On the Multivariate Circulant Rational Covariance Extension Problem

Anders Lindquist, Chiara Masiero, and Giorgio Picci

***Abstract*— Partial stochastic realization of periodic processes from finite covariance data leads to the circulant rational covariance extension problem and bilateral ARMA models. In this paper we present a convex optimization-based theory for this problem that extends and modifies previous results by Carli, Ferrante, Pavon and Picci on the AR solution, which have been successfully applied to image processing of textures. We expect that our present results will provide an enhancement of these procedures.**

I. Introduction

The rational covariance extension problem is an important problem in systems and control with an extensive literature; see, e.g., [2]–[7], [15], [17], [18], [21], [30] and references therein. Among other things, it is the basic problem in partial stochastic realization theory [3] and Toeplitz matrix completion problems. Covariance extension for *periodic* stochastic processes, on the other hand, leads to matrix completion of Toeplitz matrices with circulant structure and to partial stochastic realizations in the form of bilateral ARMA models. This connects up to a rich realization theory for reciprocal processes [22]–[25].

In [12] Carli, Ferrante, Pavon and Picci presented a maximum-entropy approach to this circulant covariance extension problem, thereby providing a procedure for determining the unique bilateral AR model matching the covariance sequence. However, recently it was discovered that the circulant covariance extension can be recast in the context of the optimization-based theory of moment problems with rational measures developed in [1], [4], [5], [7]–[9], [11], [19], [20] allowing for a complete parameterization of all bilateral ARMA realizations, and a complete theory for the scalar case was presented in [26]. The present paper provides a first step in generalizing this theory to the multivariable case.

The AR theory of [12] has been successfully applied to image processing of textures [14], [31], and we anticipate an enhancement of such methods by allowing for more general ARMA realizations. As pointed out in [26] the circulant rational convariance extension theory provides a fast approximation procedure for solving the regular rational covariance extension problem, as it is based on fast Fourier transforms (FFT), and in the present paper we shall provide numerical evidence that this also holds in the multivariable case.

The outline of the paper goes as follows. In Section II we review the regular multivariable rational covariance extension problem and harmonic analysis on the discrete unit circle. Then in Section III we present our main result on the multivariable circulant rational covariance extension problem, parametrizing the family of solutions, and in Section IV we show how logarithmic moments can be used to determine the best particular solution. Finally, in Section V we provide two numerical examples demonstrating the power of circulant covariance extension as a tool for approximation.

II. Preliminaries

A. The multivariable rational covariance extension problem

We begin by reviewing basic results from [1] in the formalism of [11]. Given a sequence $C_0, C_1, \dots, C_n$ in $\mathbb{C}^{m\times m}$ (with C_0 Hermitian symmetric) such that the block Toeplitz matrix

$$\mathbf{T}_n = \begin{bmatrix} C_0 & C_1^* & C_2^* & \cdots & C_n^* \\ C_1 & C_0 & C_1^* & \cdots & C_{n-1}^* \\ C_2 & C_1 & C_0 & \cdots & C_{n-2}^* \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ C_n & C_{n-1} & C_{n-2} & \cdots & C_0 \end{bmatrix} \tag{1}$$

is positive definite, find an infinite extension $C_{n+1}, C_{n+3}, C_{n+3}, \dots$ such that, the series expansion

$$\Phi(e^{i\theta}) = \sum_{k=-\infty}^{\infty} C_k e^{-ik\theta}, \quad C_{-k} = C_k^*, \tag{2}$$

converges for all $\theta \in [-\pi, \pi]$ to a positive $m \times m$ spectral density that takes the rational form

$$\Phi(z) = P(z)Q(z)^{-1}. \tag{3}$$

This is a moment problem since it follows from (2) that the spectral density Φ satisfies the moment conditions

$$\int_{-\pi}^{\pi} e^{ik\theta}\Phi(e^{i\theta})\frac{d\theta}{2\pi} = C_k, \quad k = 0, 1, \dots, n. \tag{4}$$

We stress that this paper is a first step in establishing a complete theory for the multivariable case. For technical reasons, we confine our ARMA models to those whose transfer function has a matrix representation with a scalar numerator polynomial. Thus, P is a symmetric trigonometric polynomial of the form

$$P(e^{i\theta}) = \sum_{k=-n}^{n} p_k e^{-ik\theta}, \quad p_{-k} = \bar{p}_k, \tag{5}$$

This research was supported by grants from VR, ACCESS, Fondazione “Ing. Aldo Gini” and the Italian Ministry for Education and Research (MIUR).

A. Lindquist is with the Department of Automation, Shanghai Jiao Tong University, Shanghai, China, and the Center for Industrial and Applied Mathematics (CIAM) and the ACCESS Linnaeus Center, KTH Royal Institute of Technology, Stockholm, Sweden, alq@kth.se

C. Masiero and Giorgio Picci are with the Department of Information Engineering, University of Padova, via Gradenigo 6/B, 35131 Padova, Italy; e-mail: masiero.chiara@dei.unipd.it and picci@dei.unipd.it, respectively.

of degree at most n, whereas Q is a symmetric trigonometric $m \times m$ matrix polynomial

$$Q(e^{i\theta}) = \sum_{k=-n}^{n} Q_k e^{-ik\theta}, \quad Q_{-k} = Q_k^*. \tag{6}$$

Let $\mathfrak{P}_+^{(m,n)}$ be the set of matrix polynomials (6) which are positive definite for all $\theta \in [-\pi, \pi]$. This is a convex cone, the closure of which we shall denote $\overline{\mathfrak{P}_+^{(m,n)}}$. Now, defining the trigonometric matrix polynomial

$$C(e^{i\theta}) = \sum_{k=-n}^{n} C_k e^{-ik\theta}, \quad C_{-k} = C_k^*, \tag{7}$$

we have

$$\langle C, Q\rangle := \int_{-\pi}^{\pi} \text{tr}\left\{C(e^{i\theta})Q(e^{i\theta})^*\right\} \frac{d\theta}{2\pi} = \sum_{k=-n}^{n} \text{tr}\{C_k Q_k^*\}. \tag{8}$$

If $Q \in \mathfrak{P}_+^{(m,n)}$, then there is a stable spectral factor

$$A(z) = A_0 z^n + A_1 z^{n-1} + \cdots + A_n \tag{9}$$

such that $Q(z) = A(z)A(z)^*$, and consequently

$$\langle C, Q\rangle = \int_{-\pi}^{\pi} \text{tr}\left\{A(e^{i\theta})C(e^{i\theta})A(e^{i\theta})^*\right\} \frac{d\theta}{2\pi} = \text{tr}\{\mathbf{A}\mathbf{T}_n\mathbf{A}^*\}, \tag{10}$$

where $\mathbf{A} := (A_0, A_1, \dots, A_n)$. Let $\mathfrak{C}_+^{(m,n)}$ be the *interior* of the dual cone of all (7) such that

$$\langle C, Q\rangle \geq 0 \quad \text{for all } Q \in \overline{\mathfrak{P}_+^{(m,n)}}. \tag{11}$$

This is an open convex cone. It follows from (10) that $C \in \mathfrak{C}_+^{(m,n)}$ if and only if $\mathbf{T}_n$ is positive definite.

Next, consider the optimization problem to maximize the generalized entropy

$$\mathbb{I}_P(\Phi) = \int_{-\pi}^{\pi} P(e^{i\theta}) \log\det \Phi(e^{i\theta}) \frac{d\theta}{2\pi} \tag{12}$$

over all Φ that are positive definite on the unit circle subject to the moment conditions (4).

Theorem 1 (Blomqvist-Lindquist-Nagamune [1]): For each $(P, C) \in \mathfrak{P}_+^{(1,n)} \times \mathfrak{C}_+^{(m,n)}$, the problem to maximize (12) subject to the moment conditions (4) has a unique solution $\hat{\Phi}$, and it has the form

$$\hat{\Phi}(z) = P(z)\hat{Q}(z)^{-1}, \tag{13}$$

where $\hat{Q} \in \mathfrak{P}_+^{(m,n)}$ is the unique solution to the dual problem to minimize

$$\mathbb{J}_P(Q) = \langle C, Q\rangle - \int_{-\pi}^{\pi} P(e^{i\theta}) \log\det Q(e^{i\theta}) \frac{d\theta}{2\pi} \tag{14}$$

over all $Q \in \mathfrak{P}_+^{(m,n)}$.

Consequently, a large subclass of all multivariable rational covariance extensions, namely those for which Φ takes the form (3), are completely parameterized by the $P \in \mathfrak{P}_+^{(1,n)}$.

B. Harmonic analysis in $\mathbb{Z}_{2N}$ and stationary periodic vector processes

The discrete Fourier transform (DFT) $\mathcal{F}$ maps a finite sequence $\mathbf{g} = \{\mathbf{g}_k;\, k = -N+1, \dots, N\}$ in $\mathbb{C}^m$, into a sequence of complex m-vectors

$$\mathbf{G}(\zeta_j) := \sum_{k=-N+1}^{N} \mathbf{g}_k \zeta_j^{-k}, \qquad j = -N+1, -N+2, \dots, N, \tag{15}$$

where $\zeta_j := e^{ij\pi/N}$. Here we have defined the discrete variable ζ taking the $2N$ values ζ_j, $j = -N+1, \dots, 0, \dots, N$ and running counterclockwise on the discrete unit circle $\mathbb{T}_{2N}$. In particular, we have $\zeta_j = (\zeta_1)^j$ and $\zeta_{-k} = \overline{\zeta_k}$. The inverse DFT $\mathcal{F}^{-1}$ is given by

$$\mathbf{g}_k = \frac{1}{2N} \sum_{j=-N+1}^{N} \zeta_j^k \mathbf{G}(\zeta_j), \quad k = -N+1, -N+2, \dots, N, \tag{16}$$

which can also be written as a Stieltjes integral

$$\mathbf{g}_k = \int_{-\pi}^{\pi} e^{ik\theta}\mathbf{G}(e^{i\theta}) d\nu(\theta), \quad k = -N+1, -N+2, \dots, N, \tag{17}$$

where ν is a step function with steps $\frac{1}{2N}$ at each ζ_k; i.e.,

$$d\nu(\theta) = \sum_{j=-N+1}^{N} \delta(e^{i\theta} - \zeta_j)\frac{d\theta}{2N}. \tag{18}$$

With $\mathbf{H}$ being the DFT of $\{\mathbf{h}_k\}$,

$$\begin{aligned} \sum_{j=-N+1}^{N} \mathbf{g}_j \mathbf{h}_j^* &= \frac{1}{2N} \sum_{k=-N+1}^{N} \mathbf{G}(\zeta_k)\mathbf{H}(\zeta_{-k})^* \\ &= \int_{-\pi}^{\pi} \mathbf{G}(e^{i\theta})\mathbf{H}(e^{i\theta})^* d\nu, \end{aligned} \tag{19}$$

which is *Plancherel's Theorem* for DFT. From this we see that

$$\langle G, H\rangle := \int_{-\pi}^{\pi} \text{tr}\left\{\mathbf{G}(e^{i\theta})\mathbf{H}(e^{i\theta})^*\right\} d\nu = \sum_{j=-N+1}^{N} \text{tr}\{\mathbf{g}_j \mathbf{h}_j^*\} \tag{20}$$

is computed exactly as in (8) despite the change of measure in the integral. Hence results such as (10) hold also with the Stieltjes measure $d\nu$.

Occasionally we shall write the discrete Fourier transform (15) in the matrix form

$$\hat{\mathbf{g}} = \mathbf{F}\mathbf{g}, \tag{21}$$

where $\hat{\mathbf{g}} := \big(\mathbf{G}(\zeta_{-N+1})^{\mathsf{T}}, \mathbf{G}(\zeta_{-N+2})^{\mathsf{T}}, \dots, \mathbf{G}(\zeta_N)^{\mathsf{T}}\big)^{\mathsf{T}}$, $\mathbf{g} := \big(\mathbf{g}_{-N+1}{}^{\mathsf{T}}, \mathbf{g}_{-N+2}{}^{\mathsf{T}}, \dots, \mathbf{g}_N{}^{\mathsf{T}}\big)^{\mathsf{T}}$ and $\mathbf{F}$ is the nonsingular $2mN \times 2mN$ block Vandermonde matrix

$$\mathbf{F} = \begin{bmatrix} \zeta_{-N+1}^{N-1} I_m & \zeta_{-N+1}^{N-2} I_m & \cdots & \zeta_{-N+1}^{-N} I_m \\ \vdots & \vdots & \cdots & \vdots \\ \zeta_0^{N-1} I_m & \zeta_0^{N-2} I_m & \cdots & \zeta_0^{-N} I_m \\ \vdots & \vdots & \cdots & \vdots \\ \zeta_N^{N-1} I_m & \zeta_N^{N-2} I_m & \cdots & \zeta_N^{-N} I_m \end{bmatrix}. \tag{22}$$

Likewise, it follows from (16) that

$$\mathbf{g} = \frac{1}{2N}\mathbf{F}^*\hat{\mathbf{g}}, \tag{23}$$

i.e., $\mathcal{F}^{-1}$ corresponds to $\frac{1}{2N}\mathbf{F}^*$. Consequently, $\mathbf{F}\mathbf{F}^* = 2N\,\mathbf{I}$, and hence $\mathbf{F}^{-1} = \frac{1}{2N}\mathbf{F}^*$ and $(\mathbf{F}^*)^{-1} = \frac{1}{2N}\mathbf{F}$.

Next consider be a zero-mean stationary m-dimensional process $\{y(t)\}$ defined on $\mathbb{Z}_{2N}$; i.e., a stationary process defined on a finite interval $[-N+1,\, N]$ of the integer line $\mathbb{Z}$ and extended to all of $\mathbb{Z}$ as a periodic stationary process with period $2N$. Let $C_{-N+1}, C_{-N+2}, \dots, C_N$ be the $m \times m$ covariance lags $C_k := \mathbb{E}\{y(t+k)y(t)^*\}$, and define its discrete Fourier transformation

$$\Phi(\zeta_j) := \sum_{k=-N+1}^{N} C_k \zeta_j^{-k}, \qquad j = -N+1, \dots, N, \tag{24}$$

which is a positive, Hermitian matrix-valued function of ζ. Then, as seen from (16) and (17),

$$\begin{aligned} C_k &= \frac{1}{2N}\sum_{j=-N+1}^{N} \zeta_j^k \Phi(\zeta_j) \\ &= \int_{-\pi}^{\pi} e^{ik\theta}\Phi(e^{i\theta})d\nu, \quad k = -N+1, \dots, N. \end{aligned} \tag{25}$$

The $m \times m$ matrix function Φ is the *spectral density* of the vector process y. In fact, let

$$\hat{y}(\zeta_k) := \sum_{t=-N+1}^{N} y(t)\zeta_k^{-t}, \quad k = -N+1, \dots, N, \tag{26}$$

be the discrete Fourier transformation of the process y. Since $\frac{1}{2N}\sum_{t=-N+1}^{N}(\zeta_k\zeta_\ell^*)^t = \delta_{k\ell}$, the random variables (26) are uncorrelated, and

$$\frac{1}{2N}\mathbb{E}\{\hat{y}(\zeta_k)\hat{y}(\zeta_\ell)^*\} = \Phi(\zeta_k)\delta_{k\ell}. \tag{27}$$

This yields a spectral representation of y analogous to the usual one, namely

$$y(t) = \frac{1}{2N}\sum_{k=-N+1}^{N} \zeta_k^t\, \hat{y}(\zeta_k) = \int_{-\pi}^{\pi} e^{ik\theta} d\hat{y}(\theta), \tag{28}$$

where $d\hat{y} := \hat{y}(e^{i\theta})d\nu$.

C. Block-circulant matrices

Circulant block matrices are block Toeplitz matrices with a special circulant structure

$$\mathrm{Circ}\{\Lambda_0, \Lambda_1, \dots, \Lambda_\nu\} = \begin{bmatrix} \Lambda_0 & \Lambda_\nu & \Lambda_{\nu-1} & \cdots & \Lambda_1 \\ \Lambda_1 & \Lambda_0 & \Lambda_\nu & \cdots & \Lambda_2 \\ \Lambda_2 & \Lambda_1 & \Lambda_0 & \cdots & \Lambda_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \Lambda_\nu & \Lambda_{\nu-1} & \Lambda_{\nu-2} & \cdots & \Lambda_0 \end{bmatrix},$$

where the block columns (or, equivalently, block rows) are shifted cyclically, and where $\Lambda_0, \Lambda_1, \dots, \Lambda_\nu$ here are taken to be complex matrices. In the multivariable circulant rational covariance extension problem we consider *Hermitian* circulant matrices

$$\mathbf{M} := \mathrm{Circ}\{M_0, M_1, M_2, \dots, M_N, M_{N-1}^*, \dots, M_1^*\}, \tag{29}$$

which can be represented in the form

$$\mathbf{M} = \sum_{k=-N+1}^{N} S^{-k} \otimes M_k, \quad M_{-k} = M_k^* \tag{30}$$

where $\otimes$ is the Kronecker product and S is the nonsingular $2N \times 2N$ cyclic shift matrix

$$S := \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix}. \tag{31}$$

The $m \times m$ pseudo-polynomial

$$M(\zeta) = \sum_{k=-N+1}^{N} M_k \zeta^{-k}, \quad M_{-k} = M_k^* \tag{32}$$

is called the *symbol* of $\mathbf{M}$. Let $\mathbf{S}$ be the $2mN \times 2mN$ cyclic shift matrix

$$\mathbf{S} = S \otimes I_m = \begin{bmatrix} 0 & I_m & 0 & \dots & 0 \\ 0 & 0 & I_m & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & I_m \\ I_m & 0 & 0 & 0 & 0 \end{bmatrix}. \tag{33}$$

Clearly $\mathbf{S}^{2N} = \mathbf{S}^0 = \mathbf{I} := I_{2mN}$, and

$$\mathbf{S}^{k+2N} = \mathbf{S}^k, \qquad \mathbf{S}^{2N-k} = \mathbf{S}^{-k} = (\mathbf{S}^k)^\mathsf{T}. \tag{34}$$

Moreover,

$$\mathbf{S}\mathbf{M}\mathbf{S}^* = \mathbf{M}, \tag{35}$$

is both necessary and sufficient for $\mathbf{M}$ to be circulant. With $\mathbf{g} := \big(\mathbf{g}_{-N+1}{}^\mathsf{T}, \mathbf{g}_{-N+2}{}^\mathsf{T}, \dots, \mathbf{g}_N{}^\mathsf{T}\big)^\mathsf{T}$, we have

$$[\mathbf{S}\mathbf{g}]_k = \mathbf{g}_{k+1}, \quad k \in \mathbb{Z}_{2N}. \tag{36}$$

Then, in view of (15), $\zeta\mathcal{F}(\mathbf{g})(\zeta) = \mathcal{F}(\mathbf{S}\mathbf{g})(\zeta)$, from which we have

$$\mathcal{F}(\mathbf{M}\mathbf{g})(\zeta) = M(\zeta)\mathcal{F}(\mathbf{g})(\zeta), \tag{37}$$

where the $m \times m$ matrix fuction $M(\zeta)$ is the symbol (32) of the circulant matrix $\mathbf{M}$. An important property of circulant block matrices is that they can be block-diagonalized by the discrete Fourier transform. More precisely, it follows from (37) that

$$\mathbf{M} = \frac{1}{2N}\mathbf{F}^*\mathrm{diag}\big(M(\zeta_{-N+1}), \dots, M(\zeta_N)\big)\mathbf{F}, \tag{38}$$

where 'diag' denotes block diagonal. Hence the inverse is

$$\mathbf{M}^{-1} = \frac{1}{2N}\mathbf{F}^*\mathrm{diag}\big(M(\zeta_{-N+1})^{-1}, \dots, M(\zeta_N)^{-1}\big)\mathbf{F}, \tag{39}$$

and, since

$$\begin{aligned} \mathbf{S} &= \frac{1}{2N}\mathbf{F}^*\mathrm{diag}\big(\zeta_{-N+1}, \dots, \zeta_N\big)\mathbf{F} \\ \mathbf{S}^* &= \frac{1}{2N}\mathbf{F}^*\mathrm{diag}\big(\zeta_{-N+1}^{-1}, \dots, \zeta_N^{-1}\big)\mathbf{F}, \end{aligned}$$

we have

$$\mathbf{S}\mathbf{M}^{-1}\mathbf{S}^* = \mathbf{M}^{-1}.$$

Hence $\mathbf{M}^{-1}$ is also a circulant block matrix with symbol $M(\zeta)^{-1}$. In general, in view of the circulant property (30) and (34), quotients of symbols are themselves pseudo-polynomials of degree at most N and hence symbols. More generally, if $\mathbf{A}$ and $\mathbf{B}$ are circulant block matrices of the same dimension with symbols $A(\zeta)$ and $B(\zeta)$ respectively, then $\mathbf{AB}$ and $\mathbf{A}+\mathbf{B}$ are circulant matrices with symbols $A(\zeta)B(\zeta)$ and $A(\zeta)+B(\zeta)$, respectively. In fact, the circulant matrices of a fixed dimension form an algebra, and the DFT is an *algebra homomorphism* of the set of circulant matrices onto the pseudo-polynomials of degree at most N in the variable $\zeta \in \mathbb{T}_{2N}$.

III. The multivariable circulant rational covariance extension problem

Given $C := (C_0, C_1, \dots, C_n) \in \mathfrak{C}_+^{(m,n)}$ for some $n < N$, find an $m \times m$ spectral density Φ of the form (3) such that

$$\int_{-\pi}^{\pi} e^{ik\theta}\Phi(e^{i\theta})d\nu = C_k, \quad k = 0, 1, 2, \dots, n. \tag{40}$$

It turns out that this yields an extension

$$C_k = \int_{-\pi}^{\pi} e^{ik\theta}\Phi(e^{i\theta})d\nu, \quad k = n+1, n+2, \dots, N \tag{41}$$

such that the banded Hermitian block-circulant matrix

$$\mathbf{C} = \text{Circ}\{C_0, C_1, \dots, C_n, 0, \dots, 0, C_n^*, \dots, C_1^*\} \tag{42}$$

with symbol (7) is extended to a Hermitian block-circulant matrix

$$\mathbf{\Sigma} := \text{Circ}\{C_0, C_1, C_2, \dots, C_N, C_{N-1}^*, \dots, C_2^*, C_1^*\} \tag{43}$$

that is positive definite with symbol Φ.

We now proceed to solve the multivariable circulant rational covariance extension problem in terms of the symbols, and then interpret the results in terms of matrices.

A. Circulant rational covariance extension in terms of symbols

Define the cone $\mathfrak{P}_+^{(m,n)}(N) \supset \mathfrak{P}_+^{(m,n)}$ of $m \times m$ matrix-valued trigonometric polynomials (6) such that

$$Q(\zeta_k) > 0 \quad k = -N+1, -N+2, \dots, N. \tag{44}$$

Then $\mathfrak{P}_+^{(m,n)}(N) \supset \mathfrak{P}_+^{(m,n)}(2N) \supset \mathfrak{P}_+^{(m,n)}(4N) \supset \cdots \supset \mathfrak{P}_+^{(m,n)}$, and the corresponding dual cones satisfy

$$\mathfrak{C}_+^{(m,n)}(N) \subset \mathfrak{C}_+^{(m,n)}(2N) \subset \mathfrak{C}_+^{(m,n)}(4N) \subset \cdots \subset \mathfrak{C}_+^{(m,n)}. \tag{45}$$

Theorem 2: Let $C \in \mathfrak{C}_+^{(m,n)}(N)$. Then, for each $P \in \mathfrak{P}_+^{(1,n)}(N)$, there is a unique $Q \in \mathfrak{P}_+^{(m,n)}(N)$ such that

$$\Phi = PQ^{-1} \tag{46}$$

satisfies the moment conditions (40).

Theorem 2 follows from the following theorem, which also provides an algorithm for computing the solution.

Theorem 3: For each $(P, C) \in \mathfrak{P}_+^{(1,n)}(N) \times \mathfrak{C}_+^{(m,n)}(N)$, the problem to maximize the functional

$$\mathbb{I}_P(\Phi) = \int_{-\pi}^{\pi} P(e^{i\theta}) \log\det\Phi(e^{i\theta})d\nu \tag{47}$$

subject to the moment conditions (40) has a unique solution $\hat{\Phi}$, and it has the form

$$\hat{\Phi}(z) = P(z)\hat{Q}(z)^{-1}, \tag{48}$$

where $\hat{Q} \in \mathfrak{P}_+^{(m,n)}(N)$ is the unique solution to the dual problem to minimize

$$\mathbb{J}_P(Q) = \langle C, Q\rangle - \int_{-\pi}^{\pi} P(e^{i\theta}) \log\det Q(e^{i\theta})d\nu \tag{49}$$

over all $Q \in \mathfrak{P}_+^{(m,n)}(N)$.

The proofs of Theorems 2 and 3 follow the lines of [26] and will be given in [27]. It can also be shown that the moment map sending $Q \in \mathfrak{P}_+^{(m,n)}(N)$ to $C \in \mathfrak{C}_+^{(m,n)}(N)$ is a diffeomorphism.

B. Circulant rational covariance extension in terms of matrices

Next we reformulate the optimization problems in terms of circulant matrices. To this end, we define the circulant matrix

$$\mathbf{\Sigma} = \frac{1}{2N}\mathbf{F}^*\text{diag}\big(\Phi(\zeta_{-N+1}), \dots, \Phi(\zeta_N)\big)\mathbf{F} \tag{50}$$

with symbol (46) and the banded numerator matrix

$$\mathbf{P} = \frac{1}{2N}\mathbf{F}^*\text{diag}\big(I_m \otimes P(\zeta_{-N+1}), \dots, I_m \otimes P(\zeta_N)\big)\mathbf{F} \tag{51}$$

of degree at most n with symbol $P(\zeta)I_m$, where the scalar pseudo-polynomial P is given by (5). It can also be shown that

$$\log\mathbf{\Sigma} = \frac{1}{2N}\mathbf{F}^*\text{diag}\big(\log\Phi(\zeta_{-N+1}), \dots, \log\Phi(\zeta_N)\big)\mathbf{F}. \tag{52}$$

Therefore, since $\log\det\Phi = \text{tr}\log\Phi$, the primal functional (47) may be written

$$\begin{aligned}\int_{-\pi}^{\pi} &P(e^{i\theta}) \log\det\Phi(e^{i\theta})d\nu \\ &= \frac{1}{2N}\sum_{j=-N+1}^{N} \text{tr}\left\{P(\zeta_j)\log\Phi(\zeta_j)\right\} \\ &= \frac{1}{2N}\text{tr}\{\mathbf{P}\log\mathbf{\Sigma}\}\end{aligned} \tag{53}$$

and the moment conditions (40) as

$$\frac{1}{2N}\text{tr}\{\mathbf{S}^k\mathbf{\Sigma}\} = C_k, \quad k = 0, 1, \dots, n, \tag{54}$$

or, equivalently, as

$$\mathbf{E}_n{}^{\mathsf{T}}\mathbf{\Sigma}\mathbf{E}_n = \mathbf{T}_n, \quad \text{where } \mathbf{E}_n = \begin{bmatrix}\mathbf{I}_{mn}\\ \mathbf{0}\end{bmatrix}. \tag{55}$$

Consequently, the primal problem amounts to maximizing $\text{tr}\{\mathbf{P}\log\mathbf{\Sigma}\}$ over all Hermitian, positive definite $2mN \times 2mN$ block matrices subject to (54) or (55). This reduces to the primal problem presented in [12] in the special case $P \equiv 1$, except that in [12] there is an extra condition insuring that $\mathbf{\Sigma}$ is circulant. However, in [13] it was shown that this condition is automatically satisfied and is hence not needed.

Similarly the dual functional (49) can be written

$$\int_{-\pi}^{\pi} C(e^{i\theta})Q(e^{i\theta})^* d\nu - \int_{-\pi}^{\pi} P(e^{i\theta}) \log\det Q(e^{i\theta}) d\nu \\ = \frac{1}{2N}\mathrm{tr}\{\mathbf{CQ}\} - \frac{1}{2N}\mathrm{tr}\{\mathbf{P}\log\mathbf{Q}\}, \tag{56}$$

where

$$\mathbf{Q} = \frac{1}{2N}\mathbf{F}^*\mathrm{diag}\big(Q(\zeta_{-N+1}),\dots,Q(\zeta_N)\big)\mathbf{F} \tag{57}$$

and $\mathbf{C}$ is the banded circulant block matrix (42) formed from $C_0, C_1, \dots, C_n$. Therefore, given $C \in \mathfrak{C}_+(N)$, it follows from Theorem 2 that, for each Hermitian, positive-definite circulant block matrix $\mathbf{P}$ with symbol of the form $P(\zeta)I_m$, where P is a pseudo-polynomial of degree at most n, there is a unique $\mathbf{\Sigma}$ given by

$$\mathbf{\Sigma} = \mathbf{Q}^{-1}\mathbf{P}, \tag{58}$$

where $\mathbf{Q}$ is the unique solution of the problem to minimize

$$\mathbb{J}_{\mathbf{P}}(\mathbf{Q}) = \frac{1}{2N}\mathrm{tr}\{\mathbf{CQ}\} - \frac{1}{2N}\mathrm{tr}\{\mathbf{P}\log\mathbf{Q}\} \tag{59}$$

over all Hermitian, circulant block-banded matrices

$$\mathbf{Q} = \mathrm{Circ}\{Q_0, Q_1, \dots, Q_n, 0, \dots, 0, Q_n^*, Q_{n-1}^*, \dots, Q_1^*\}$$

that are positive definite. For the maximum-entropy solution corresponding to $\mathbf{P} = \mathbf{I}$ this reduces to an optimization problem that is different from the one presented in [12].

As observed in [12] the condition $\mathbf{T}_n > 0$ is necessary, but not a sufficient, for feasibility of the circulant block-banded covariance extension problem. In the present setting we see that the Toeplitz condition $\mathbf{T}_n > 0$ is equivalent to $C \in \mathfrak{C}_+^{(m,n)}$, whereas, by Theorem 2, $C \in \mathfrak{C}_+^{(m,n)}(N)$ is required for feasibility. Since $\mathfrak{C}_+^{(m,n)}(N) \subset \mathfrak{C}_+^{(m,n)}$, it follows that the Toeplitz condition cannot be sufficient in general. However, as proved in [12], feasibility is achieved for a sufficiently large N. This can also be seen by noting that the set $\{\zeta_j;\ j = -N+1, \dots, N\}$ becomes dense on the unit circle as $N \to \infty$, and therefore $\mathfrak{P}_+(N) \to \mathfrak{P}_+$. Consequently, $\mathfrak{C}_+(N) \to \mathfrak{C}_+$, and the convergence is monotone in the sense of (45). Therefore, since $\mathfrak{C}_+$ is an open set, there is an N_0 such that any $C \in \mathfrak{C}_+$ will sooner or later end up in $\mathfrak{C}_+(N)$ and remain there as $N \geq N_0$ increases.

IV. Determining $\mathbf{P}$ from Logarithmic Moments

We have parameterized a large class of solutions to the multivariable circulant rational covariance extension problem in a smooth manner by the numerator trigonometric polynomials $P \in \mathfrak{P}_+^{(1,n)}(N)$, or, equivalently, by their corresponding banded circulant matrices $\mathbf{P}$. Next, we show how P can be determined from the logarithmic moments

$$\gamma_k = \int_{-\pi}^{\pi} e^{ik\theta} \log\det\Phi(e^{i\theta}) d\nu, \quad k = 1, 2, \dots, n. \tag{60}$$

Such moments are known as *cepstral coefficients* in speech processing. Let $\Gamma(\zeta)$ be the pseudo-polynomial

$$\Gamma(\zeta) = \sum_{k=-n}^{n} \gamma_k \zeta^{-k}, \tag{61}$$

where $\gamma_{-k} = \bar{\gamma}_k$, $k = 1, 2, \dots, n$ and γ_0 is real.

Consider the problem of finding the spectral density Φ, or, equivalently, the circulant block matrix $\mathbf{\Sigma}$, that maximizes the entropy gain

$$\mathbb{I}(\Phi) = \int_{-\pi}^{\pi} \log\det\Phi(e^{i\theta}) d\nu = \frac{1}{2N}\mathrm{tr}\log\mathbf{\Sigma} \tag{62}$$

subject to the two sets of moment conditions (40) and (60). Such a problem was apparently first considered in the usual trigonometric moment setting in an unpublished technical report [29] and then, independently and in a more elaborate form, in [6], [7], [15].

Setting up the Lagrangian a straightforward calculation yields the dual problem to minimize

$$\mathbb{J}(P,Q) = \langle C, Q\rangle - \int_{-\pi}^{\pi} P(e^{i\theta}) \log\det Q(e^{i\theta}) d\nu \\ - \langle \Gamma, P\rangle + \int_{-\pi}^{\pi} P(e^{i\theta}) \log P(e^{i\theta}) d\nu, \tag{63}$$

over $(P,Q) \in \hat{\mathfrak{P}}_+^{(1,n)}(N) \times \mathfrak{P}_+^{(m,n)}(N)$, where $\hat{\mathfrak{P}}_+^{(1,n)}(N)$ is the bounded subset

$$\hat{\mathfrak{P}}_+^{(1,n)}(N) := \{P \in \mathfrak{P}_+^{(1,n)}(N) \mid p_0 = 1\} \tag{64}$$

of the cone $\mathfrak{P}_+^{(1,n)}(N)$.

The following theorem is a multivariable version of Theorem 8 in [26] and the proof is analogous.

Theorem 4: Suppose that $C \in \mathfrak{C}_+^{(m,n)}(N)$ and that $\gamma_1, \dots, \gamma_n$ are complex numbers. Then there exists a solution $(\hat{P}, \hat{Q})$ that minimizes $\mathbb{J}(P,Q)$ over all $(P,Q) \in \hat{\mathfrak{P}}_+^{(1,n)}(N) \times \mathfrak{P}_+^{(m,n)}(N)$, and, for any such solution

$$\hat{\Phi} = \hat{P}\hat{Q}^{-1} \tag{65}$$

satisfies the covariance moment conditions (40). If, in addition, $\hat{P} \in \mathfrak{P}_+^{(1,n)}(N)$, (65) also satisfies the logarithmic moment conditions (60) and is an optimal solution of the primal problem to maximize the entropy gain (62) given (40) and (60). Then $\hat{Q} \in \mathfrak{P}_+^{(m,n)}(N)$, and the solution is unique. In fact, $\mathbb{J}$ is strictly convex on $\hat{\mathfrak{P}}_+^{(1,n)}(N) \times \mathfrak{P}_+^{(m,n)}(N)$.

Provided $C \in \mathfrak{C}_+(N)$, minimizing $\mathbb{J}(P,Q)$ over all $(P,Q) \in \hat{\mathfrak{P}}_+^{(1,n)}(N) \times \mathfrak{P}_+^{(m,n)}(N)$ will always produce a spectral density with the prescribed covariance lags $C_0, C_1, \dots, C_n$. If the moments $C_0, C_1, \dots, C_n$ and $\gamma_1, \dots, \gamma_n$ come from the same theoretical spectral density, the optimal solution (65) will also match the cepstral coefficients. In practice, however, they will be estimated from different data sets, so there is no guarantee that $\hat{P}$ does not end up on the boundary of $\mathfrak{P}_+^{(1,n)}(N)$ without satisfying the logarithmic moment conditions. Then the problem needs to be regularized, leading to adjusted values of $\gamma_1, \dots, \gamma_n$ consistent with the covariances $C_0, C_1, \dots, C_n$.

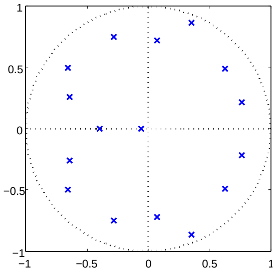

Fig. 1. Autoregressive 2×2 model, with order $n = 8$.

Such a regularization was proposed by P. Enqvist [15] in the context of the usual rational covariance extension problem. The regularized dual problem to find a pair $(P, Q) \in \hat{\mathfrak{P}}_+^{(1,n)}(N) \times \mathfrak{P}_+^{(m,n)}(N)$ minimizing

$$\mathbf{J}_\lambda(P, Q) = \mathbb{J}(P, Q) - \lambda \int_{-\pi}^{\pi} \log P(e^{i\theta}) d\nu \tag{66}$$

for some $\lambda > 0$ will always lead to a solution where $P \in \mathfrak{P}_+^{(1,n)}(N)$. Indeed, (66) will take an infinite value for $P \in \partial\mathfrak{P}_+^{(1,n)}(N)$, since then $P(\zeta_k) = 0$ for some k, and hence the minimum will be in the interior. In circulant form (66) becomes

$$\begin{aligned} \mathbf{J}_\lambda(P, Q) = &\frac{1}{2N}\text{tr}\{\mathbf{C}\mathbf{Q}\} - \frac{1}{2N}\text{tr}\{\mathbf{\Gamma}\mathbf{P}\} \\ &+ \frac{1}{2N}\text{tr}\{\mathbf{P}\log \mathbf{P}\mathbf{Q}^{-1}\} - \frac{\lambda}{2N}\text{tr}\{\log \mathbf{P}\}, \end{aligned} \tag{67}$$

where

$$\mathbf{\Gamma} = \frac{1}{2N}\mathbf{F}^*\text{diag}\big(I_m \otimes \Gamma(\zeta_{-N+1}), \dots, I_m \otimes \Gamma(\zeta_N)\big)\mathbf{F}. \tag{68}$$

Then both sets (40) and (60) of moments are matched provided one adjusts the logarithmic moments $\gamma_1, \gamma_2, \dots, \gamma_n$ to $\gamma_1 + \varepsilon_1, \gamma_2 + \varepsilon_2, \dots, \gamma_n + \varepsilon_n$, where

$$\begin{aligned} \varepsilon_k = \int_{-\pi}^{\pi} e^{ik\theta} \frac{\lambda}{P(e^{i\theta})} d\nu &= \frac{\lambda}{2N} \sum_{j=-N+1}^{N} \frac{\zeta_j^k}{P(\zeta_j)} \\ &= \frac{\lambda}{2N}\text{tr}\{\mathbf{S}^k \mathbf{P}^{-1}\}. \end{aligned} \tag{69}$$

V. Numerical examples

Given a $P \in \mathfrak{P}_+^{(1,n)}$ and a sequence $C_0, C_1, \dots, C_n$ of $m \times m$ covariance lags with a positive definite block Toeplitz matrix (1), Theorem 1 states that there is a unique $Q \in \mathfrak{P}_+^{(m,n)}$ such that $\Phi := PQ^{-1}$ satisfies the moment conditions (4). As pointed out above, for a sufficiently large N the sequence C will also belong to the somewhat smaller cone $\mathfrak{C}_+^{(m,n)}(N)$, and then, by Theorem 2, there will be a unique $Q_N \in \mathfrak{P}_+^{(m,n)}(N)$ such that $\Phi_N := PQ_N^{-1}$ satisfies (40). Next we shall give some numerical results illustrating how Φ can be approximated by Φ_N for various values of N.

In our first example Φ is a 2×2 spectral density corresponding to an AR process of order $n = 8$ with poles as depicted in Fig. 1. Given the theoretical covariance sequence $C_0, C_1, \dots, C_n$ from this Φ, we solve the corresponding circulant moment problem (40) for various values of N to obtain a bilateral AR representation of order $n = 8$ with spectral density Φ_N. Fig. 2 illustrates the approximation error $\|\Phi(e^{i\theta}) - \hat{\Phi}(e^{i\theta})\|_2$ for $N = 16, 32$ and 64. It turns out that there is no need to go for high values of N.

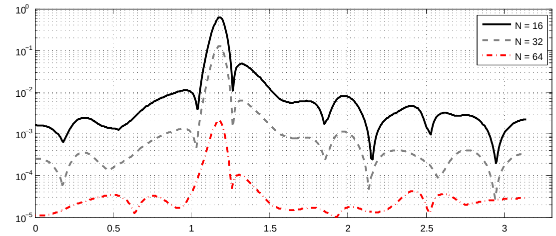

Fig. 2. Norm of the spectral estimation error for bilateral AR models with $N = 16, 32, 64$.

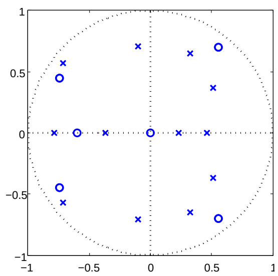

Fig. 3. ARMA 2×2 model, with order $n = 6$.

In the second example we start from a two-dimensional ARMA process with a spectral density $\Phi := PQ^{-1}$, where P is a scalar pseudo-polynomial of degree three and Q is a 2×2 matrix-valued pseudo-polynomial of degree $n = 6$. Its zero poles map is illustrated in Fig. 3. Given its covariance sequence $C_0, C_1, \dots, C_n$ and cepstral sequence $\gamma_1, \gamma_2, \dots, \gamma_n$, we apply the combined covariance and cepstral procedure described in Section IV to determine a pair (P_N, Q_N) for $n = 6$ and a correponding bilateral ARMA model. For comparison we also compute an bilateral AR approximation with $n = 12$ fixing $P = 1$. As illustrated in Fig. 4 and Fig. 5, the bilateral ARMA model of order $n = 6$ computed for $N = 32$ outperforms the bilateral AR model with $n = 12$ which is obtained by fixing $N = 64$.

VI. Conclusions

In this paper we have taken a first step in generalizing the scalar theory of rational circulant covariance extension given

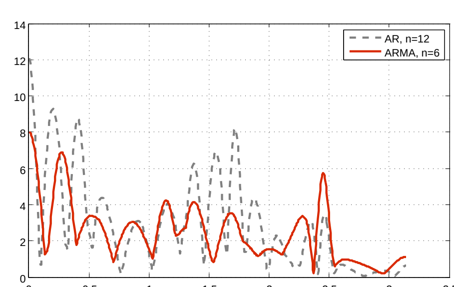

Fig. 4. Comparison between a bilateral AR of order 12 for $N = 64$ and a bilateral ARMA of order 6 for $N = 32$: norm of the approximation error.

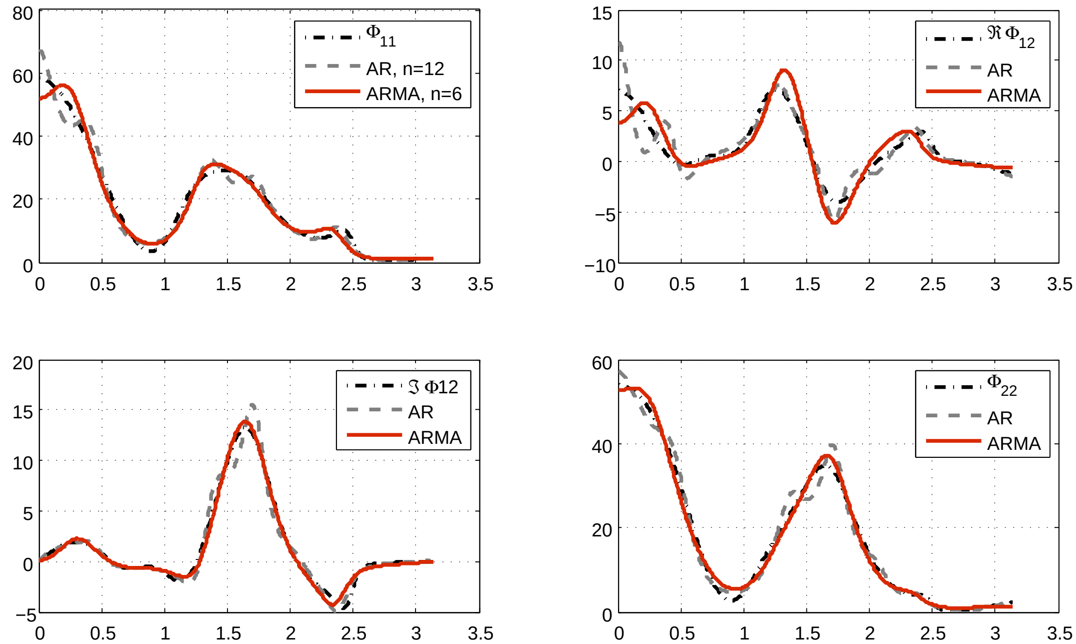

Fig. 5. Comparison between a bilateral AR of order 12 for $N = 64$ and a bilateral ARMA of order 6 for $N = 32$: estimated spectral densities.

in [26] to the multivariable case. Proofs of the theorems will be given in [27].

REFERENCES

[1] A. Blomqvist, A. Lindquist and R. Nagamune, Matrix-valued Nevanlinna-Pick interpolation with complexity constraint: An optimization approach, *IEEE Trans. Autom. Contr.* **48** (2003), 2172–2190.

[2] C. I. Byrnes, A. Lindquist, S.V. Gusev, and A. V. Matveev, A complete parameterization of all positive rational extensions of a covariance sequence, *IEEE Trans. Aut. Contr.* **AC-40** (1995) 1841-1857.

[3] C. I. Byrnes and A. Lindquist, On the partial stochastic realization problem, *IEEE Trans. Automatic Control* **AC-42** (1997), 1049–1069.

[4] C. I. Byrnes, S. V. Gusev, and A. Lindquist, A convex optimization approach to the rational covariance extension problem, *SIAM J. Control and Opt.* **37** (1999), 211-229.

[5] C. I. Byrnes, S.V. Gusev, and A. Lindquist, From finite covariance windows to modeling filters: A convex optimization approach, *SIAM Review* **43** (2001) 645–675.

[6] C. I. Byrnes, P. Enqvist, and A. Lindquist, Cepstral coefficients, covariance lags and pole-zero models for finite data strings, *IEEE Trans. on Signal Processing* **SP-50** (2001), 677–693.

[7] C. I. Byrnes, P. Enqvist, and A. Lindquist, Identifiability and well-posedness of shaping-filter parameterizations: A global analysis approach, *SIAM J. Control and Optimization* **41** (2002), 23–59.

[8] C. Byrnes, T.T. Georgiou, and A. Lindquist, A generalized entropy criterion for Nevanlinna-Pick interpolation: A convex optimization approach to certain problems in systems and control, *IEEE Trans. on Automatic Control*, **45** (2001), 822-839.

[9] C. I. Byrnes and A. Lindquist, The generalized moment problem with complexity constraint, *Int. Equations Operator Th.* **56** (2006) 163–180.

[10] C. I. Byrnes and A. Lindquist, Interior point solutions of variational problems and global inverse function theorems, *International Journal of Robust and Nonlinear Control* **17** (2007), 463–481.

[11] C. I. Byrnes and A. Lindquist, The moment problem for rational measures: convexity in the spirit of Krein, in *Modern Analysis and Application: Mark Krein Centenary Conference*, Vol. I: Operator Theory and Related Topics, Book Series: Operator Theory Advances and Applications Volume 190, Birkhäuser, 2009, pp. 157 – 169.

[12] F. P. Carli and A. Ferrante and M. Pavon and G. Picci, A Maximum Entropy Solution of the Covariance Extension Problem for Reciprocal Processes, *IEEE Trans. Automatic Control* **AC-56** (2011), 1999-2012.

[13] F. Carli, T.T. Georgiou, On the covariance completion problem under a circulant structure, *IEEE Trans. Autom. Contr* **56**(4) (2011), 918 -922.

[14] A. Chiuso and A. Ferrante and G. Picci, Reciprocal realization and modeling of textured images, *Proc. 44rd IEEE CDC*, 2005.

[15] P. Enqvist, *Spectral estimation by Geometric, Topological and Optimization Methods*, PhD thesis, KTH, Stockholm, Sweden, 2001.

[16] A. Ferrante and M. Pavon and F. Ramponi, Further results on the Byrnes-Georgiou-Lindquist generalized moment problem, *Modeling, Estimation and Control: Festschrift in honor of Giorgio Picci on the occasion of his sixty-fifth Birthday*, Springer-Verlag, 2007.

[17] T.T. Georgiou, *Partial Realization of Covariance Sequences*, Ph.D. thesis, CMST, University of Florida, Gainesville 1983.

[18] T.T. Georgiou, Realization of power spectra from partial covariances, *IEEE Trans. Acoustics, Speech Signal Proc.* **ASSP-35** (1987) 438-449.

[19] T.T. Georgiou, Solution of the general moment problem via a one-parameter imbedding, *IEEE Trans. Aut. Contr.* **AC-50** (2005) 811-826.

[20] T. T. Georgiou and A. Lindquist, Kullback-Leibler approximation of spectral density functions, *IEEE Trans. Inf. Th.* **49** (2003), 2910–2917.

[21] R. E. Kalman, Realization of Covariance Sequences, Proc. Toeplitz Memorial Conference, Tel Aviv, Israel, 1981.

[22] A.J. Krener, Reciprocal Processes and the stochastic realization problem for acausal systems, *Modeling Identification and Robust Control*, C.I. Byrnes and A. Lindquist, eds., North-Holland, 1986, 197–211.

[23] A. J. Krener, R. Frezza, and B. C. Levy, Gaussian reciprocal processes andself-adjoint differential equations of second order, *Stochastics and Stochastics Reports*, vol. 34, pp. 29-56, 1991.

[24] B. C. Levy and A. Ferrante, Characterization of stationary discrete-time Gaussian Reciprocal Processes over a finite interval, *SIAM J. Matrix Anal. Appl.* **24** (2002), 334-355.

[25] B. C. Levy and R. Frezza and A.J. Krener, Modeling and Estimation of discrete-time Gaussian Reciprocal Processes, *IEEE Trans. Automatic Control* **AC-35** (1990), 1013-1023.

[26] A. Lindquist and G. Picci, The circulant rational covariance extension problem: the complete solution, *IEEE Trans. Automatic Control* **58** (Nov. 2013).

[27] A. Lindquist, C. Masiero and G. Picci, Bilateral ARMA realization of multivariate, periodic, stationary stochastic processes, to appear.

[28] C. Masiero, forthcoming PhD thesis, University of Padova.

[29] B. R. Musicus and A. M. Kabel, Maximum entropy pole-zero estimation, Techn. Rpt 510, MIT Research Lab. Electronics, Aug. 1985.

[30] M. Pavon and . Ferrante, *On the Geometry of Maximum Entropy Problems*, provisionally accepted for publication in *SIAM REVIEW*, available in `http://arxiv.org/abs/1112.5529`, 2012.

[31] G. Picci and F. Carli, Modelling and simulation of images by reciprocal processes, *Proc. Tenth International Conference on Computer Modeling and Simulation UKSIM 2008*, 513–518.