$$F(\lambda) = F_0(\lambda) + Q(\lambda)V(\lambda) \tag{12}$$

where $Q(\lambda)$ is analytic in $\mathbb{D}$,

$$F_0(\lambda) = H(I - \lambda A)^{-1}B \tag{13}$$
$$V(\lambda) = D + C\lambda(I - \lambda A)^{-1}B, \tag{14}$$

and C, D chosen so that

$$\begin{bmatrix} E^{-\frac{1}{2}}AE^{\frac{1}{2}} & E^{-\frac{1}{2}}B \\ CE^{\frac{1}{2}} & D \end{bmatrix}$$

is a unitary matrix, with E being the unique solution of the Lyapunov equation $E - AEA^* = BB^*$.
(iii) The set of $\mathcal{F}$-functions in (ii) can be explicitely parametrized via a sequence of "one-step-extensions" specified by a sequence of suitable contractive matrix-parameters.

Statement (i) is of course a direct consequence of Proposition 1 but helps in setting up the corresponce between input spectra consistent with Σ and the interpolation problem specified in statement (ii). Aside from the correspondence between the two sets, i.e., of input spectra and of solutions to a "positive-real Nehari" problem, statement (ii) also includes the fact that Σ is in essence the "Pick matrix" corresponding to this interpolation problem. Hence, the solution sets are nonempty when $\Sigma \geq 0$, i.e., Σ is a state covariance. The parametrization claimed in (iii) is in the style of classical analytic interpolation and is explained in detail in the course of the proof.

IV. Proofs of Proposition 1, Theorem 1, and Corollary 1

Proposition 1 expresses the solvability of a linear matrix equation as a rank condition and is interesting in its own right. It will be shown first. Then we give the proof Theorem 1. The necessity part is completed in this section. However, the proof of the sufficiency part relies on Theorem 2 which will be established in the next section (Section V).

Proof: [Proposition 1] The sufficiency is immediate by noting that if (9) has a solution H then the matrices

$$\begin{bmatrix} \Delta & B \\ B^* & 0 \end{bmatrix} \tag{15}$$

and

$$\begin{bmatrix} 0 & B \\ B^* & 0 \end{bmatrix} \tag{16}$$

are congruent since they are related by

$$\begin{bmatrix} I & H^* \\ 0 & I \end{bmatrix}\begin{bmatrix} \Delta & B \\ B^* & 0 \end{bmatrix}\begin{bmatrix} I & 0 \\ H & I \end{bmatrix} = \begin{bmatrix} 0 & B \\ B^* & 0 \end{bmatrix}.$$

The converse can be established using Roth's theorem [12]. Roth's theorem gives conditions for the solvability of $\Delta = BX + YC$ where Δ, B, C are known and have compatible dimensions. When specialized to our case where $C = B^*$ it asserts that (8) implies the existence of matrices X, Y of suitable size so that $\Delta = BX + YB^*$. Our proof is complete by selecting $H = (X + Y^*)/2$.

However, a simple and direct proof, without reference to Roth's result (or techniques used in [12]) can be constructed along the following lines. Through congruence transformation the matrix (15) can be brought into the form given below:

$$\left[\begin{array}{c|c} \Delta & B \\ \hline B^* & 0 \end{array}\right] \sim \left[\begin{array}{cc|cc} 0 & 0 & I & 0 \\ 0 & \Delta_1 & 0 & 0 \\ \hline I & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \tag{17}$$

where the size of I is equal to the rank of B. It is now easy to see that in order for the rank of (15) to be equal to that of (16), Δ_1 must be the zero matrix. The congruence transformation used in the first step can be analyzed to recover a solution H. ■

Proof: [Theorem 1] We first address the necessity ("only if") part in the theorem. We substitute $G(\lambda)$ in (3) with $B + \lambda AG(\lambda)$ to obtain

$$\begin{aligned} \Sigma &= B\int_{-\pi}^{\pi}\frac{d\mu(\theta)}{2\pi}B^* + \int_{-\pi}^{\pi}\left(e^{j\theta}AG(e^{j\theta})\frac{d\mu(\theta)}{2\pi}\right)B^* \\ &\quad +B\int_{-\pi}^{\pi}\left(\frac{d\mu(\theta)}{2\pi}G(e^{j\theta})^*A^*e^{-j\theta}\right) \\ &\quad +A\int_{-\pi}^{\pi}\left(G(e^{j\theta})\frac{d\mu(\theta)}{2\pi}G(e^{j\theta})^*\right)A^* \\ &= BH + H^*B^* + A\Sigma A^*, \end{aligned} \tag{18}$$

where

$$\begin{aligned} H &= \int_{-\pi}^{\pi}\left(\frac{d\mu(\theta)}{2\pi}\left(\frac{1}{2}B^* + G(e^{j\theta})^*A^*e^{-j\theta}\right)\right) \\ &= \left[\begin{array}{cccc} \frac{1}{2}R_0 & R_1 & R_2 & \ldots \end{array}\right]\left[\begin{array}{c} B^* \\ B^*A^* \\ B^*(A^*)^2 \\ \vdots \end{array}\right] \end{aligned} \tag{19}$$

with R_ℓ, $\ell = 0, 1, \ldots$ the covariance-lags/Fourier-coefficients of $d\mu$ given in (2). Then (7) follows from (18) and Proposition 1.

The sufficiency ("if" part) of the theorem is contained in claim (ii) of Theorem 2. The proof of Theorem 2 is constructive, and the needed theory and proof developed in Section V. ■

Proof: [of Corollary 1] First we assume that Σ is of the form given in (10) with W commuting with A. It follows that

$$\begin{aligned} \Sigma - A\Sigma A^* &= WE + EW^* - A(WE + EW^*)A^* \\ &= WE - WAEA^* + EW^* - AEA^*W^* \\ &= WBB^* + BB^*W^*. \end{aligned} \tag{20}$$

If $H^* := WB$, because of Proposition 1, equation (7) follows.

For the converse, we begin with (7). The equation $\Sigma - A\Sigma A^* = BH + H^*B^*$ must hold, again because of Proposition 1. Since the pair (A, B) is controllable, i.e., A is cyclic with B a cyclic vector, we can always solve the equation

$$\left[\begin{array}{cccc} B & AB & \ldots & A^{n-1}B \end{array}\right]\left[\begin{array}{c} w_0 \\ w_1 \\ \vdots \\ w_{n-1} \end{array}\right] = H^*$$

for $w_0, \ldots, w_{n-1}$, and define

$$W = w_0 I + w_1 A + \ldots + w_{n-1}A^{n-1}.$$

This matrix is polynomial in A and therefore commutes with A. Hence

$$\Sigma - A\Sigma A^* = BB^*W^* + WBB^*.$$

Replacing BB^* by $E - AEA^*$ allows us to retrace the algebra that led to (20), backwards. We conclude that

$$\Sigma - WE - EW^* = A(\Sigma - WE - EW^*)A^*.$$

Since the eigenvalues of A are in $\mathbb{D}$ it follows that the unique solution to $X = AXA^*$ is the zero matrix. Hence, $\Sigma = WE + EW^*$ as claimed in the corollary. ■

V. Analytic interpolation

It is helpful to note that the statements of Theorems 1 and 2 are not affected when we transform the data according to:

$$(A, B, \Sigma) \mapsto (T^{-1}AT, T^{-1}B, T^{-1}\Sigma(T^{-1})^*) \tag{21}$$

with T an invertible matrix. This simply corresponds to a change of coordinates for the state space of (1). Hence, we can select such a transformation so that the controllability grammian

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

becomes the identity matrix. (If $E \neq I$ simply take T to be the Hermitian square root of E.) Therefore, without loss of generality we assume in the sequel that $E = I$, or equivalently, that

$$AA^* + BB^* = I. \tag{22}$$

A. Rational inner (all-pass) functions

Consider A, B normalized to satisfy (22). It follows that the matrix $[A,\ B]$ can be completed into a unitary matrix

$$U := \left[\begin{array}{cc} A & B \\ C & D \end{array}\right]. \tag{23}$$

In fact $C \in \mathbb{C}^{m\times n}$, and $D \in \mathbb{C}^{m\times m}$) are determined up to a left unitary transformation. We note that such a left-unitary transformation does not affect the subsequent analysis and hence, from this point on, we assume that the selection of C, D represents any such particular completion.

For future reference we list all algebraic relations between A, B, C, D inherited by the fact that U is unitary:

$$\begin{aligned} BB^* &= I - AA^*, && (24)\\ DD^* &= I - CC^*, && (25)\\ AC^* &= -BD^*, \text{ as well as} && (26)\\ C^*C &= I - A^*A, && (27)\\ D^*D &= I - B^*B, && (28)\\ A^*B &= -\ C^*D. && (29) \end{aligned}$$

Proposition 2: If U in (23) is unitary, the matrix-valued function

$$V(\lambda) = D + \lambda C(I - \lambda A)^{-1}B \tag{30}$$

is inner.

Recall that a square matrix-valued function is *inner*, (which is referred to as *all-pass* in the engineering literature,) if it is bounded and analytic in $\mathbb{D}$ with unitary boundary radial limits. In our case where $V(\lambda)$ is rational, it is continuous on the boundary and $V(\lambda)V^*(\lambda^{-1}) = I$.

Proof: Clearly $V(\lambda)$ is analytic in a disc of radius larger than 1. We verify that V is unitary on the unit circle by letting $\lambda = e^{j\theta}$ in the following algebraic identity:

$$\begin{aligned} V(\lambda)V^*(\lambda^{-1}) &= DD^* + \lambda C(I - \lambda A)^{-1}BD^*\\ &\quad +\lambda^{-1}DB^*(I - \lambda^{-1}A^*)^{-1}C^*\\ &\quad +C(I - \lambda A)^{-1}BB^*(I - \lambda^{-1}A^*)^{-1}C^*\\ &= I - CC^* + C\lambda A(I - \lambda A)^{-1}C^*\\ &\quad +C(I - \lambda^{-1}A^*)^{-1}\lambda^{-1}A^*C^*\\ &\quad +C(I - \lambda A)^{-1}(I - AA^*)(I - \lambda^{-1}A^*)^{-1}C^*\\ &= I. \end{aligned}$$

For the first step of the derivation we used (25), (24), and (26), while for the last step we used the identity

$$\begin{aligned} 0 &= M(I - M)^{-1} + I + (I - N)^{-1}N \\ &\quad - (I - M)^{-1}(I - MN)(I - N)^{-1} \end{aligned} \tag{31}$$

for $M = \lambda A$ and $N = \lambda^{-1}A^*$. ■

B. The analytic interpolation problem

We now translate condition (19) into an analytic interpolation constraint for a corresponding positive real function.

Proposition 3: Assume that the function

$$F(\lambda) = \frac{1}{2}R_0 + \lambda R_1 + \lambda^2 R_2 + \ldots \tag{32}$$

is in the class $\mathcal{F}$. Then, its Taylor/Fourier series coefficients R_ℓ $(\ell = 0, 1, \ldots)$ satisfy equation (19), i.e.,

$$H = \begin{bmatrix} \frac{1}{2}R_0 & R_1 & R_2 & \ldots \end{bmatrix} \begin{bmatrix} B^* \\ B^*A^* \\ B^*(A^*)^2 \\ \vdots \end{bmatrix} \tag{33}$$

if and only if

$$F(\lambda) = F_0(\lambda) + Q(\lambda)V(\lambda) \tag{34}$$

with $F_0(\lambda) := HG(\lambda)$ and $Q(\lambda)$ analytic in $\mathbb{D}$.

Proof: First assume that the R_ℓ's, which are uniformly bounded since $F(\lambda) \in \mathcal{F}$, satisfy equation (33). It follows that the negative Fourier coefficients of $F(\lambda)G(\lambda)^*$ are

$$H(A^\ell)^* \text{ for } \ell = 0, 1, 2, \ldots, \tag{35}$$

and hence independent of the particular $F(\lambda)$. The negative Fourier coefficients of $F_0(\lambda)G(\lambda)^*$,

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} HG(e^{j\theta})G(e^{j\theta})^* e^{j\ell\theta}d\theta = H(A^\ell)^* \text{ for } \ell = 0, 1, 2, \ldots,$$

are identical to those of $F_0(\lambda)G(\lambda)^*$. It follows that

$$(F(\lambda) - F_0(\lambda))\, G(\lambda)^* = \Gamma_1\lambda + \Gamma_2\lambda + \ldots$$

is analytic in $\mathbb{D}$, and then, that

$$\begin{aligned} (F(\lambda) - F_0(\lambda))\, V(\lambda)^* &= (F(\lambda) - F_0(\lambda))\left(D^* + \lambda^{-1}G(\lambda)^*C^*\right)\\ &= (F(\lambda) - F_0(\lambda))\, D^* +\\ &\qquad \lambda^{-1}\left(\Gamma_1\lambda + \Gamma_2\lambda + \ldots\right)C^*\\ &= Q(\lambda) \end{aligned}$$

is also analytic in $\mathbb{D}$. Thus, (34) holds.

Now assume that (34) holds. To prove (33) it suffices to show that the 0th Fourier coefficient of $F(\lambda)G(\lambda)^*$ is simply H. We first note that

$$\begin{aligned} \frac{1}{2\pi}\int_{-\pi}^{\pi} F_0(e^{j\theta})G(e^{j\theta})^*d\theta &= \frac{1}{2\pi}\int_{-\pi}^{\pi} HG(e^{j\theta})G(e^{j\theta})^*d\theta\\ &= H. \end{aligned}$$

Then, making use of (34),

$$\begin{aligned} \lim_{r\nearrow 1}\frac{1}{2\pi}\int_{-\pi}^{\pi} F(e^{j\theta})G(e^{j\theta})^*d\theta &= H +\\ &\lim_{r\nearrow 1}\frac{1}{2\pi}\int_{-\pi}^{\pi} Q(e^{j\theta})V(e^{j\theta})G(e^{j\theta})^*d\theta \end{aligned}$$

We claim that the last integral is equal to zero. To show this we first use equations (27) and (29) to derive the identity

$$\begin{aligned} (\lambda I - A^*)^{-1}C^*V(\lambda) &= -(\lambda I - A^*)^{-1}A^*B\\ &\quad +(I - \lambda^{-1}A^*)^{-1}(I - A^*A)(I - \lambda A)^{-1}B\\ &= -(I - \lambda A)^{-1}B\\ &= G(\lambda), \end{aligned} \tag{36}$$

where, for the last step, we also used (31) with $M = \lambda A$ and $N = \lambda^{-1}A^*$. Using (36) we deduce that the integrand of (36)

$$\begin{aligned} Q(\lambda)(G(\lambda)V(\lambda)^*)^* &= Q(\lambda)((\lambda I - A^*)^{-1}C^*)^*\\ &= Q(\lambda)\lambda C(I - \lambda A)^{-1} \end{aligned}$$

which is analytic in $\mathbb{D}$. Thus the integrand is zero, and the 0th Fourier coefficient of $F(\lambda)G(\lambda)^*$ is equal to H as claimed. ■

Remark 2: The condition given as (34) represents an analytic interpolation constraint on F which is analogous to the well-known Nehari problem in interpolation with $\mathcal{H}_\infty$ functions [5]. However, although a standard Möbius transformation

$$F(\lambda) \mapsto s(\lambda) = (I - F(\lambda))(I + F(\lambda))^{-1} \tag{37}$$

maps $\mathcal{F}$-functions bijectively into contractive matrix-valued function in $\mathcal{H}_\infty$, it **does not** translate (34) into the analogous problem for $\mathcal{H}_\infty$-functions due to non-commutativity of the matrix functions involved. Hence, we need to develop an analogous theory from scratch. This is done next.

We point out that a direct correspondence between the two via linear fractional transformation works only in the case where $m = 1$ and $V(\lambda)$ is scalar, or when $V(\lambda)$ is a scalar multiple of the identity. In such a case, (37) translates the problem into $s(\lambda) = s_0(\lambda) + q(\lambda)V(\lambda)$, where s_0 is obtained from F_0 via (37), and existing results from the literature can be used to parametrize all solutions (as noted in [7]).

C. Co-invariant subspaces and the "positive-real Nehari"

Throughout, $\mathcal{H}_2^m$ represents *row* vector-valued functions in $\mathcal{H}_2$. The "forward shift" corresponds to multiplication by λ and is denoted by S, while the "backward shift" is given by

$$S^* : \mathcal{H}_2^\ell \to \mathcal{H}_2^\ell : x(\lambda) \mapsto \mathbf{\Pi}_{\mathcal{H}_2}\lambda^{-1}x(\lambda)$$

and is precisely the adjoint operator to S. Subspaces which are invariant under S^* are precisely the ones that are orthogonal to invariant subspaces of the forward shift S, e.g. see [11]. These are of the form

$$\mathcal{K} := \mathcal{H}_2^m \ominus \mathcal{H}_2^m V(\lambda)$$

with $V(\lambda)$ arbitrary inner function, and are referred to as "co-invariant subspaces".

Analytic interpolation *a la Sarason* amounts to specifying an operator on a suitable co-invariant subspace and seeking an extension to the whole of $\mathcal{H}_2^m$ which commutes with S^*. If this is possible, then the operator can then be described via multiplication with an anti-analytic function (followed by projection onto $\mathcal{H}_2^m$). Typically, contractiveness (or, positivity) of the "compressed" operator on the co-invariant subspace *and commutativity of this operator* with S^* are sufficient for existence of a contractive (or positive, respectively) "lifting". This is the content of the celebrated Sarason-Sz. Nagy-Foias commutant lifting theorem [11]. However, in our setting the lifting theorem is not directly applicable. This is due to lack of commutativity between the compressed operator and S^* because of the matricial nature of the interpolation conditions. Thus, we are forced to develop the needed steps from scratch. We use a *Schur-like parametrization* of interpolants via successive one-step extensions—an idea which of course is quite standard in analytic interpolation. Following standard parlance, by "positive/non-negative" we mean that *real part is positive/non-negative definite*, also, in order to avoid excessive technical details many of the statements and arguments will be special to the case of interest where co-invariant subspaces are finite-dimentional, i.e., $V(\lambda)$ is rational.

Let $A, B, V(\lambda), G(\lambda), \mathcal{K}$ be as above. First we note that the orthogonal projection onto $\mathcal{K}$ is given by

$$\mathbf{\Pi}_{\mathcal{K}} : \mathcal{H}_2^{1\times m} \to \mathcal{K} : x(\lambda) \mapsto \left(\mathbf{\Pi}_{(\mathcal{H}_2^{1\times m})^\perp} x(\lambda)V(\lambda)^*\right) V(\lambda).$$

To see this, note that since $V(\lambda)$ is inner, $\mathbf{\Pi}_{\mathcal{K}}$ defined above is idempotant and Hermitian—hence a projection. It is easy to verify that its kernel is precisely $\mathcal{H}_2^{1\times m}V(\lambda)$ and therefore $\mathbf{\Pi}_{\mathcal{K}}$ is the orthogonal projection onto $\mathcal{K}$ as claimed.

Proposition 4: The rows of $G(\lambda)$ form a basis for $\mathcal{K}$.

Proof: We first claim that any element in $\mathcal{K}$ is of the form

$$v(\lambda I - A^*)^{-1}C^*V(\lambda) \tag{38}$$

where $v \in \mathbb{C}^{1\times n}$. To see this note that

$$\begin{aligned}\mathbf{\Pi}_{\mathcal{K}} \quad : \quad & x_0 + x_1\lambda + \ldots \mapsto \\ & \left(\mathbf{\Pi}_{(\mathcal{H}_2^{1\times m})^\perp}(x_0 + x_1\lambda + \ldots)(D^* + \lambda^{-1}B^*C^* + \ldots)\right) V(\lambda) \\ & = v(\lambda^{-1}C^* + \lambda^{-2}A^*C^* + \ldots)V(\lambda)\end{aligned}$$

where $v = x_0B^* + x_1B^*A^* + \ldots$. Next, from equation (36) we see that

$$G(\lambda) = (\lambda I - A^*)^{-1}C^*V(\lambda). \tag{39}$$

In view of (38), the rows of $G(\lambda)$ span $\mathcal{K}$. Finally, if $vG(\lambda) = 0$ for some $v \in \mathbb{C}^{1\times n}$, then necessarily $v = 0$ because (A, B) is controllable. Hence the rows of $G(\lambda)$ are linearly independent and form a basis for $\mathcal{K}$ as claimed. ■

We now consider certain facts about non-negative operators on $\mathcal{H}_2^m$ which commute with S^*. Commutativity with S^* forces a lower triangular Toeplitz structure when expressed with respect to the standard basis in $\mathcal{H}_2^m$:

$$T = \begin{bmatrix} \frac{1}{2}R_0^* & 0 & 0 & \ldots \\ R_1^* & \frac{1}{2}R_0^* & 0 & \ldots \\ R_2^* & R_1^* & \frac{1}{2}R_0^* & \ldots \end{bmatrix}.$$

Alternatively, such an operator can be represented by multiplication with the conjugate of a corresponding function $F(\lambda) = \frac{1}{2}R_0 + R_1\lambda + \ldots$ followed by "projection":

$$T : \mathcal{H}_2^m \to \mathcal{H}_2^m : x(\lambda) \to (x(\lambda)F^*(\lambda))_+$$

where $(\cdot)_+$ denotes truncation of the negative Fourier coefficients. In general, neither the product $x(\lambda)F^*(\lambda)$ is a square-integrable nor the operator T bounded. However, if F is in $\mathcal{F}$ the corresponding operator is defined on a dense subset of $\mathcal{H}_2^m$ and is non-negative. Conversely, if this Toeplitz operator is non-negative, then F is in $\mathcal{F}$. This is well known, e.g., [4].

On $\mathcal{K}$, which is finite-dimensional, T is well defined and bounded. In fact, if the 0th Fourier coefficient of $G(\lambda)F(\lambda)^*$ is H^*, then

$$T(G(\lambda)) = (G(\lambda)F(\lambda)^*)_+ = (I-\lambda A)^{-1}H^*,$$

and if $\langle\cdot,\cdot\rangle$ denotes the standard inner product in $\mathcal{H}_2$,

$$\begin{aligned} W &:= \langle G(\lambda), T(G(\lambda))\rangle \\ &= \begin{bmatrix} B & AB & A^2B & \dots \end{bmatrix} \times \\ &\quad \begin{bmatrix} \frac{1}{2}R_0^* & 0 & 0 & \dots \\ R_1^* & \frac{1}{2}R_0^* & 0 & \dots \\ R_2^* & R_1^* & \frac{1}{2}R_0^* & \dots \\ \vdots & \vdots & \vdots & \end{bmatrix} \begin{bmatrix} B^* \\ B^*A^* \\ B^*(A^*)^2 \\ \vdots \end{bmatrix} \\ &= \begin{bmatrix} H^* & AH^* & A^2H^* & \dots \end{bmatrix} \begin{bmatrix} B^* \\ B^*A^* \\ B^*(A^*)^2 \\ \vdots \end{bmatrix}. \end{aligned}$$

Evidently, H satisfies (19), W satisfies

$$W - AWA^* = H^*B^*, \tag{40}$$

and $\Sigma = 2\Re(W) = W + W^*$ satisfies

$$\Sigma - A\Sigma A^* = BH + H^*B^*. \tag{41}$$

Because $F \in \mathcal{F}$, Σ is non-negative definite. Conversely, (40) is a consequence of the operator being representable as a multiplication by the adjoint of a function $F_0(\lambda) = H(I-\lambda A)^{-1}B$. Any analytic function F which agrees with F_0, in that it defines the same operator, relates to F_0 via (34) as stated in Proposition 3. It turns out that if the real part is non-negative there is always a particular such interpolant F which belongs to $\mathcal{F}$. This is the content of the next statement which is the analog of Nehari's theorem for the case of positive real functions.

Theorem 3: Let $V(\lambda) = D + C\lambda(I-\lambda A)^{-1}B$ be an inner function and let $F_0(\lambda) = H(I-\lambda A)^{-1}B$ be such that the real part of the operator

$$\mathcal{K} \to \mathcal{K} : vG(\lambda) \mapsto \mathbf{\Pi}_{\mathcal{K}} vG(\lambda)F(\lambda)^*,$$

with $v \in \mathbb{C}^{1\times n}$, is non-negative definite. Then there exists a function $F(\lambda) \in \mathcal{F}$ such that

$$F(\lambda) = F_0(\lambda) + Q(\lambda)V(\lambda)$$

with $Q(\lambda)$ is analytic in $\mathbb{D}$.

This theorem is essentially equivalent to part (ii) of Theorem 2. In fact, our preceeding discussion shows that Σ is the corresponding "Pick matrix", i.e., the real part of the "compressed" operator. Although analogous to the standard Nehari theorem (e.g., see [5]), we do not know a direct way that it can be established from existing literature. A proof is given next.

D. One-step extension

We now complete the proof of Theorems 2 and 3. The key technical steps are packaged into two lemmas.

We first recall the definition of the *Schur complement* pivoting about the nonsingular block entry N_{11} of a partitioned matrix

$$N = \begin{bmatrix} N_{11} & N_{12} \\ N_{21} & N_{22} \end{bmatrix},$$

which is

$$[N : N_{11}] := N_{22} - N_{21}N_{11}^{-1}N_{12},$$

and also recall the fact that the *signature* of N which is the triple (p,n,z) of numbers of its positive, negative and zero-eigenvalues, is equal to the sum of the signatures of N_{11} and $[N : N_{11}]$.

The lemma below requires that $\Sigma > 0$ as opposed to only ≥ 0. This restriction will be relaxed by a separate argument after the lemma is applied.

Lemma 1: Let U be a unitary matrix partitioned as in (23), with $A \in \mathbb{C}^{n\times n}$, $D \in \mathbb{C}^{m\times m}$, $B, C' \in \mathbb{C}^{n\times m}$, and $m \leq n$. Further assume that the rank of B is equal to m. If Σ is a positive definite matrix such that the identity

$$\Sigma - A\Sigma A^* = BH + H^*B^*$$

holds for a suitable matrix H, then the matrix

$$\begin{aligned} \rho \; := \; & HB + B^*H^* - HA\Sigma^{-1}A^*H^* \\ & + (D^* - HA\Sigma^{-1}C^*)(C\Sigma^{-1}C^*)^{-1}(D - C\Sigma^{-1}A^*H) \end{aligned} \tag{42}$$

is nonnegative definite.

Proof: Let $\Sigma^{\frac{1}{2}}$ denote the Hermitian square root of Σ, and define

$$\mathcal{J} := \begin{bmatrix} I_n & 0 & 0 \\ 0 & 0 & I_m \\ 0 & I_m & 0 \end{bmatrix}, \text{ and}$$

$$T := \begin{bmatrix} \Sigma^{-\frac{1}{2}}A\Sigma^{\frac{1}{2}} & \Sigma^{-\frac{1}{2}}B & \Sigma^{-\frac{1}{2}}H^* \\ C\Sigma^{-\frac{1}{2}} & 0 & D \\ HA\Sigma^{-\frac{1}{2}} & -I & HB \end{bmatrix}.$$

It can be readily checked that

$$T\mathcal{J}T^* = \begin{bmatrix} I_n & \begin{matrix} 0 & 0 \end{matrix} \\ \begin{matrix} 0 \\ 0 \end{matrix} & \boxed{R} \end{bmatrix}$$

where the 2×2-block submatrix R is

$$R = \begin{bmatrix} C\Sigma^{-1}C^* & C\Sigma^{-1}A^*H^* - D \\ HA\Sigma^{-1}C^* - D^* & HA\Sigma^{-1}A^*H^* - HB - B^*H^* \end{bmatrix}.$$

The signature of $\mathcal{J}$ is $(n+m, m, 0)$ (i.e., it has $n+m$ positive eigenvalues, m negative ones, and no 0-eigenvalue). Therefore $T\mathcal{J}T^*$ can have at most $n+m$ positive eigenvalues. By observing the structure of $T\mathcal{J}T^*$ we see that R can have at most m positive eigenvalues. Because U is unitary and hence $B^*B + D^*D = I_m$, the condition that B has rank equal to m is equivalent to $\|D\| < 1$, which in turn is equivalent to the rank of C being equal to m. Therefore, the (1,1)-block entry of R which is $m \times m$, is positive definite. It follows that the Schur complement of R when pivoting on the (1,1)-block entry can have no positive eigenvalues. But this Schur complement is precisely equal to $-\rho$ which therefore is negative semi-definite. ∎

The next lemma gives an explicit construction of a one-step-extension of a compressed operator to one in a larger co-invariant subspace.

Lemma 2: Let $V(\lambda) = D + C\lambda(I - \lambda A)^{-1}B$ be an inner function with A, B, C, D selected as before so that U in (23) is unitary, let

$$\begin{aligned}\mathcal{K}_0 &:= \mathcal{H}_2^m \ominus \mathcal{H}_2^m V(\lambda),\\ \mathcal{K}_1 &:= \mathcal{H}_2^m \ominus \mathcal{H}_2^m \lambda V(\lambda),\end{aligned}$$

and let $F_0(\lambda) = H(I - \lambda A)^{-1}B$ be such that the real part of the operator

$$\mathcal{W} : \mathcal{K}_0 \to \mathcal{K}_0 : x(\lambda) \mapsto \mathbf{\Pi}_{\mathcal{K}_0} x(\lambda) F_0^*(\lambda)$$

is nonnegative definite. Then, there exists a constant matrix Q_0 such that the function

$$F_1(\lambda) = H_1(I - \lambda A_1)^{-1}B_1,$$

with $H_1 = \begin{bmatrix} H & Q_0 \end{bmatrix}$, satisfies

$$\mathbf{\Pi}_{\mathcal{K}_0} x(\lambda) F_1(\lambda)^* \;=\; \mathbf{\Pi}_{\mathcal{K}_0} x(\lambda) F_0(\lambda)^*, \tag{43}$$

and the real part of the operator

$$\mathcal{W}_1 \;=\; \mathcal{K}_1 \to \mathcal{K}_1 : x(\lambda) \mapsto \mathbf{\Pi}_{\mathcal{K}_1} x(\lambda) F_1(\lambda)^*,$$

is nonnegative definite.

The subspace $\mathcal{K}_1$ in the lemma can be expressed as

$$\mathcal{K}_1 \;=\; \mathcal{H}_2^m \ominus \mathcal{H}_2^m V_1(\lambda)$$

with $V_1(\lambda) = D_1 + C_1\lambda(I - \lambda A_1)^{-1}B_1$ and

$$\begin{aligned} A_1 &= \begin{bmatrix} A & 0 \\ C & 0 \end{bmatrix}, \qquad (44)\\ B_1 &= \begin{bmatrix} B \\ D \end{bmatrix},\\ C_1 &= \begin{bmatrix} 0 & I \end{bmatrix},\\ D_1 &= 0, \text{ and}\\ U_1 &= \begin{bmatrix} A_1 & B_1 \\ C_1 & D_1 \end{bmatrix}\end{aligned}$$

a unitary matrix as before. Indeed, $U_1U_1^* = U_1^*U_1 = I$ as well as $V_1(\lambda) = \lambda V(\lambda)$, can both be verified by straightforward algebra. Thus, inductive reasoning via successive one-step extensions of $F_0(\lambda)$ into $F_i(\lambda)$, $i = 1, 2, \ldots$ (in the obvious notation) allows us to claim the existence of a suitable interpolant on the whole of $\mathcal{H}_2$. We first proceed with the proof of the lemma.

Proof: The claim that the compression of $F_1(\lambda)^*$ on $\mathcal{K}_0$ agrees with $F_0(\lambda)^*$ as in (43), is a direct consequence of the fact $F_1(\lambda) = F_0(\lambda) + Q_0V(\lambda)$. This latter identity can be verified by straightforward algebra.

If Σ represents two times the real part of $\mathcal{W}$, then expressed with respect to the standard basis $G(\lambda) = (I - \lambda A)^{-1}B$ for $\mathcal{K}_0$ is

$$\Sigma = W + W^*,$$

where W and Σ satisfy (40) and (41). Similarly, if Σ_1 is two times the real part of $\mathcal{W}_1$ then

$$\begin{aligned}\Sigma_1 &= A_1\Sigma_1A_1^* + B_1H_1 + H_1^*B_1^* \qquad (45)\\ &= \begin{bmatrix} A \\ C \end{bmatrix} \Sigma \begin{bmatrix} A^* & C^* \end{bmatrix} + \qquad (46)\\ &\quad \begin{bmatrix} B \\ D \end{bmatrix} \begin{bmatrix} H & Q_0 \end{bmatrix} + \begin{bmatrix} H^* \\ {Q_0}^* \end{bmatrix} \begin{bmatrix} B^* & D^* \end{bmatrix}.\end{aligned}$$

Thus, to complete the proof we only need to verify the existence of a Q_0 for which the right hand side of the above equation is nonnegative definite. For simplicity we assume that Σ is positive definite and at the end of the proof we complete the argument for the case of Σ singular as well.

Using the expression for Σ_1 in (46) we compute that

$$U^*\Sigma_1U = \qquad (47)$$
$$\begin{bmatrix} \Sigma & C^*{Q_0}^* + A^*H^* \\ HA + Q_0C & HB + B^*H^* + Q_0D + D^*{Q_0}^* \end{bmatrix}.$$

Since the $(1,1)$ block entry Σ is positive definite, Σ_1 is non-negative definite iff the Schur complement of the above matrix, with pivot the $(1,1)$ entry, is non-negative definite. This Schur complement, after collecting constant/linear/quadratic terms in Q_0 together, is

$$\omega(Q_0) \;:=\; N + Q_0L + L^*{Q_0}^* - Q_0M{Q_0}^* \tag{48}$$

where

$$\begin{aligned} N &:= HB + B^*H^* - HA\Sigma^{-1}A^*H^* \qquad (49)\\ L &:= (D - C\Sigma^{-1}A^*H^*) \text{ and}\\ M &:= C\Sigma^{-1}C^*.\end{aligned}$$

This is a quadratic function of Q_0 and will attain a maximum, in the positive-definite sense, for $Q_0 = L^*M^{-1}$. The maximal value is $N + L^*M^{-1}L$ which is precisely the expression for ρ in (42). From Lemma 1, $\rho \geq 0$. In fact,

$$\begin{aligned}\omega(Q_0) \geq 0 \;&\Leftrightarrow\; N + L^*M^{-1}L \geq\\ &\qquad Q_0M{Q_0}^* - Q_0L - L^*{Q_0}^* + L^*M^{-1}L\\ &\Leftrightarrow\; \rho \geq (Q_0M^{\frac{1}{2}} - L^*M^{-\frac{1}{2}})(M^{\frac{1}{2}}{Q_0}^* - M^{-\frac{1}{2}}L)\\ &\Leftrightarrow\; 1 \geq \|(M^{\frac{1}{2}}{Q_0}^* - M^{-\frac{1}{2}}L)\rho^{-\frac{1}{2}}\|.\end{aligned}$$

Therefore, $\omega(Q_0)$ is nonnegative definite for all Q_0 of the form

$$Q_0 = L^*M^{-1} + \rho^{\frac{1}{2}}sM^{-\frac{1}{2}}, \tag{50}$$

where s is a contractive matrix, i.e., $\|s\| \leq 1$. In fact, s is the analog of a Schur parameter (otherwise known as Szegö parameter and partial reflection coefficient) which is encountered in one-step-extension solutions to interpolation problems. The family of admissible Q_0's is a matrix-ball

$$\mathcal{B}(c, r_\ell, r_r) := \{Q_0 : Q_0 = c + r_\ell s r_r \text{ with } \|s\| \leq 1\}$$

around a central value

$$c = L^* M^{-1}, \tag{51}$$

and left and right matrix radii

$$r_\ell = \rho^{\frac{1}{2}} \text{ and } r_r = M^{-\frac{1}{2}}, \tag{52}$$

respectively.

We now explain that the above claims and formulae remain valid even when Σ is singular. (Evidently, inverses have to be replaced by generalized inverses.) We will argue as follows: If Σ is singular, then we use $\Sigma_\epsilon := \Sigma + \epsilon I$ instead. This is positive definite for all $\epsilon > 0$ and satisfies

$$\Sigma_\epsilon = A\Sigma_\epsilon A^* + BH_\epsilon + H_\epsilon^* B^*$$

where $H_\epsilon = H + \epsilon B^*$. (This is because of (22).) Hence, there is a matrix ball $\mathcal{B}_\epsilon$ of admissible Q_0's rendering (47) nonnegative definite when Σ_ϵ is used. Denote by c_ϵ the center of the ball. We show below that if $\epsilon_1 \geq \epsilon_2$ then $\mathcal{B}_{\epsilon_1} \supseteq \mathcal{B}_{\epsilon_2}$. Using this fact the proof is complete since, the limit $c_0 := \lim_{\epsilon\to 0} c_\epsilon$ exists and for $Q_0 = c_0$, the matrix

$$\Phi_\epsilon := \left[\begin{array}{cc} \Sigma_\epsilon & C^* {Q_0}^* + A^* H_\epsilon^* \\ H_\epsilon A + Q_0 C & H_\epsilon B + B^* H_\epsilon^* + Q_0 D + D^* {Q_0}^* \end{array}\right]$$

is nonnegative definite for all $\epsilon > 0$. Thus, by continuity, it is nonnegative for $\epsilon = 0$ as well.

It remains to establish our assertion that if $\epsilon_1 \geq \epsilon_2$ then $\mathcal{B}_{\epsilon_1} \supset \mathcal{B}_{\epsilon_2}$. We first note that

$$\Upsilon := \left[\begin{array}{cc} I & A^* B \\ B^* A & 2B^* B \end{array}\right]$$

is nonnegative definite. To see this simply observe that the Schur complement pivoting on the $(2,2)$ entry is

$$\begin{aligned} 2B^*B - B^*AA^*B &= 2B^*B - B^*B + B^*BB^*B \\ &= B^*B + B^*BB^*B \geq 0. \end{aligned}$$

Now note that

$$\Phi_{\epsilon_1} = \Phi_{\epsilon_2} + (\epsilon_1 - \epsilon_2)\Upsilon.$$

Thus, if a Q_0 makes $\Phi_{\epsilon_2} \geq 0$ then it also makes $\Phi_{\epsilon_1} \geq 0$. This completes the proof. ∎

We now compile the above into a proof for Theorems 2 and 3.

Proof: [Theorem 2 (i)]: This is a direct consequence of (7) and Proposition 1.

Proof: [Theorem 2 (ii) — except for existence]: First, we have shown in Proposition 3 that the algebraic constraint (19) on the coefficients of an $\mathcal{F}$-function is equivalent to the analytic interpolation condition (34). Further, the relationship between an $\mathcal{F}$-function $F(\lambda)$ and a positive measure $d\mu$ was shown in Theorem 1, namely that a function in $\mathcal{F}$ of the form (34) gives rise to a measure which satisfies (3). Conversely, a nonnegative measure $d\mu$ which satisfies (3) gives rise to an $\mathcal{F}$-function which meets (19), and hence by the proposition, (34). The only remaining step is to show that if $\Sigma \geq 0$ then such a pair of $d\mu$ and $F(\lambda)$ exists. This is next.

Proof of Theorem 2 (iii), of existence in (ii), and of Theorem 3:] The sequence $\mathcal{K}_i$ $(i = 1, 2, \ldots)$ is dense in $\mathcal{H}_2$ by virtue of the fact that $\lambda^i V(\lambda) \to 0$ uniformly on compact subsets of $\mathbb{D}$. Since Σ is nonnegative definite, using Lemma 2 we can extend the "compressed operator" W into a positive operator W_i on

$$\mathcal{K}_i = \mathcal{H}_2^m \ominus \mathcal{H}_2^m \lambda^i V(\lambda),$$

successively for $i = 1, 2, \ldots$ by suitable choice of Q_i at each step. Then, W_i corresponds to multiplication by the conjugate of

$$\begin{aligned} F_i(\lambda) &= H_i(I - \lambda A_i)^{-1} B_i \\ &= F_0(\lambda) + \left(Q_0 + Q_1\lambda + \ldots Q_{i-1}\lambda^{i-1}\right) V(\lambda) \end{aligned}$$

(in the obvious notation, consistent with (44) and (34),) while $H_i = [H,\ Q_0,\ Q_1,\ \ldots, Q_{i-1}]$ and $Q_i = c_i + r_{\ell,i} s_i r_{r,i}$ with $\|s_i\| \leq 1$.

We claim that for any choice of such s_i's, $F_i(\lambda)$ tends to a function which is analytic in $\mathbb{D}$. It suffices to prove that the Q_i's are uniformly bounded. To this end we consider the general form of Σ_{i+1}—the real part of W_{i+1} for $i = 1, 2, \ldots$. It is straightforward to verify that

$$\Sigma_2 = \left[\begin{array}{ccc} \Sigma & \gamma^* & \gamma^* \\ \gamma & \delta & Q_1^* \\ \gamma & Q_1 & \delta \end{array}\right]$$

where $\delta := HB + B^*H^* + Q_0 D + D^* Q_0$ and $\gamma = HA + Q_0 C$. Similarly, in general, Σ_{i+1} is of the form

$$\left[\begin{array}{ccccc} \Sigma & \gamma^* & \gamma^* & \ldots & \gamma^* \\ \gamma & \delta & Q_1^* & \ldots & Q_1^* \\ \gamma & Q_1 & \delta & \ldots & Q_2^* \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \gamma & Q_1 & Q_2 & \ldots & \delta \end{array}\right].$$

Thus, as long as $\Sigma_{i+1} \geq 0$ for $i = 1, 2, \ldots$,

$$\left[\begin{array}{cc} \delta & Q_i^* \\ Q_i & \delta \end{array}\right] \geq 0 \tag{53}$$

because this is a submatrix of Σ_{i+1}. It follows that

$$\|Q_i\| \leq \|\delta\|. \tag{54}$$

In order to prove that (53) implies (54), let x, y be a pair of unit norm maximal singular vectors for Q_i, (i.e., unit norm vectors such that $Q_i x = \|Q_i\| y$), and consider the value of the quadratic form defined by the nonnegative matrix in (53), when evaluated on

$$\left[\begin{array}{c} x \\ e^{j\phi} y \end{array}\right]$$

with $e^{j\phi}$ chosen so that $e^{-j\phi}y^*Q_ix$ is real and negative. The value is

$$x^*\delta x + y^*\delta y + 2\Re\left(e^{-j\phi}y^*Q_ix\right) \quad \leq \quad 2\|\delta\| - 2\|Q_i\|.$$

Therefore, unless (54) holds, the sign of the above expression violates (53). Thus (54) is true for any value of i, and provides the uniform bound we claimed.

Finally, we have established that $F_i(\lambda)$ converges to a function $F(\lambda)$ which is analytic in $\mathbb{D}$. On the other hand, multiplication by its conjugate defines a positive operator on a dense subset of $\mathcal{H}_2^m$. These two imply that the limit $F(\lambda)$ belongs to $\mathcal{F}$. This completes our proof. ∎

Remark 3: Left and/or right radii $r_{\ell,i}, r_{r,i}$, at some step i, may be singular. In such a case, the corresponding Schur parameter s_i needs only be specified by a contractive mapping from the range of $r_{r,i}$ to the orthogonal complement of the null space of $r_{\ell,i}$. Evidently, in the case where the radii become the zero matrices, for some value of i, the extension is unique and specified by the central solution $Q_k = c_k$, for $k \geq i$. For the analogous parametrization of standard block-Toeplitz extensions see [3].

VI. Final remarks

The characterization of the algebraic structure of state covariances, as well as the characterization of all input spectra which are consistent with a given state covariance, have significant implications in time-series analysis. The pertinent theory for the case where the underlying system (A, B) has a scalar input and trivial dynamics (e.g., A is an ordinary shift matrix) is the basis of many of the so-called "modern nonlinear spectral analysis methods" which exploit the Toeplitz structure of the state covariance (e.g., Burg's maximum entropy method, Capon's method, Pisarenko's, MUSIC, ESPRIT, etc.—see [13]). In our earlier work [7], we considered the case of scalar input and nontrivial dynamics and derived a variety of algorithms which encompass these ealier "modern nonlinear methods". Typically, the algorithms represent specific selections for the coefficients Q in our parametrization of input spectra. It was noted in [6, 7] that judicious choice of the filter parameters (A, B) (or, first-order filter-banks as in [1]) leads to algorithms with resolution superior than state of the art. The present work lays out the mathematical basis of the multivariable case. The multivariable case should prove useful in the spectral analysis of vectorial time-series, with potential applications in high resolution imaging with polarimetric synthetic aperture radar, multifrequency radar, etc.

References

[1] C. I. Byrnes, T.T. Georgiou, and A. Lindquist, A new approach to spectral estimation: A tunable high-resolution spectral estimator, *IEEE Trans. on Signal Proc.*, **48(11):** 3189-3206, November 2000.

[2] P. Delsarte, Y. Genin, and Y. Kamp, Orthogonal polynomial matrices on the unit circle, *IEEE Trans. on Circuits and Systems,* **CAS-25**: 149-160, 1978.

[3] P. Delsarte, Y. Genin, and Y. Kamp, Generalized Schur representation of matrix-valued functions, *SIAM J. Alg. Disc. Meth.*, **2(2):** 94-107, 1981.

[4] P. Delsarte, Y. Genin, and Y. Kamp, Schur parameterization of positive definite block-Toeplitz systems, *SIAM J. Appl. Math.*, **36(1):** 34-46, 1979.

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

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

[7] T.T. Georgiou, Spectral Estimation via Selective Harmonic Amplification, *IEEE Trans. on Automatic Contr.*, January 2001, **46(1)**: 29-42.

[8] T.T. Georgiou, Subspace analysis of covariances, Proceedings of the IEEE Conf. on Accoustics, Speech, and Signal Processing, Istanbul, June 2000.

[9] K. Hoffman, **Banach Spaces of Analytic Functions,** Dover Publications, NY (unabridged republication in 1988 of earlier printing by Englewood Cliffs, NJ), 1962.

[10] P. Masani, Recent trends in multivariate prediction theory, in **Multivariate Analysis**, P.R. Krishnaiah, Ed., Academic Press, pp. 351-382, 1966.

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

[12] W.E. Roth, The equations $AX - YB = C$ and $AX - XB = C$ in matrices, *Proc. Amer. Math. Soc.*, **59**: 39-44, 1976.

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