Modeling of Stationary Periodic Time Series by ARMA Representations

Anders Lindquist and Giorgio Picci

Dedicated to Boris Teodorovich Polyak on the occasion of his 80th birthday

Abstract This is a survey of some recent results on the rational circulant covariance extension problem: Given a partial sequence $(c_0, c_1, \dots, c_n)$ of covariance lags $c_k = \mathbb{E}\{y(t+k)\overline{y(t)}\}$ emanating from a stationary periodic process $\{y(t)\}$ with period $2N > 2n$, find all possible rational spectral functions of $\{y(t)\}$ of degree at most $2n$ or, equivalently, all bilateral and unilateral ARMA models of order at most n, having this partial covariance sequence. Each representation is obtained as the solution of a pair of dual convex optimization problems. This theory is then reformulated in terms of circulant matrices and the connections to reciprocal processes and the covariance selection problem is explained. Next it is shown how the theory can be extended to the multivariate case. Finally, an application to image processing is presented.

Keywords Discrete moment problem • Periodic processes • Circulant covariance extension • Bilateral ARMA models • Image processing

1 Introduction

The rational covariance extension problem to determine a rational spectral density given a finite number of covariance lags has been studied in great detail [2, 5–7, 9, 10, 17, 19, 20, 24, 35], and it can be formulated as a (truncated) trigonometric moment problem with a degree constraint. Among other things, it is the basic problem in partial stochastic realization theory [2] and certain Toeplitz

A. Lindquist (✉)
Shanghai Jiao Tong University, Shanghai, China

Royal Institute of Technology, Stockholm, Sweden
e-mail: alq@kth.se

G. Picci
University of Padova, Padova, Italy
e-mail: picci@dei.unipd.it

B. Goldengorin (ed.), *Optimization and Its Applications in Control and Data Sciences*, Springer Optimization and Its Applications 115,
DOI 10.1007/978-3-319-42056-1_9

matrix completion problems. In particular, it provides a parameterization of the family of (unilateral) autoregressive moving-average (ARMA) models of stationary stochastic processes with the same finite sequence of covariance lags. We also refer the reader to the recent monograph [31], in which this problem is discussed in the context of stochastic realization theory.

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

$$\sum_{k=-n}^{n} q_k y(t-k) = \sum_{k=-n}^{n} p_k e(t-k)$$

for a stochastic processes $\{y(t)\}$, where $\{e(t)\}$ is the corresponding conjugate process. This connects up to a rich realization theory for reciprocal processes [26–29]. As we shall see there are also (forward and backward) unilateral ARMA representations for periodic processes.

In [12] a maximum-entropy approach to this circulant covariance extension problem was presented, providing a procedure for determining the unique bilateral AR model matching the covariance sequence. However, more recently it was discovered that the circulant covariance extension problem can be recast in the context of the optimization-based theory of moment problems with rational measures developed in [1, 3, 4, 6, 8–10, 21, 22] allowing for a complete parameterization of all bilateral ARMA realizations. This led to a complete theory for the scalar case [30], which was then extended to the multivariable case in [32]. Also see [38] for modifications of this theory to skew periodic processes and [37] for fast numerical procedures.

The AR theory of [12] has been successfully applied to image processing of textures [13, 36], and we anticipate an enhancement of such methods by allowing for more general ARMA realizations.

The present survey paper is to a large extent based on [30, 32] and [12]. In Sect. 2 we begin by characterizing stationary periodic processes. In Sect. 3 we formulate the rational covariance extension problem for periodic processes as a moment problem with atomic measure and present the solution in the context of the convex optimization approach of [1, 3, 4, 6, 8–10]. These results are then reformulated in terms of circulant matrices in Sect. 4 and interpreted in term of bilateral ARMA models in Sect. 5 and in terms of unilateral ARMA models in Sect. 6. In Sect. 7 we investigate the connections to reciprocal processes of order n [12] and the covariance selection problem of Dempster [15]. In Sect. 8 we consider the situation when both partial covariance data and logarithmic moment (cepstral) data is available. To simplify the exposition the theory has so far been developed in the context of scalar processes, but in Sect. 9 we show how it can be extended to the multivariable case. All of these results are illustrated by examples taken from [30] and [32]. Section 10 is devoted to applications in image processing.

2 Periodic Stationary Processes

Consider a zero-mean full-rank stationary process $\{y(t)\}$, in general complex-valued, 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 $2N$ so that

$$y(t + 2kN) = y(t) \tag{1}$$

almost surely. By stationarity there is a representation

$$y(t) = \int_{-\pi}^{\pi} e^{it\theta} d\hat{y}(\theta), \qquad \text{where } \mathbb{E}\{|d\hat{y}|^2\} = dF(\theta), \tag{2}$$

(see, e.g., [31, p. 74]), and therefore

$$c_k := \mathbb{E}\{y(t+k)\overline{y(t)}\} = \int_{-\pi}^{\pi} e^{ik\theta} dF(\theta). \tag{3}$$

Also, in view of (1),

$$\int_{-\pi}^{\pi} e^{it\theta} \left(e^{i2N\theta} - 1\right) d\hat{y} = 0,$$

and hence

$$\int_{-\pi}^{\pi} \left|e^{i2N\theta} - 1\right|^2 dF = 0,$$

which shows that the support of dF must be contained in $\{k\pi/N;\ k = -N+1, \dots, N\}$. Consequently the spectral density of $\{y(t)\}$ consists of point masses on the discrete unit circle $\mathbb{T}_{2N} := \{\zeta_{-N+1}, \zeta_{-n+2}, \dots, \zeta_N\}$, where

$$\zeta_k = e^{ik\pi/N}. \tag{4}$$

More precisely, define the function

$$\Phi(\zeta) = \sum_{k=-N+1}^{N} c_k \zeta^{-k} \tag{5}$$

on $\mathbb{T}_{2N}$. This is the discrete Fourier transform (DFT) of the sequence $(c_{-N+1}, \dots, c_N)$, which can be recovered by the inverse DFT

$$c_k = \frac{1}{2N} \sum_{j=-N+1}^{N} \Phi(\zeta_j)\zeta_j^k = \int_{-\pi}^{\pi} e^{ik\theta} \Phi(e^{i\theta}) d\nu, \tag{6}$$

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

$$d\nu(\theta) = \sum_{j=-N+1}^{N} \delta(e^{i\theta} - \zeta_j)\frac{d\theta}{2N}. \tag{7}$$

Consequently, by (3), $dF(\theta) = \Phi(e^{i\theta})d\nu(\theta)$. We note in passing that

$$\int_{-\pi}^{\pi} e^{ik\theta} d\nu(\theta) = \delta_{k0}, \tag{8}$$

where δ_{k0} equals one for $k = 0$ and zero otherwise. To see this, note that, for $k \neq 0$,

$$\begin{aligned}(1-\zeta_k)\int_{-\pi}^{\pi} e^{ik\theta} d\nu &= \frac{1}{2N}\sum_{j=-N+1}^{N}\left(\zeta_k^j - \zeta_k^{j+1}\right)\\ &= \frac{1}{2N}\left(\zeta_k^{-N+1} - \zeta_k^{N+1}\right) = 0.\end{aligned}$$

Since $\{y(t)\}$ is stationary and full rank, the Toeplitz matrix

$$\mathbf{T}_n = \begin{bmatrix} c_0 & \bar{c}_1 & \bar{c}_2 & \cdots & \bar{c}_n \\ c_1 & c_0 & \bar{c}_1 & \cdots & \bar{c}_{n-1} \\ c_2 & c_1 & c_0 & \cdots & \bar{c}_{n-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_n & c_{n-1} & c_{n-2} & \cdots & c_0 \end{bmatrix} \tag{9}$$

is positive definite for all $n \in \mathbb{Z}$. However, this condition is not sufficient for $c_0, c_1, \dots, c_n$ to be a bona-fide covariance sequence of a periodic process, as can be seen from the following simple example. Consider a real-valued periodic stationary process y of period four. Then

$$\mathbb{E}\left\{\begin{bmatrix} y(1) \\ y(2) \\ y(3) \\ y(4) \end{bmatrix}\begin{bmatrix} y(1) & y(2) & y(3) & y(4) \end{bmatrix}\right\} = \begin{bmatrix} c_0 & c_1 & c_2 & c_3 \\ c_1 & c_0 & c_1 & c_2 \\ c_2 & c_1 & c_0 & c_1 \\ c_3 & c_2 & c_1 & c_0 \end{bmatrix}.$$

Then looking at the covariance matrix for two periods, we obtain

$$\mathbb{E}\left\{\begin{bmatrix} y(1) \\ y(2) \\ \vdots \\ y(8) \end{bmatrix} \begin{bmatrix} y(1) \; y(2) \cdots y(8) \end{bmatrix}\right\} = \begin{bmatrix} c_0 & c_1 & c_2 & c_3 & c_0 & c_1 & c_2 & c_3 \\ c_1 & c_0 & c_1 & c_2 & c_1 & c_0 & c_1 & c_2 \\ c_2 & c_1 & c_0 & c_1 & c_2 & c_1 & c_0 & c_1 \\ c_3 & c_2 & c_1 & c_0 & c_3 & c_2 & c_1 & c_0 \\ c_0 & c_1 & c_2 & c_3 & c_0 & c_1 & c_2 & c_3 \\ c_1 & c_0 & c_1 & c_2 & c_1 & c_0 & c_1 & c_2 \\ c_2 & c_1 & c_0 & c_1 & c_2 & c_1 & c_0 & c_1 \\ c_3 & c_2 & c_1 & c_0 & c_3 & c_2 & c_1 & c_0 \end{bmatrix},$$

which is a Toeplitz matrix only when $c_3 = c_1$. Therefore the condition $c_3 = c_1$ is necessary. Consequently

$$\mathbf{T}_8 = \begin{bmatrix} c_0 & c_1 & c_2 & c_1 & c_0 & c_1 & c_2 & c_1 \\ c_1 & c_0 & c_1 & c_2 & c_1 & c_0 & c_1 & c_2 \\ c_2 & c_1 & c_0 & c_1 & c_2 & c_1 & c_0 & c_1 \\ c_1 & c_2 & c_1 & c_0 & c_1 & c_2 & c_1 & c_0 \\ c_0 & c_1 & c_2 & c_1 & c_0 & c_1 & c_2 & c_1 \\ c_1 & c_0 & c_1 & c_2 & c_1 & c_0 & c_1 & c_2 \\ c_2 & c_1 & c_0 & c_1 & c_2 & c_1 & c_0 & c_1 \\ c_1 & c_2 & c_1 & c_0 & c_1 & c_2 & c_1 & c_0 \end{bmatrix}$$

is a *circulant matrix*, where the columns are shifted cyclically, the last component moved to the top. Circulant matrices will play a key role in the following.

3 The Covariance Extension Problem for Periodic Processes

Suppose that we are given a partial covariance sequence $c_0, c_1, \dots, c_n$ with $n < N$ such that the Toeplitz matrix $\mathbf{T}_n$ is positive definite. Consider the problem of finding and extension $c_{n+1}, c_{n+2}, \dots, c_N$ so that the corresponding sequence $c_0, c_1, \dots, c_N$ is the covariance sequence of a stationary process of period $2N$.

In general this problem will have infinitely many solutions, and, for reasons that will become clear later, we shall restrict our attention to spectral functions (5) which are rational in the sense that

$$\Phi(\zeta) = \frac{P(\zeta)}{Q(\zeta)}, \tag{10}$$

where P and Q are Hermitian pseudo-polynomials of degree at most n, that is of the form

$$P(\zeta) = \sum_{k=-n}^{n} p_k \zeta^{-k}, \quad p_{-k} = \bar{p}_k. \tag{11}$$

Let $\mathfrak{P}_+(N)$ be the cone of all pseudo-polynomials (11) that are positive on the discrete unit circle $\mathbb{T}_{2N}$, and let $\mathfrak{P}_+ \subset \mathfrak{P}_+(N)$ be the subset of pseudo-polynomials (11) such that $P(e^{i\theta}) > 0$ for all $\theta \in [-\pi, \pi]$. Moreover let $\mathfrak{C}_+(N)$ be the dual cone of all partial covariance sequences $\mathbf{c} = (c_0, c_1, \dots, c_n)$ such that

$$\langle \mathbf{c}, \mathbf{p} \rangle := \sum_{k=-n}^{n} c_k \bar{p}_k > 0 \quad \text{for all } P \in \overline{\mathfrak{P}_+(N)} \setminus \{0\},$$

and let $\mathfrak{C}_+$ be defined in the same way as the dual cone of $\mathfrak{P}_+$. It can be shown [25] that $\mathbf{c} \in \mathfrak{C}_+$ is equivalent to the Toeplitz condition $\mathbf{T}_n > 0$. Since $\mathfrak{P}_+ \subset \mathfrak{P}_+(N)$, we have $\mathfrak{C}_+(N) \subset \mathfrak{C}_+$, so in general $\mathbf{c} \in \mathfrak{C}_+(N)$ is a stricter condition than $\mathbf{T}_n > 0$.

The proof of the following theorem can be found in [30].

Theorem 1. *Let $\mathbf{c} \in \mathfrak{C}_+(N)$. Then, for each $P \in \mathfrak{P}_+(N)$, there is a unique $Q \in \mathfrak{P}_+(N)$ such that*

$$\Phi = \frac{P}{Q}$$

satisfies the moment conditions

$$\int_{-\pi}^{\pi} e^{ik\theta} \Phi(e^{i\theta}) d\nu(\theta) = c_k, \quad k = 0, 1, \dots, n. \tag{12}$$

Consequently the family of solutions (10) of the covariance extension problem stated above are parameterized by $P \in \mathfrak{P}_+(N)$ in a bijective fashion. From the following theorem we see that, for any $P \in \mathfrak{P}_+(N)$, the corresponding unique $Q \in \mathfrak{P}_+(N)$ can be obtained by convex optimization. We refer the reader to [30] for the proofs.

Theorem 2. *Let $\mathbf{c} \in \mathfrak{C}_+(N)$ and $P \in \mathfrak{P}_+(N)$. Then the problem to maximize*

$$\mathbb{I}_P(\Phi) = \int_{-\pi}^{\pi} P(e^{i\theta}) \log \Phi(e^{i\theta}) d\nu \tag{13}$$

subject to the moment conditions (12) has a unique solution, namely (10), where Q is the unique optimal solution of the problem to minimize

$$\mathbb{J}_P(Q) = \langle \mathbf{c}, \mathbf{q} \rangle - \int_{-\pi}^{\pi} P(e^{i\theta}) \log Q(e^{i\theta}) d\nu \tag{14}$$

over all $Q \in \mathfrak{P}_+(N)$, where $\mathbf{q} := (q_0, q_1, \dots, q_n)$. The functional $\mathbb{J}_P$ is strictly convex.

Theorems 1 and 2 are discrete versions of corresponding results in [6, 9]. The solution corresponding to $P = 1$ is called the *maximum-entropy solution* by virtue of (13).

Remark 3. As $N \to \infty$ the process y looses it periodic character, and its spectral density Φ_∞ becomes continuous and defined on the whole unit circle so that

$$\int_{-\pi}^{\pi} e^{ik\theta} \Phi_\infty(e^{i\theta}) \frac{d\theta}{2\pi} = c_k, \quad k = 0, 1, \ldots, n. \tag{15}$$

In fact, denoting by Q_N the solution of Theorem 1, it was shown in [30] that $\Phi_\infty = P/Q_\infty$, where, for each fixed P,

$$Q_\infty = \lim_{N \to \infty} Q_N$$

is the unique Q such that $\Phi_\infty = P/Q$ satisfies the moment conditions (15).

4 Reformulation in Terms of Circulant Matrices

Circulant matrices [14] are Toeplitz matrices with a special circulant structure

$$\mathrm{Circ}\{\gamma_0, \gamma_1, \ldots, \gamma_\nu\} = \begin{bmatrix} \gamma_0 & \gamma_\nu & \gamma_{\nu-1} & \cdots & \gamma_1 \\ \gamma_1 & \gamma_0 & \gamma_\nu & \cdots & \gamma_2 \\ \gamma_2 & \gamma_1 & \gamma_0 & \cdots & \gamma_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \gamma_\nu & \gamma_{\nu-1} & \gamma_{\nu-2} & \cdots & \gamma_0 \end{bmatrix}, \tag{16}$$

where the columns (or, equivalently, rows) are shifted cyclically, and where $\gamma_0, \gamma_1, \ldots, \gamma_\nu$ here are taken to be complex numbers. In our present covariance extension problem we consider *Hermitian* circulant matrices

$$\mathbf{M} := \mathrm{Circ}\{m_0, m_1, m_2, \ldots, m_N, \bar{m}_{N-1}, \ldots, \bar{m}_2, \bar{m}_1\}, \tag{17}$$

which can be represented in form

$$\mathbf{M} = \sum_{k=-N+1}^{N} m_k \mathbf{S}^{-k}, \quad m_{-k} = \bar{m}_k \tag{18}$$

where $\mathbf{S}$ is the nonsingular $2N \times 2N$ cyclic shift matrix

$$\mathbf{S} := \begin{bmatrix} 0 & 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}. \tag{19}$$

The pseudo-polynomial

$$M(\zeta) = \sum_{k=-N+1}^{N} m_k \zeta^{-k}, \quad m_{-k} = \bar{m}_k \tag{20}$$

is called the *symbol* of $\mathbf{M}$. Clearly $\mathbf{S}$ is itself a circulant matrix (although not Hermitian) with symbol $S(\zeta) = \zeta$. A necessary and sufficient condition for a matrix $\mathbf{M}$ to be circulant is that

$$\mathbf{S}\mathbf{M}\mathbf{S}^{\mathsf{T}} = \mathbf{M}. \tag{21}$$

Hence, since $\mathbf{S}^{-1} = \mathbf{S}^{\mathsf{T}}$, the inverse of a circulant matrix is also circulant. More generally, if $\mathbf{A}$ and $\mathbf{B}$ are circulant 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—more precisely, a commutative *-algebra with the involution * being the conjugate transpose—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 $\zeta \in \mathbb{T}_{2N}$. Consequently, circulant matrices commute, and, if $\mathbf{M}$ is a circulant matrix with symbol $M(\zeta)$, then $\mathbf{M}^{-1}$ is circulant with symbol $M(\zeta)^{-1}$.

The proof of the following proposition is immediate.

Proposition 4. *Let $\{y(t);\ t = -N+1, \dots, N\}$ be a stationary process with period $2N$ and covariance lags (3), and let $\mathbf{y}$ be the $2N$-dimensional stochastic vector $\mathbf{y} = [y(-N+1), y(-N+2), \cdots, y(N)]^{\mathsf{T}}$. Then, with * denoting conjugate transpose,*

$$\boldsymbol{\Sigma} := \mathbb{E}\{\mathbf{y}\mathbf{y}^*\} = \operatorname{Circ}\{c_0, c_1, c_2, \dots, c_N, \bar{c}_{N-1}, \dots, \bar{c}_2, \bar{c}_1\} \tag{22}$$

is a $2N \times 2N$ Hermitian circulant matrix with symbol $\Phi(\zeta)$ given by (5).

The covariance extension problem of Sect. 3, called the *circulant rational covariance extension problem*, can now be reformulated as a matrix extension problem. The given covariance data $\mathbf{c} = (c_0, c_1, \dots, c_n)$ can be represented as a circulant matrix

$$\mathbf{C} = \operatorname{Circ}\{c_0, c_1, \dots, c_n, 0, \dots, 0, \bar{c}_n, \bar{c}_{n-1}, \dots, \bar{c}_1\} \tag{23}$$

with symbol

$$C(\zeta) = \sum_{k=-n}^{n} c_k \zeta^{-k}, \tag{24}$$

where the unknown covariance lags $c_{n+1}, c_{n+2}, \dots, c_N$ in (22), to be determined, here are replaced by zeros. A circulant matrix of type (23) is called *banded of order n*. We recall that $n < N$. From now one we drop the attribute ‘Hermitian’ since we shall only consider such circulant matrices in the sequel. A banded circulant matrix of order n will thus be determined by $n + 1$ (complex) parameters.

The next lemma establishes the connection between circulant matrices and their symbols.

Lemma 5. *Let* $\mathbf{M}$ *be a circulant matrix with symbol* $M(\zeta)$*. Then*

$$\mathbf{M} = \mathbf{F}^* \mathrm{diag}\big(M(\zeta_{-N+1}), M(\zeta_{-N+2}), \dots, M(\zeta_N)\big)\mathbf{F}, \tag{25}$$

where $\mathbf{F}$ *is the unitary matrix*

$$\mathbf{F} = \frac{1}{\sqrt{2N}} \begin{bmatrix} \zeta_{-N+1}^{N-1} & \zeta_{-N+1}^{N-2} & \cdots & \zeta_{-N+1}^{-N} \\ \vdots & \vdots & \cdots & \vdots \\ \zeta_0^{N-1} & \zeta_0^{N-2} & \cdots & \zeta_0^{-N} \\ \vdots & \vdots & \cdots & \vdots \\ \zeta_N^{N-1} & \zeta_N^{N-2} & \cdots & \zeta_N^{-N} \end{bmatrix}. \tag{26}$$

Moreover, if $M(\zeta_k) > 0$ *for all k, then*

$$\log \mathbf{M} = \mathbf{F}^* \mathrm{diag}\big(\log M(\zeta_{-N+1}), \log M(\zeta_{-N+2}), \dots, \log M(\zeta_N)\big)\mathbf{F}. \tag{27}$$

Proof. The discrete Fourier transform $\mathcal{F}$ maps a sequence $(g_{-N+1}, g_{-N+2}, \dots, g_N)$ into the sequence of complex numbers

$$G(\zeta_j) := \sum_{k=-N+1}^{N} g_k \zeta_j^{-k}, \qquad j = -N+1, -N+2, \dots, N. \tag{28}$$

The sequence $\mathbf{g}$ can be recovered from G by the inverse transform

$$g_k = \int_{-\pi}^{\pi} e^{ik\theta} G(e^{i\theta}) d\nu(\theta), \quad k = -N+1, -N+2, \dots, N. \tag{29}$$

This correspondence can be written

$$\hat{\mathbf{g}} = \mathbf{F}\mathbf{g}, \tag{30}$$

where $\hat{\mathbf{g}} := (2N)^{-\frac{1}{2}}\big(G(\zeta_{-N+1}),\dots,G(\zeta_N)\big)^{\mathsf{T}}$, $\mathbf{g} := (g_{-N+1},\dots,g_N)^{\mathsf{T}}$, and $\mathbf{F}$ is the nonsingular $2N\times 2N$ Vandermonde matrix (26). Clearly $\mathbf{F}$ is unitary. Since

$$\mathbf{Mg} = \sum_{k=-N+1}^{N} m_k \mathbf{S}^{-k}$$

and $[\mathbf{S}^{-k}\mathbf{g}]_j = g_{j-k}$, where $g_{k+2N} = g_k$, we have

$$\begin{aligned}\mathfrak{F}(\mathbf{Mg}) &= \sum_{j=-N+1}^{N} \zeta^{-j} \sum_{k=-N+1}^{N} m_k g_{j-k}\\ &= \sum_{k=-N+1}^{N} m_k\zeta^{-k} \sum_{j=-N+1}^{N} g_{j-k}\zeta^{-(j-k)} = M(\zeta)\mathfrak{F}\mathbf{g},\end{aligned}$$

which yields

$$\sqrt{2N}(\mathbf{FMg})_j = M(\zeta_j)\sqrt{2N}(\mathbf{Fg})_j, \quad j=-N+1,-N+2,\dots,N,$$

from which (25) follows. Finally, since, as a function of $z\in\mathbb{C}$, $\log M(z)$ is analytic in the neighborhood of each $M(\zeta_k) > 0$, the eigenvalues of $\log\mathbf{M}$ are just the real numbers $\log M(\zeta_k)$, $k = -N+1,\dots,N$, by the spectral mapping theorem [16, p. 557], and hence (27) follows.

We are now in a position to reformulate Theorems 1 and 2 in terms of circulant matrices. To this end first note that, in view of Lemma 5, the cone $\mathfrak{P}_+(N)$ corresponds to the class of positive-definite banded $2N\times 2N$ circulant matrices $\mathbf{P}$ of order n. Moreover, by Plancherel's Theorem for DFT, which is a simple consequence of (8), we have

$$\sum_{k=-n}^{n} c_k\bar{p}_k = \frac{1}{2N}\sum_{j=-N+1}^{N} C(\zeta_j)P(\zeta_j),$$

and hence, by Lemma 5,

$$\langle \mathbf{c},\mathbf{p}\rangle = \frac{1}{2N}\operatorname{tr}(\mathbf{CP}), \tag{31}$$

where tr denotes trace.

Consequently, $\mathbf{c}\in\mathfrak{C}_+(N)$ if and only if $\operatorname{tr}(\mathbf{CP}) > 0$ for all nonzero, positive-semidefinite, banded $2N\times 2N$ circulant matrices $\mathbf{P}$ of order n. Moreover, if $\mathbf{Q}$ and $\mathbf{P}$ are circulant matrices with symbols $P(\zeta)$ and $Q(\zeta)$, respectively, then, by Lemma 5, $P(\zeta)/Q(\zeta)$ is the symbol of $\mathbf{Q}^{-1}\mathbf{P}$. Therefore Theorem 1 has the following matrix version.

Theorem 6. *Let* $\mathbf{c} \in \mathfrak{C}_+(N)$*, and let* $\mathbf{C}$ *be the corresponding circulant matrix* (23)*. Then, for each positive-definite banded* $2N \times 2N$ *circulant matrices* $\mathbf{P}$ *of order* n*, there is unique positive-definite banded* $2N \times 2N$ *circulant matrices* $\mathbf{Q}$ *of order* n *such that*

$$\boldsymbol{\Sigma} = \mathbf{Q}^{-1}\mathbf{P} \tag{32}$$

is a circulant extension (22) *of* $\mathbf{C}$*.*

In the same way, Theorem 2 has the following matrix version, as can be seen by applying Lemma 5.

Theorem 7. *Let* $\mathbf{c} \in \mathfrak{C}_+(N)$*, and let* $\mathbf{C}$ *be the corresponding circulant matrix* (23)*. Moreover, let* $\mathbf{P}$ *be a positive-definite banded* $2N \times 2N$ *circulant matrix of order* n*. Then the problem to maximize*

$$\mathscr{I}_{\mathbf{P}}(\boldsymbol{\Sigma}) = \operatorname{tr}(\mathbf{P} \log \boldsymbol{\Sigma}) \tag{33}$$

subject to

$$\mathbf{E}_n{}^T \boldsymbol{\Sigma} \mathbf{E}_n = \mathbf{T}_n, \quad \text{where } \mathbf{E}_n = \begin{bmatrix} \mathbf{I}_n \\ \mathbf{0} \end{bmatrix} \tag{34}$$

has a unique solution, namely (32)*, where* $\mathbf{Q}$ *is the unique optimal solution of the problem to minimize*

$$\mathscr{J}_{\mathbf{P}}(\mathbf{q}) = \operatorname{tr}(\mathbf{C}\mathbf{Q}) - \operatorname{tr}(\mathbf{P} \log \mathbf{Q}) \tag{35}$$

over all positive-definite banded $2N \times 2N$ *circulant matrices* $\mathbf{Q}$ *of order* n*, where* $\mathbf{q} := (q_0, q_1, \dots, q_n)$*. The functional* $\mathscr{J}_{\mathbf{P}}$ *is strictly convex.*

5 Bilateral ARMA Models

Suppose now that we have determined a circulant matrix extension (32). Then there is a stochastic vector $\mathbf{y}$ formed from the a stationary periodic process with corresponding covariance lags (3) so that

$$\boldsymbol{\Sigma} := \mathbb{E}\{\mathbf{y}\mathbf{y}^*\} = \operatorname{Circ}\{c_0, c_1, c_2, \dots, c_N, \bar{c}_{N-1}, \dots, \bar{c}_2, \bar{c}_1\}.$$

Let $\hat{\mathbb{E}}\{y(t) \mid y(s),\, s \neq t\}$ be the wide sense conditional mean of $y(t)$ given all $\{y(s),\, s \neq t\}$. Then the error process

$$d(t) := y(t) - \hat{\mathbb{E}}\{y(t) \mid y(s),\, s \neq t\} \tag{36}$$

is orthogonal to all random variables $\{y(s),\ s \neq t\}$, i.e., $\mathbb{E}\{y(t)\,\overline{d(s)}\} = \sigma^2\,\delta_{ts}$, $t, s \in \mathbb{Z}_{2N} := \{-N+1, -N+2, \dots, N\}$, where σ^2 is a positive number. Equivalently, $\mathbb{E}\{\mathbf{y}\mathbf{d}^*\} = \sigma^2\mathbf{I}$, where $\mathbf{I}$ is the $2N \times 2N$ identity matrix. Setting $\mathbf{e} := \mathbf{d}/\sigma^2$, we then have

$$\mathbb{E}\{\mathbf{e}\mathbf{y}^*\} = \mathbf{I}, \tag{37}$$

i.e., the corresponding process e is the *conjugate process* of y [33]. Interpreting (36) in the mod $2N$ arithmetics of $\mathbb{Z}_{2N}$, $\mathbf{y}$ admits a linear representation of the form

$$\mathbf{G}\mathbf{y} = \mathbf{e}, \tag{38}$$

where $\mathbf{G}$ is a $2N \times 2N$ Hermitian circulant matrix with ones on the main diagonal. Since $\mathbf{G}\mathbb{E}\{\mathbf{y}\mathbf{y}^*\} = \mathbb{E}\{\mathbf{e}\mathbf{y}^*\} = \mathbf{I}$, $\mathbf{G}$ is also positive definite and the covariance matrix $\boldsymbol{\Sigma}$ is given by

$$\boldsymbol{\Sigma} = \mathbf{G}^{-1}, \tag{39}$$

which is circulant, since the inverse of a circulant matrix is itself circulant. In fact, a stationary process $\mathbf{y}$ is full-rank periodic in $\mathbb{Z}_{2N}$, if and only if $\boldsymbol{\Sigma}$ is a Hermitian positive definite circulant matrix [12].

Since $\mathbf{G}$ is a Hermitian circulant matrix, it has a symbol

$$G(\zeta) = \sum_{k=-N+1}^{N} g_k \zeta^{-k}, \quad g_{-k} = \bar{g}_k,$$

and the linear equation can be written in the autoregressive (AR) form

$$\sum_{k=-N+1}^{N} g_k y(t-k) = e(t). \tag{40}$$

However, in general $\mathbf{G}$ is not banded and $n << N$, and therefore (40) is not a parsimonious representation. Instead using the solution (32), we have $\mathbf{G} = \mathbf{P}^{-1}\mathbf{Q}$, where $\mathbf{P}$ and $\mathbf{Q}$ are banded of order n with symbols

$$P(\zeta) = \sum_{k=-n}^{n} p_k \zeta^{-k} \quad \text{and} \quad Q(\zeta) = \sum_{k=-n}^{n} q_k \zeta^{-k},$$

and hence (38) can be written

$$\mathbf{Q}\mathbf{y} = \mathbf{P}\mathbf{e},$$

or equivalently in the ARMA form

$$\sum_{k=-n}^{n} q_k y(t-k) = \sum_{k=-n}^{n} p_k e(t-k). \tag{41}$$

Consequently, by Theorem 6, there is a unique bilateral ARMA model (41) for each banded positive-definite Hermitian circulant matrix $\mathbf{P}$ of order n, provided $\mathbf{c} \in \mathfrak{C}_+(N)$. Of course, we could use the maximum-entropy solution with $\mathbf{P} = \mathbf{I}$ leading to an AR model

$$\sum_{k=-n}^{n} q_k y(t-k) = e(t). \tag{42}$$

Next, to illustrate the accuracy of bilateral AR modeling by the methods described so far we give some simulations from [30], provided by Chiara Masiero. Given an AR model of order $n = 8$ with poles as depicted in Fig. 1, we compute a covariance sequence $\mathbf{c} = (c_0, c_1, \dots, c_n)$ with $n = 8$, which is then used to solve the optimization problem (35) with $\mathbf{P} = \mathbf{I}$ to obtain a bilateral AR approximations of degree eight for various choices of N. In Fig. 2, the top picture depicts the spectral density for $N = 128$ together with the true spectral density (dashed line), and the bottom picture illustrates how the estimation error decreases with increasing N.

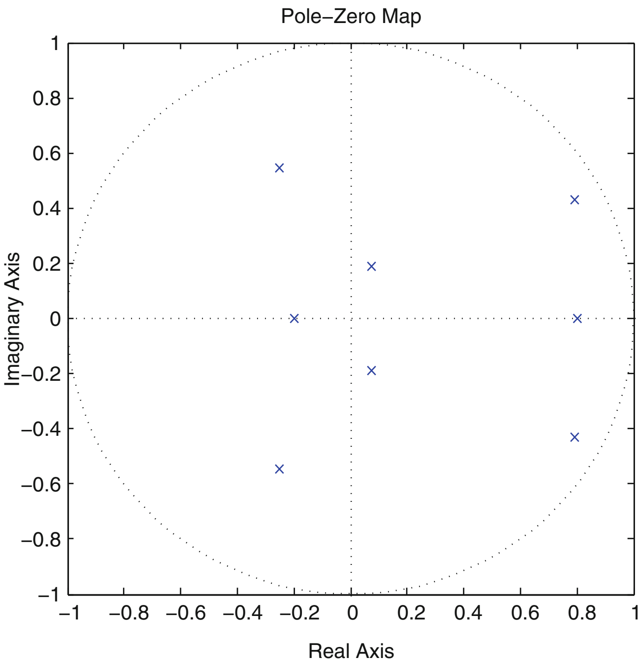


Fig. 1 Poles of true AR model

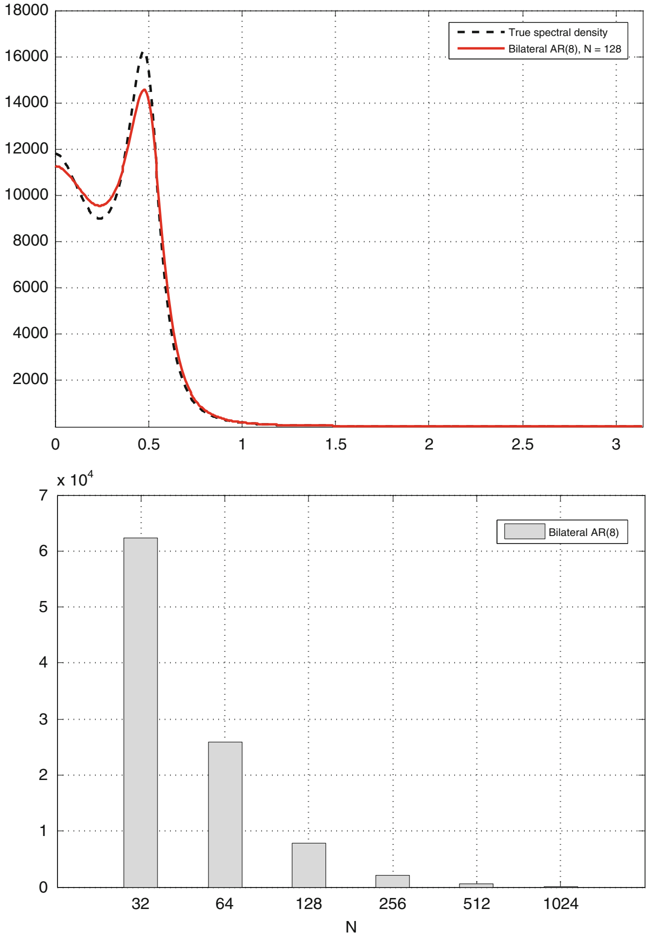


Fig. 2 Bilateral AR approximation: (*top*) spectrum for N = 128 and true spectrum (*dashed*); (*bottom*) errors for N=32, 64, 128, 256, 512 and 1024

6 Unilateral ARMA Models and Spectral Factorization

As explained in Sect. 2, a periodic process y has a discrete spectrum, and Theorem 1 provides values of

$$\Phi(z) = \frac{P(z)}{Q(z)}$$

only in the discrete points $z \in \mathbb{T}_{2N} := \{\zeta_{-N+1}, \zeta_{-n+2}, \dots, \zeta_N\}$. Since Φ takes positive values on $\mathbb{T}_{2N}$, there is a trivial discrete factorization

$$\Phi(\zeta_k) = W(\zeta_k)W(\zeta_k)^* \quad k = -N+1, \dots, N. \tag{43}$$

Defining

$$W_k = \frac{1}{2N} \sum_{j=-N+1}^{N} W(\zeta_j)\zeta_j^k, \quad k = -N+1, \dots, N,$$

we can write (43) in the form

$$\Phi(\zeta) = W(\zeta)W(\zeta)^*, \tag{44}$$

where $W(\zeta)$ is the discrete Fourier transform

$$W(\zeta) = \sum_{k=-N+1}^{N} W_k \zeta^{-k}.$$

Formally substituting the variable $z \in \mathbb{T}$ in place of ζ in W, we obtain a spectral factorization equation

$$\tilde{\Phi}(z) = W(z)W(z)^*, \quad z \in \mathbb{T}, \tag{45}$$

defined on the whole unit circle, where the continuous spectral density $\tilde{\Phi}(z)$, frequency sampled with sampling interval $\frac{\pi}{N}$, satisfies $\tilde{\Phi}(\zeta) = \Phi(\zeta)$ on $\mathbb{T}_{2N}$. This is a spectral density of a non-periodic stationary process but should not be confused with Φ_∞ in Remark 3, which is the unique continuous Φ with numerator polynomial P and the same first $n+1$ covariance lags as the periodic process y, i.e.,

$$\int_{-\pi}^{\pi} e^{ik\theta} \Phi_\infty(e^{i\theta}) \frac{d\theta}{2\pi} = c_k, \quad k = 0, 1, \dots, n.$$

In fact, although

$$\int_{-\pi}^{\pi} e^{ik\theta} \tilde{\Phi}(e^{i\theta}) d\nu(\theta) = c_k, \quad k = 0, 1, \dots, n, \tag{46}$$

the non-periodic process with spectral density $\tilde{\Phi}$ has the covariance lags

$$\tilde{c}_k = \int_{-\pi}^{\pi} e^{ik\theta} \tilde{\Phi}(e^{i\theta}) \frac{d\theta}{2\pi}, \quad k = 0, 1, \dots, n,$$

which differ from $c_0, c_1, \dots, c_n$. However, setting $\Delta\theta_j := \theta_j - \theta_{j-1}$ where $e^{i\theta_j} = \zeta_j$, we see from (4) that $\Delta\theta_j = \pi/N$ and that the integral (46) with $\tilde{\Phi}$ fixed is the Riemann sum

$$\sum_{j=-N+1}^{N} e^{ik\theta_j} \tilde{\Phi}(\zeta_j) \frac{\Delta\theta_j}{2\pi}$$

converging to $\tilde{c}_k$ for $k = 0, 1, \dots, n$ as $N \to \infty$.

By Proposition 4, $\Phi(\zeta)$ is the symbol of the circulant covariance matrix $\boldsymbol{\Sigma}$, and hence (44) can be written in the matrix form

$$\boldsymbol{\Sigma} = \mathbf{W}\mathbf{W}^*, \tag{47}$$

where $\mathbf{W}$ is the circulant matrix with symbol $W(\zeta)$. The spectral density (45) has a unique outer spectral factor $W(z)$; see, e.g., [31]. As explained in detail in [11], this corresponds in the discrete setting to $W(\zeta)$ taking the form

$$W(\zeta) = \sum_{k=0}^{N} W_k \zeta^{-k}, \tag{48}$$

which in turn corresponds to $\mathbf{W}$ being *lower-triangular circulant*, i.e.,

$$\mathbf{W} = \mathrm{Circ}\{W_0, W_1, \ldots, W_N, 0, \ldots, 0\}. \tag{49}$$

Note that a lower-triangular circulant matrix is not lower triangular as the circulant structure has to be preserved. Since $\boldsymbol{\Sigma}$ is invertible, then so is $\mathbf{W}$.

Next define the periodic stochastic process $\{w(t),\ t = -N + 1 \ldots, N\}$ for which $\mathbf{w} = [w(-N+1), w(-N+2), \ldots, w(N)]^{\mathsf{T}}$ is given by

$$\mathbf{w} = \mathbf{W}^{-1}\mathbf{y}. \tag{50}$$

Then, in view of (47), we obtain $\mathbb{E}\{\mathbf{w}\mathbf{w}^*\} = \mathbf{I}$, i.e., the process w is a white noise process. Consequently we have the unilateral representation

$$y(t) = \sum_{k=0}^{N} W_k w(t-k)$$

in terms of white noise.

To construct an ARMA model we appeal to the following result, which is easy to verify in terms of symbols but, as demonstrated in [11], also holds for block circulant matrices considered in Sect. 9.

Lemma 8. *There exists an integer N_0 such that the following holds for $N \geq N_0$. A positive definite, Hermitian, circulant matrix $\mathbf{M}$ admits a factorization $\mathbf{M} = \mathbf{V}\mathbf{V}^*$, where $\mathbf{V}$ is of a banded lower-diagonal circulant matrix of order $n < N$, if and only if $\mathbf{M}$ is bilaterally banded of order n.*

By Theorem 6, $\boldsymbol{\Sigma} = \mathbf{Q}^{-1}\mathbf{P}$, where $\mathbf{Q}$ and $\mathbf{P}$ are banded, positive definite, Hermitian, circulant matrices of order n. Hence, for N sufficiently large, by Lemma 8 there are factorizations

$$\mathbf{Q} = \mathbf{A}\mathbf{A}^* \quad \text{and} \quad \mathbf{P} = \mathbf{B}\mathbf{B}^*,$$

where $\mathbf{A}$ and $\mathbf{B}$ are banded lower-diagonal circulant matrices of order n. Consequently, $\boldsymbol{\Sigma} = \mathbf{A}^{-1}\mathbf{B}(\mathbf{A}^{-1}\mathbf{B})^*$, i.e.,

$$\mathbf{W} = \mathbf{A}^{-1}\mathbf{B}, \tag{51}$$

which together with (50) yields $\mathbf{A}\mathbf{y} = \mathbf{B}\mathbf{w}$, i.e., the unilateral ARMA model

$$\sum_{k=0}^{n} a_k y(t-k) = \sum_{k=0}^{n} b_k w(t-k). \tag{52}$$

Since $\mathbf{A}$ is nonsingular, $a_0 \neq 0$, and hence we can normalize by setting $a_0 = 1$. In particular, if $\mathbf{P} = \mathbf{I}$, we obtain the AR representation

$$\sum_{k=0}^{n} a_k y(t-k) = b_0 w(t). \tag{53}$$

Symmetrically, there is factorization

$$\boldsymbol{\Sigma} = \bar{\mathbf{W}}\bar{\mathbf{W}}^*, \tag{54}$$

where $\bar{\mathbf{W}}$ is upper-diagonal circulant, i.e. the transpose of a lower-diagonal circulant matrix, and a white-noise process

$$\bar{\mathbf{w}} = \bar{\mathbf{W}}^{-1}\mathbf{y}. \tag{55}$$

Likewise there are factorizations

$$\mathbf{Q} = \bar{\mathbf{A}}\bar{\mathbf{A}}^* \quad \text{and} \quad \mathbf{P} = \bar{\mathbf{B}}\bar{\mathbf{B}}^*,$$

where $\bar{\mathbf{A}}$ and $\bar{\mathbf{B}}$ are banded upper-diagonal circulant matrices of order n. This yields a backward unilateral ARMA model

$$\sum_{k=-n}^{0} \bar{a}_k y(t-k) = \sum_{k=-n}^{0} \bar{b}_k \bar{w}(t-k). \tag{56}$$

These representations should be useful in the smoothing problem for periodic systems [29].

7 Reciprocal Processes and the Covariance Selection Problem

Let $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{X}$ be subspaces in a certain common ambient Hilbert space of zero mean second order random variables. We say that $\mathbf{A}$ and $\mathbf{B}$ are *conditionally orthogonal* given $\mathbf{X}$ if

$$\alpha - \hat{\mathbb{E}}\{\alpha \mid \mathbf{X}\} \perp \beta - \hat{\mathbb{E}}\{\beta \mid \mathbf{X}\}, \quad \forall \alpha \in \mathbf{A}, \forall \beta \in \mathbf{B} \tag{57}$$

(see, e.g., [31]), which we denote $\mathbf{A} \perp \mathbf{B} \mid \mathbf{X}$, and which clearly is equivalent to

$$\mathbb{E}\left\{\hat{\mathbb{E}}\{\alpha \mid \mathbf{X}\}\overline{\hat{\mathbb{E}}\{\beta \mid \mathbf{X}\}}\right\} = \mathbb{E}\{\alpha\overline{\beta}\}, \quad \forall \alpha \in \mathbf{A}, \forall \beta \in \mathbf{B}. \tag{58}$$

Conditional orthogonality is the same as conditional uncorrelatedness, and hence conditional independence in the Gaussian case.

Let $\mathbf{y}_{[t-n,t)}$ and $\mathbf{y}_{(t,t+n]}$ be the n-dimensional random column vectors obtained by stacking $y(t-n), y(t-n+1)\dots, y(t-1)$ and $y(t+1), y(t+2)\dots, y(t+n)$, respectively, in that order. In the same way, $\mathbf{y}_{[t-n,t]}$ is obtained by appending $y(t)$ to $\mathbf{y}_{[t-n,t)}$ as the last element, etc. Here and in the following the sums $t-k$ and $t+k$ are to be understood modulo $2N$. For any interval $(t_1, t_2) \subset [-N+1, N]$, we denote by $(t_1, t_2)^c$ the complementary set in $[1, 2N]$.

Definition 9. A *reciprocal process of order n* on $(-N, N]$ is a process $\{y(t);\ t = -N+1, \dots, N\}$ such that

$$\hat{\mathbb{E}}\{y(t) \mid y(s),\ s \neq t\} = \hat{\mathbb{E}}\{y(t) \mid \mathbf{y}_{[t-n,t)} \vee \mathbf{y}_{(t,t+n]}\} \tag{59}$$

for $t \in (-N, N]$.

This is a generalization introduced in [12] of the concept of *reciprocal process* [23], which can be trivially extended to vector processes. In fact, a reciprocal process in the original sense is here a reciprocal process of order one. This concept does not require stationarity, although here it will always be assumed.

It follows from [31, Proposition 2.4.2 (iii)] that $\{y(t)\}$ is reciprocal of order n if and only if

$$\hat{\mathbb{E}}\{y(t) \mid y(s),\ s \in [t-n,\ t+n]^c\} = \hat{\mathbb{E}}\{y(t) \mid \mathbf{y}_{[t-n,t)} \vee \mathbf{y}_{(t,t+n]}\} \tag{60}$$

for $t \in [-N+1, N]$. In particular, the estimation error

$$\begin{aligned} d(t) &:= y(t) - \hat{\mathbb{E}}\{y(t) \mid y(s),\ s \neq t\} \\ &= y(t) - \hat{\mathbb{E}}\{y(t) \mid \mathbf{y}_{[t-n,t)} \vee \mathbf{y}_{(t,t+n]}\} \end{aligned} \tag{61}$$

must clearly be orthogonal to all random variables $\{y(s),\ s \neq t\}$; i.e. $\mathbb{E}\{d(t)\overline{y(s)}\} = \sigma^2\delta_{st}$, where σ^2 is the variance of $d(t)$. Then $e(t) := d(t)/\sigma^2$ is the (normalized) conjugate process of y satisfying (37), i.e.,

$$\mathbb{E}\{e(t)\overline{y(s)}\} = \delta_{ts}. \tag{62}$$

Since $e(t+k)$ is a linear combination of the components of the random vector $\mathbf{y}_{[t+k-n,t+k+n]}$, it follows from (62) that both $e(t+k)$ and $e(t-k)$ are orthogonal to $e(t)$ for $k > n$. Hence the process $\{e(t)\}$ has correlation bandwidth n, i.e.,

$$\mathbb{E}\{e(t+k)\,e(t)^*\} = 0 \quad \text{for } n < |k| < 2N-n,\; k \in [-N+1, N], \tag{63}$$

and consequently $(\mathbf{y}, \mathbf{e})$ satisfies (38), where $\mathbf{G}$ is banded of order n, which corresponds to an AR representation (42).

Consequently, the AR solutions of the rational circulant covariance extension problem are precisely the ones corresponding to a reciprocal process $\{y(t)\}$ of order n. Next we demonstrate how this representation is connected to the *covariance selection problem* of Dempster [15] by deriving a generalization of this seminal result.

Let $J := \{j_1, \dots, j_p\}$ and $K := \{k_1, \dots, k_q\}$ be two subsets of $\{-N+1, -N+2, \dots, N\}$, and define $\mathbf{y}_J$ and $\mathbf{y}_K$ as the subvectors of $\mathbf{y} = (y_{-N+1}, y_{-N+2}, \cdots, y_N)^{\mathsf{T}}$ with indices in J and K, respectively. Moreover, let

$$\check{\mathbf{Y}}_{J,K} := \text{span}\{y(t);\; t \notin J,\; t \notin K\} = \check{\mathbf{Y}}_J \cap \check{\mathbf{Y}}_K,$$

where $\check{\mathbf{Y}}_J := \text{span}\{y(t);\; t \notin J\}$. With a slight misuse of notation, we shall write

$$\mathbf{y}_J \perp \mathbf{y}_K \mid \check{\mathbf{Y}}_{J,K}, \tag{64}$$

to mean that the subspaces spanned by the components of $\mathbf{y}_J$ and $\mathbf{y}_K$, respectively, are conditionally orthogonal given $\check{\mathbf{Y}}_{J,K}$. This condition can be characterized in terms of the inverse of the covariance matrix $\boldsymbol{\Sigma} := \mathbb{E}\{\mathbf{y}\mathbf{y}^*\} = \left[\sigma_{ij}\right]_{i,j=-N+1}^{N}$ of y.

Theorem 10. *Let* $\mathbf{G} := \boldsymbol{\Sigma}^{-1} = \left[g_{ij}\right]_{i,j=1}^{N}$ *be the concentration matrix of the random vector* y*. Then the conditional orthogonality relation* (64) *holds if and only if* $g_{jk} = 0$ *for all* $(j,k) \in J \times K$*.*

Proof. Let E_J be the $2N \times 2N$ diagonal matrix with ones in the positions $(j_1, j_1), \dots, (j_m, j_m)$ and zeros elsewhere and let E_K be defined similarly in terms of index set K. Then $\check{\mathbf{Y}}_J$ is spanned by the components of $\mathbf{y} - E_J\mathbf{y}$ and $\check{\mathbf{Y}}_K$ by the components of $\mathbf{y} - E_K\mathbf{y}$. Let

$$\tilde{\mathbf{y}}_K := \mathbf{y}_K - \hat{\mathbb{E}}\{\mathbf{y}_K \mid \check{\mathbf{Y}}_K\},$$

and note that its $q \times q$ covariance matrix

$$\tilde{\boldsymbol{\Sigma}}_K := \mathbb{E}\{\tilde{\mathbf{y}}_K\tilde{\mathbf{y}}_K^*\}$$

must be positive definite, for otherwise some linear combination of the components of $\mathbf{y}_K$ would belong to $\check{\mathbf{Y}}_K$. Let $\tilde{\mathbf{y}}_K = G_K\mathbf{y}$ for some $q \times 2N$ matrix G_K. Since $\tilde{\mathbf{y}}_K \perp \check{\mathbf{Y}}_K$,

$$\mathbb{E}\{\tilde{\mathbf{y}}_K(\mathbf{y} - E_K\mathbf{y})^*\} = 0$$

and therefore $\mathbb{E}\{\tilde{\mathbf{y}}_K\mathbf{y}^*\} = G_K\boldsymbol{\Sigma}$ must be equal to $\mathbb{E}\{\tilde{\mathbf{y}}_K(E_K\mathbf{y})^*\}$, which, by $\tilde{\mathbf{y}}_K \in \check{\mathbf{Y}}_K^{\perp}$, in turn equals

$$\mathbb{E}\{\tilde{\mathbf{y}}_K(E_K\mathbf{y})^*\} = \mathbb{E}\{\tilde{\mathbf{y}}_K\hat{\mathbb{E}}\{(E_K\mathbf{y})^* \mid \check{\mathbf{Y}}_K^{\perp}\}\}.$$

However, since the nonzero components of $\hat{\mathbb{E}}\{E_K\mathbf{y} \mid \check{\mathbf{Y}}_K^{\perp}\}$ are those of $\tilde{\mathbf{y}}_K$, there is an $2N \times q$ matrix Π_K with the unit vectors e'_{k_i}, $i = 1, \dots, q$, as the rows such that

$$\hat{\mathbb{E}}\{E_K\mathbf{y} \mid \check{\mathbf{Y}}_K^{\perp}\} = \Pi_K\tilde{\mathbf{y}}_K,$$

and hence

$$\mathbb{E}\{\tilde{\mathbf{y}}_K(E_K\mathbf{y})^*\} = \mathbb{E}\{\tilde{\mathbf{y}}_K\tilde{\mathbf{y}}_K^*\}\Pi_K^* = \tilde{\boldsymbol{\Sigma}}_K\Pi_K^*.$$

Consequently, $G_K\boldsymbol{\Sigma} = \tilde{\boldsymbol{\Sigma}}_K\Pi_K^*$, i.e.,

$$G_K = \tilde{\boldsymbol{\Sigma}}_K\Pi_K^*\boldsymbol{\Sigma}^{-1}.$$

In the same way, $\tilde{\mathbf{y}}_J = G_J\mathbf{y}$, where G_J is the $q \times 2N$ matrix

$$G_J = \tilde{\boldsymbol{\Sigma}}_J\Pi_J^*\boldsymbol{\Sigma}^{-1},$$

and therefore

$$\mathbb{E}\{\tilde{\mathbf{y}}_J\tilde{\mathbf{y}}_K^*\} = \tilde{\boldsymbol{\Sigma}}_J\Pi_J^*\boldsymbol{\Sigma}^{-1}\Pi_K\tilde{\boldsymbol{\Sigma}}_K,$$

which is zero if and only if $\Pi_J^*\boldsymbol{\Sigma}^{-1}\Pi_K = 0$, i.e., $g_{jk} = 0$ for all $(j,k) \in J \times K$.

It remains to show that $\mathbb{E}\{\tilde{\mathbf{y}}_J\tilde{\mathbf{y}}_K^*\} = 0$ is equivalent to (64), which in view of (58), can be written

$$\mathbb{E}\left\{\hat{\mathbb{E}}\{\mathbf{y}_J \mid \check{\mathbf{Y}}_{J,K}\}\hat{\mathbb{E}}\{\mathbf{y}_K \mid \check{\mathbf{Y}}_{J,K}\}^*\right\} = \mathbb{E}\{\mathbf{y}_J\mathbf{y}_K^*\}.$$

However,

$$\mathbb{E}\{\tilde{\mathbf{y}}_J\tilde{\mathbf{y}}_K^*\} = \mathbb{E}\{\mathbf{y}_J\mathbf{y}_K^*\} - \mathbb{E}\left\{\hat{\mathbb{E}}\{\mathbf{y}_J \mid \check{\mathbf{Y}}_J\}\hat{\mathbb{E}}\{\mathbf{y}_K \mid \check{\mathbf{Y}}_K\}^*\right\},$$

so the proof will complete if we show that

$$\mathbb{E}\left\{\hat{\mathbb{E}}\{\mathbf{y}_J \mid \check{\mathbf{Y}}_J\}\hat{\mathbb{E}}\{\mathbf{y}_K \mid \check{\mathbf{Y}}_K\}^*\right\} = \mathbb{E}\left\{\hat{\mathbb{E}}\{\mathbf{y}_J \mid \check{\mathbf{Y}}_{J,K}\}\hat{\mathbb{E}}\{\mathbf{y}_K \mid \check{\mathbf{Y}}_{J,K}\}^*\right\} \tag{65}$$

the proof of which follows precisely the lines of Lemma 2.6.9 in [31, p. 56].

Taking J and K to be singletons we recover as a special case Dempster's original result [15].

To connect back to Definition 9 of a reciprocal process of order n, use the equivalent condition (60) so that, with $J = \{t\}$ and $K = [t-n,\, t+n]^c$, $\mathbf{y}_J = y(t)$ and $\mathbf{y}_K$ are conditionally orthogonal given $\check{\mathbf{Y}}_{J,K} = \mathbf{y}_{[t-n,t)} \vee \mathbf{y}_{(t,t+n]}$. Then $J \times K$ is the set $\left\{t \times [t-n,\, t+n]^c\,;\, t \in (-N,\, N]\right\}$, and hence Theorem 10 states precisely that the circulant matrix $\mathbf{G}$ is banded of order n. We stress that in general $\mathbf{G} = \boldsymbol{\Sigma}^{-1}$ is not banded, as the underlying process $\{y(t)\}$ is not reciprocal of degree n, and we then have an ARMA representation as explained in Sect. 5.

8 Determining P with the Help of Logarithmical Moments

We have shown that the solutions of the circulant rational covariance extension problem, as well as the corresponding bilateral ARMA models, are completely parameterized by $P \in \mathfrak{P}_+(N)$, or, equivalently, by their corresponding banded circulant matrices $\mathbf{P}$. This leads to the question of how to determine the $\mathbf{P}$ from given data.

To this end, suppose that we are also given the logarithmic moments

$$\gamma_k = \int_{-\pi}^{\pi} e^{ik\theta} \log \Phi(e^{i\theta}) d\nu, \quad k = 1, 2, \ldots, n. \tag{66}$$

In the setting of the classical trigonometric moment problem such moments are known as *cepstral coefficients*, and in speech processing, for example, they are estimated from observed data for purposes of design.

Following [30] and, in the context of the trigonometric moment problem, [7, 10, 18, 34], we normalize the elements in $\mathfrak{P}_+(N)$ to define $\tilde{\mathfrak{P}}_+(N) := \{P \in \mathfrak{P}_+(N) \mid p_0 = 1\}$ and consider the problem to find a nonnegative integrable Φ maximizing

$$\mathbb{I}(\Phi) = \int_{-\pi}^{\pi} \log \Phi(e^{i\theta}) d\nu = \frac{1}{2N} \sum_{j=-N+1}^{N} \log \Phi(\zeta_j) \tag{67}$$

subject to the moment constraints (6) and (66). It is shown in [30] that if there is a maximal Φ that is positive on the unit circle, it is given by

$$\Phi(\zeta) = \frac{P(\zeta)}{Q(\zeta)}, \tag{68}$$

where (P, Q) is the unique solution of the dual problem to minimize

$$\mathbb{J}(P, Q) = \langle \mathbf{c}, \mathbf{q} \rangle - \langle \boldsymbol{\gamma}, \mathbf{p} \rangle + \int_{-\pi}^{\pi} P(e^{i\theta}) \log \left(\frac{P(e^{i\theta})}{Q(e^{i\theta})} \right) d\nu \tag{69}$$

over all $(P, Q) \in \tilde{\mathfrak{P}}_+(N) \times \mathfrak{P}_+(N)$, where $\boldsymbol{\gamma} = (\gamma_0, \gamma_1, \ldots, \gamma_n)$ and $\mathbf{p} = (p_0, p_1, \ldots, p_n)$ with $\gamma_0 = 0$ and $p_0 = 1$.

The problem is that the dual problem might have a minimizer on the boundary so that there is no stationery point in the interior, and then the constraints will in general not be satisfied [30]. Therefore the problem needs to be regularized in the style of [17]. More precisely, we consider the regularized problem to minimize

$$\mathbb{J}_\lambda(P,Q) = \mathbb{J}(P,Q) - \lambda \int_{-\pi}^{\pi} \log P(e^{i\theta}) d\nu \tag{70}$$

for some suitable $\lambda > 0$ over all $(P,Q) \in \tilde{\mathfrak{P}}_+(N) \times \mathfrak{P}_+(N)$. Setting $\mathbf{J}_\lambda(\mathbf{P},\mathbf{Q}) := 2N\mathbb{J}_\lambda(P,Q)$, (70) can be written

$$\mathbf{J}_\lambda(\mathbf{P},\mathbf{Q}) = \mathrm{tr}\{\mathbf{C}\mathbf{Q}\} - \mathrm{tr}\{\boldsymbol{\Gamma}\mathbf{P}\} + \mathrm{tr}\{\mathbf{P}\log \mathbf{P}\mathbf{Q}^{-1}\} - \lambda\, \mathrm{tr}\{\log \mathbf{P}\}, \tag{71}$$

where $\boldsymbol{\Gamma}$ is the Hermitian circulant matrix with symbol

$$\Gamma(\zeta) = \sum_{k=-n}^{n} \gamma_k \zeta^{-k}, \quad \gamma_{-k} = \bar{\gamma}_k. \tag{72}$$

Therefore, in the circulant matrix form, the regularized dual problem amounts to minimizing (71) over all banded Hermitian circulant matrices $\mathbf{P}$ and $\mathbf{Q}$ of order n subject to $p_0 = 1$. It is shown in [30] that

$$\boldsymbol{\Sigma} = \mathbf{Q}^{-1}\mathbf{P}, \tag{73}$$

or, equivalently in symbol form (68), maximizes

$$\mathbf{I}(\boldsymbol{\Sigma}) = \mathrm{tr}\{\log \boldsymbol{\Sigma}\} = \log\det \boldsymbol{\Sigma}, \tag{74}$$

or, equivalently (67), subject to (6) and (66), the latter constraint modified so that the logarithmic moment γ_k is exchanged for $\gamma_k + \varepsilon_k$, $k = 1, 2, \dots, n$, where

$$\varepsilon_k = \int_{-\pi}^{\pi} e^{ik\theta} \frac{\lambda}{\hat{P}(e^{i\theta})} d\nu = \frac{\lambda}{2N} \mathrm{tr}\{\mathbf{S}^k \hat{\mathbf{P}}^{-1}\}, \tag{75}$$

$\hat{P}$ being the optimal P.

The following example from [30], provided by Chiara Masiero, illustrates the advantages of this procedure. We start from an ARMA model with $n = 8$ poles and three zeros distributed as in Fig. 3, from which we compute $\mathbf{c} = (c_0, c_1, \dots, c_n)$ and $\boldsymbol{\gamma} = (\gamma_1, \dots, \gamma_n)$ for various choices of the order n. First we determine the maximum entropy solution from $\mathbf{c}$ with $n = 12$ and $N = 1024$. The resulting spectral function Φ is depicted in the top plot of Fig. 4 together with the true spectrum. Next we compute Φ by the procedure in this section using $\mathbf{c}$ and $\boldsymbol{\gamma}$ with $n = 8$ and $N = 128$. The result is depicted in the bottom plot of Fig. 4 again together with the true spectrum. This illustrates the advantage of bilateral ARMA modeling as compared to bilateral AR modeling, as a much lower value on N provides a better approximation, although n is smaller.

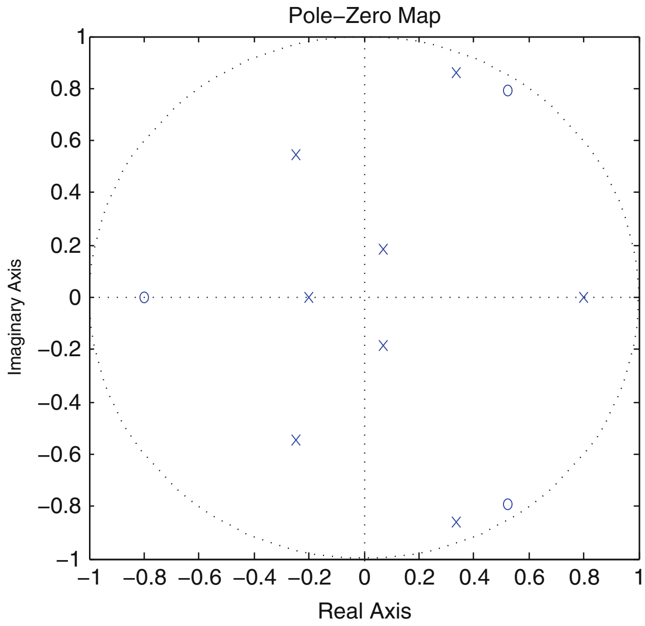


Fig. 3 Poles and zeros of true ARMA model

9 Extensions to the Multivariate Case

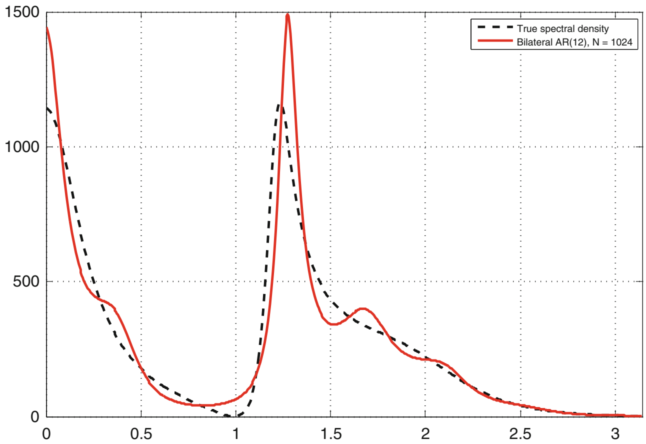
To simplify notation we have so far restricted our attention to scalar stationary periodic processes. We shall now demonstrate that most of the results can be simply extended to the multivariate case, provided we restrict the analysis to scalar pseudo-polynomials $P(\zeta)$. In fact, most of the equations in the previous section will remain intact if we allow ourselves to interpret the scalar quantities as matrix-valued ones.

Let $\{y(t)\}$ be a zero-mean stationary m-dimensional process 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 $2N$. Moreover, 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

$$\Phi(\zeta_j) := \sum_{k=-N+1}^{N} C_k \zeta_j^{-k}, \qquad j = -N+1, \dots, N, \tag{76}$$

which is a positive, Hermitian matrix-valued function of ζ. Then, by the inverse discrete Fourier transformation,

$$C_k = \frac{1}{2N} \sum_{j=-N+1}^{N} \zeta_j^k \Phi(\zeta_j) = \int_{-\pi}^{\pi} e^{ik\theta} \Phi(e^{i\theta}) d\nu, \quad k = -N+1, \dots, N, \tag{77}$$



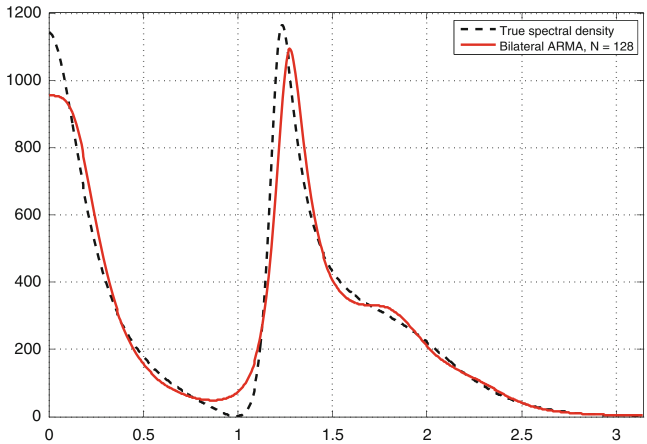


Fig. 4 Bilateral approximations with true spectrum (*dashed*): (*top*) bilateral AR with $n = 12$ and $N = 1024$; (*bottom*) bilateral ARMA with $n = 8$ and $N = 128$ using both covariance and logarithmic moment estimates

where the Stieljes measure $d\nu$ is given by (7). The $m \times m$ matrix function Φ is the *spectral density* of the vector process y. In fact, let

$$\hat{y}(\zeta_k) := \sum_{t=-N+1}^{N} y(t)\zeta_k^{-t}, \quad k = -N+1, \dots, N, \tag{78}$$

be the discrete Fourier transformation of the process y. Since

$$\frac{1}{2N} \sum_{t=-N+1}^{N} (\zeta_k \zeta_\ell^*)^t = \delta_{k\ell}$$

by (8), the random variables (78) are uncorrelated, and

$$\frac{1}{2N}\mathbb{E}\{\hat{y}(\zeta_k)\hat{y}(\zeta_\ell)^*\} = \Phi(\zeta_k)\delta_{k\ell}. \tag{79}$$

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

$$y(t) = \frac{1}{2N}\sum_{k=-N+1}^{N} \zeta_k^t\,\hat{y}(\zeta_k) = \int_{-\pi}^{\pi} e^{ik\theta} d\hat{y}(\theta), \tag{80}$$

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

Next, we define the class $\mathfrak{P}_+^{(m,n)}(N)$ of $m \times m$ Hermitian pseudo-polynomials

$$Q(\zeta) = \sum_{k=-n}^{n} Q_k \zeta^{-k}, \quad Q_{-k} = Q_k^* \tag{81}$$

of degree at most n that are positive definite on the discrete unit circle $\mathbb{T}_{2N}$, and let $\mathfrak{P}_+^{(m,n)} \subset \mathfrak{P}_+^{(m,n)}(N)$ be the subset of all (81) such that $Q(e^{i\theta})$ is positive define for all $\theta \in [-\pi, \pi]$. Moreover let $\mathfrak{C}_+^{(m,n)}(N)$ be the dual cone of all $C = (C_0, C_1, \dots, C_n)$ such that

$$\langle C, Q\rangle := \sum_{k=-n}^{n} \operatorname{tr}\{C_k Q_k^*\} > 0 \quad \text{for all } Q \in \overline{\mathfrak{P}_+^{(m,n)}(N)} \setminus \{0\},$$

and let $\mathfrak{C}_+^{(m,n)} \supset \mathfrak{C}_+^{(m,n)}(N)$ be defined as the dual cone of $\mathfrak{P}_+^{(m,n)}$. Analogously to the scalar case it can be shown that $C \in \mathfrak{C}_+^{(m,n)}$ if and only if the block-Toeplitz matrix

$$\mathbf{T}_n = \begin{bmatrix} C_0 & C_1^* & C_2^* & \cdots & C_n^* \\ C_1 & C_0 & C_1^* & \cdots & C_{n-1}^* \\ C_2 & C_1 & C_0 & \cdots & C_{n-2}^* \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ C_n & C_{n-1} & C_{n-2} & \cdots & C_0 \end{bmatrix} \tag{82}$$

is positive definite [32], a condition that is necessary, but in general not sufficient, for $C \in \mathfrak{C}_+^{(m,n)}(N)$ to hold.

The basic problem is the following. Given the sequence $C = (C_0, C_1, \dots, C_n) \in \mathfrak{C}_+^{(m,n)}(N)$ of $m \times m$ covariance lags, find an extension $C_{n+1}, C_{n+2}, \dots, C_N$ with $C_{-k} = C_k^*$ such that the spectral function Φ defined by (76) has the rational form

$$\Phi(\zeta) = P(\zeta)Q(\zeta)^{-1}, \quad P \in \mathfrak{P}_+^{(1,n)}(N),\ Q \in \mathfrak{P}_+^{(m,n)}(N). \tag{83}$$

Theorem 11. *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*

$$\Phi = PQ^{-1} \tag{84}$$

satisfies the moment conditions

$$\int_{-\pi}^{\pi} e^{ik\theta} \Phi(e^{i\theta}) d\nu = C_k, \quad k = 0, 1, \ldots, n. \tag{85}$$

Theorem 11 is a direct consequence of the following theorem, which also provides an algorithm for computing the solution.

Theorem 12. *For each $(C, P) \in \mathfrak{C}_+^{(m,n)}(N) \times \mathfrak{P}_+^{(1,n)}(N)$, the problem to maximize the functional*

$$\mathbb{I}_P(\Phi) = \int_{-\pi}^{\pi} P(e^{i\theta}) \log\det \Phi(e^{i\theta}) d\nu \tag{86}$$

subject to the moment conditions (85) has a unique solution $\hat{\Phi}$, and it has the form

$$\hat{\Phi}(\zeta) = P(\zeta)\hat{Q}(\zeta)^{-1}, \tag{87}$$

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

$$\mathbb{J}_P(Q) = \langle C, Q\rangle - \int_{-\pi}^{\pi} P(e^{i\theta}) \log\det Q(e^{i\theta}) d\nu \tag{88}$$

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

The proofs of Theorems 11 and 12 follow the lines of [32]. 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.

To formulate a matrix version of Theorems 11 and 12 we need to introduce (Hermitian) block-circulant matrices

$$\mathbf{M} = \sum_{k=-N+1}^{N} S^{-k} \otimes M_k, \quad M_{-k} = M_k^* \tag{89}$$

where $\otimes$ is the Kronecker product and S is the nonsingular $2N \times 2N$ cyclic shift matrix (19). The notation $\mathbf{S}$ will now be reserved for the $2mN \times 2mN$ block-shift matrix

$$\mathbf{S} = S \otimes I_m = \begin{bmatrix} 0 & I_m & 0 & \ldots & 0 \\ 0 & 0 & I_m & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & I_m \\ I_m & 0 & 0 & 0 & 0 \end{bmatrix}. \tag{90}$$

As before $\mathbf{S}^{2N} = \mathbf{S}^0 = \mathbf{I} := I_{2mN}, \mathbf{S}^{k+2N} = \mathbf{S}^k$, and $\mathbf{S}^{2N-k} = \mathbf{S}^{-k} = (\mathbf{S}^k)^{\mathsf{T}}$. Moreover

$$\mathbf{SMS}^* = \mathbf{M} \tag{91}$$

is both necessary and sufficient for $\mathbf{M}$ to be $m \times m$ block-circulant. The symbol of $\mathbf{M}$ is the $m \times m$ pseudo-polynomial

$$M(\zeta) = \sum_{k=-N+1}^{N} M_k \zeta^{-k}, \quad M_{-k} = M_k^*. \tag{92}$$

We shall continue using the notation

$$\mathbf{M} := \mathrm{Circ}\{M_0, M_1, M_2, \dots, M_N, M_{N-1}^*, \dots, M_1^*\} \tag{93}$$

also for (Hermitain) block-circulant matrices.

The problem can now be reformulated in the following way. Given the banded block-circulant matrix

$$\mathbf{C} = \sum_{k=-n}^{n} S^{-k} \otimes C_k, \quad C_{-k} = C_k^* \tag{94}$$

of order n, find an extension $C_{n+1}, C_{n+2}, \dots, C_N$ such that the block-circulant matrix

$$\boldsymbol{\Sigma} = \sum_{k=-N+1}^{N} S^{-k} \otimes C_k, \quad C_{-k} = C_k^* \tag{95}$$

has the symbol (83).

To proceed we need a block-circulant version of Lemma 5.

Lemma 13. *Let* $\mathbf{M}$ *be a block-circulant matrix with symbol* $M(\zeta)$*. Then*

$$\mathbf{M} = \mathbf{F}^* \mathrm{diag}\big(M(\zeta_{-N+1}), M(\zeta_{-N+2}), \dots, M(\zeta_N)\big)\mathbf{F}, \tag{96}$$

where $\mathbf{F}$ *is the unitary* $2mN \times 2mN$ *matrix*

$$\mathbf{F} = \frac{1}{\sqrt{2N}} \begin{bmatrix} \zeta_{-N+1}^{N-1} I_m & \zeta_{-N+1}^{N-2} I_m & \cdots & \zeta_{-N+1}^{-N} I_m \\ \vdots & \vdots & \cdots & \vdots \\ \zeta_0^{N-1} I_m & \zeta_0^{N-2} I_m & \cdots & \zeta_0^{-N} I_m \\ \vdots & \vdots & \cdots & \vdots \\ \zeta_N^{N-1} I_m & \zeta_N^{N-2} I_m & \cdots & \zeta_N^{-N} I_m \end{bmatrix}. \tag{97}$$

Moreover, if $M(\zeta_k)$ is positive definite for all k, then

$$\log \mathbf{M} = \mathbf{F}^* \operatorname{diag}\big(\log M(\zeta_{-N+1}), \log M(\zeta_{-N+2}), \dots, \log M(\zeta_N)\big)\mathbf{F}, \tag{98}$$

where diag stands for block diagonal.

The proof of Lemma 13 will be omitted, as it follows the same lines as that of Lemma 5 with straight-forward modification to the multivariate case. Clearly the inverse

$$\mathbf{M}^{-1} = \mathbf{F}^* \operatorname{diag}\big(M(\zeta_{-N+1})^{-1}, M(\zeta_{-N+2})^{-1}, \dots, M(\zeta_N)^{-1}\big)\mathbf{F} \tag{99}$$

is also block-circulant, and

$$\mathbf{S} = \mathbf{F}^* \operatorname{diag}\big(\zeta_{-N+1} I_m, \zeta_{-N+2} I_m, \dots, \zeta_N I_m\big)\mathbf{F}. \tag{100}$$

However, unlike the scalar case, block-circulant matrices do not commute in general.

Given Lemma 13, we are now in a position to reformulate Theorems 11 and 12 in matrix from.

Theorem 14. *Let $C \in \mathfrak{C}_+^{(m,n)}(N)$, and let $\mathbf{C}$ be the corresponding block-circulant matrix (94) and (82) the corresponding block-Toeplitz matrix. Then, for each positive-definite banded $2mN \times 2mN$ block-circulant matrices*

$$\mathbf{P} = \sum_{k=-n}^{n} S^{-k} \otimes p_k I_m, \quad p_{-k} = \bar{p}_k \tag{101}$$

of order n, where $P(\zeta) = \sum_{k=-n}^{n} p_k \zeta^{-k} \in \mathfrak{P}_+^{(1,n)}(N)$, there is a unique sequence $Q = (Q_0, Q_1, \dots, Q_n)$ of $m \times m$ matrices defining a positive-definite banded $2mN \times 2mN$ block-circulant matrix

$$\mathbf{Q} = \sum_{k=-n}^{n} S^{-k} \otimes Q_k, \quad Q_{-k} = Q_k^* \tag{102}$$

of order n such that

$$\boldsymbol{\Sigma} = \mathbf{Q}^{-1}\mathbf{P} \tag{103}$$

is a block-circulant extension (95) of $\mathbf{C}$. The block-circulant matrix (103) is the unique maximizer of the function

$$\mathscr{I}_{\mathbf{P}}(\boldsymbol{\Sigma}) = \operatorname{tr}(\mathbf{P} \log \boldsymbol{\Sigma}) \tag{104}$$

subject to

$$\mathbf{E}_n{}^T \boldsymbol{\Sigma} \mathbf{E}_n = \mathbf{T}_n, \quad \text{where } \mathbf{E}_n = \begin{bmatrix} \mathbf{I}_{mn} \\ \mathbf{0} \end{bmatrix}. \tag{105}$$

Moreover, $\mathbf{Q}$ *is the unique optimal solution of the problem to minimize*

$$\mathscr{J}_{\mathbf{P}}(\mathbf{Q}) = \text{tr}(\mathbf{CQ}) - \text{tr}(\mathbf{P} \log \mathbf{Q}) \tag{106}$$

over all positive-definite banded $2mN \times 2mN$ *block-circulant matrices* (102) *of order* n*. The functional* $\mathscr{J}_{\mathbf{P}}$ *is strictly convex.*

For $\mathbf{P} = \mathbf{I}$ we obtain the maximum-entropy solution considered in [12], where the primal problem to maximize $\mathscr{I}_{\mathbf{I}}$ subject to (105) was presented. In [12] there was also an extra constraint (91), which, as we can see, is not needed, since it is automatically fulfilled. For this reason the dual problem presented in [12] is more complicated than merely minimizing $\mathscr{J}_{\mathbf{I}}$.

Next suppose we are also given the (scalar) logarithmic moments (66) and that $C \in \mathfrak{C}_+^{(m,n)}(N)$. Then, if the problem to maximize $\text{tr}\{\log \boldsymbol{\Sigma}\}$ subject to (105) and (66) over all positive-definite block-circulant matrices (95) has a solution, then it has the form

$$\boldsymbol{\Sigma} = \mathbf{Q}^{-1}\mathbf{P} \tag{107}$$

where the $(\mathbf{P}, \mathbf{Q})$ is a solution of the dual problem to minimize

$$\mathbf{J}(\mathbf{P}, \mathbf{Q}) = \text{tr}\{\mathbf{CQ}\} - \text{tr}\{\boldsymbol{\Gamma}\mathbf{P}\} + \text{tr}\{\mathbf{P} \log \mathbf{PQ}^{-1}\}, \tag{108}$$

over all positive-definite block-circulant matrices of the type (101) and (102) with the extra constrain $p_0 = 1$, where $\boldsymbol{\Gamma}$ is the block-circulant matrix formed in the style of (102) from

$$\Gamma(\zeta) = \sum_{k=-n}^{n} \gamma_k \zeta^{-k}, \quad \gamma_{-k} = \bar{\gamma}_k. \tag{109}$$

However, the minimum of (108) may end up on the boundary, in which case the constraint (66) may fail to be satisfied. Therefore, as in the scalar case, we need to regularize the problem by instead minimizing

$$\mathbf{J}_\lambda(\mathbf{P}, \mathbf{Q}) = \text{tr}\{\mathbf{CQ}\} - \text{tr}\{\boldsymbol{\Gamma}\mathbf{P}\} + \text{tr}\{\mathbf{P} \log \mathbf{PQ}^{-1}\} - \lambda\, \text{tr}\{\log \mathbf{P}\}. \tag{110}$$

This problem has a unique optimal solution (107) satisfying (105), but not (66). The appropriate logarithmic moment constraint is obtained as in the scalar case by exchanging γ_k for $\gamma_k + \varepsilon_k$ for each $k = 1, 2, \ldots, n$, where ε_k is given by (75).

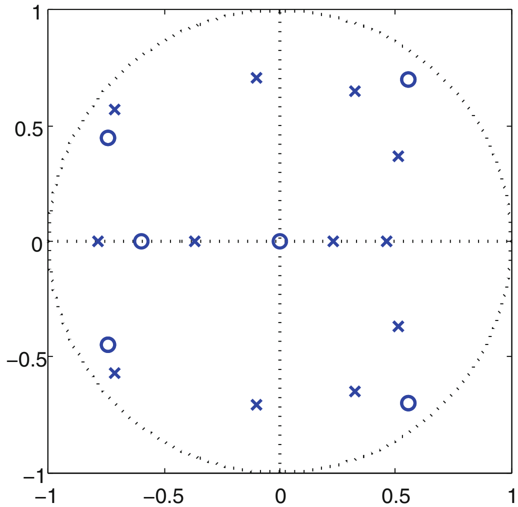


Fig. 5 Poles and zeros of an ARMA 2×2 model of order $n = 6$

Again each solution leads to an ARMA model

$$\sum_{k=-n}^{n} Q_k y(t-k) = \sum_{k=-n}^{n} p_k e(t-k), \tag{111}$$

where $\{e(t)\}$ is the conjugate process of $\{y(t)\}$, $Q_0, Q_1, \dots, Q_n$ are $m \times m$ matrices, whereas $p_0, p_1, \dots, p_n$ are scalar with $p_0 = 1$.

We illustrate this theory with a simple example from [32], where a covariance sequence $C := (C_0, C_1, \dots C_n)$ and a cepstral sequence $\boldsymbol{\gamma} := (\gamma_1, \gamma_2, \dots, \gamma_n)$ have been computed 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 and poles are illustrated in Fig. 5.

Given C and $\boldsymbol{\gamma}$, we apply the procedure in this section to determine a pair $(\mathbf{P}, \mathbf{Q})$ of order $n = 6$. For comparison we also compute an bilateral AR approximation with $n = 12$ fixing $\mathbf{P} = \mathbf{I}$. As illustrated in Fig. 6, the bilateral ARMA model of order $n = 6$ computed with $N = 32$ outperforms the bilateral AR model of order $n = 12$ with $N = 64$.

The results of Sect. 5 can also be generalized to the multivariate case along the lines described in [11].

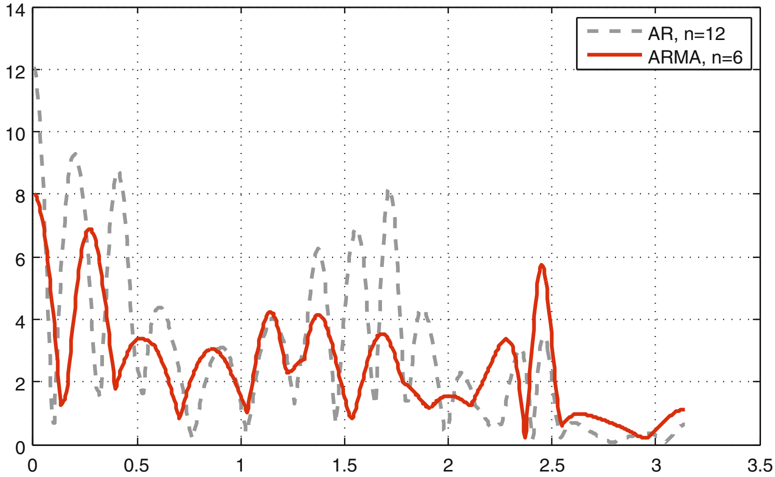
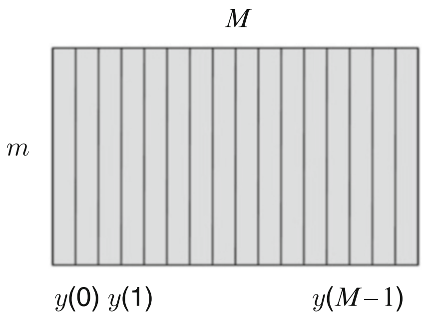


Fig. 6 The norm of the approximation error for a bilateral AR of order 12 for $N = 64$ and a bilateral ARMA of order 6 for $N = 32$

Fig. 7 An image modeled as a reciprocal vector process



10 Application to Image Processing

In [12] the circulant maximum-entropy solution has been used to model spatially stationary images (*textures*) [40] in terms of (vector-valued) stationary periodic processes. The image could be thought of as an $m \times M$ matrix of pixels where the columns form a m-dimensional reciprocal process $\{y(t)\}$, which can extended to a periodic process with period $M > N$ outside the interval $[0, N]$; see Fig. 7.

This imposes the constraint $C_{M-k} = C_k{}^{\mathsf{T}}$ on the covariance lags $C_k := E\{y(t+k)y(t)^{\mathsf{T}}\}$, leading to a circulant Toeplitz matrix. The problem considered in [12] is to model the process $\{y(t)\}$ given (estimated) $C_0, C_1, \ldots, C_n$, where $n < N$ with an efficient low-dimensional model. This is precisely a problem of the type considered in Sect. 9.

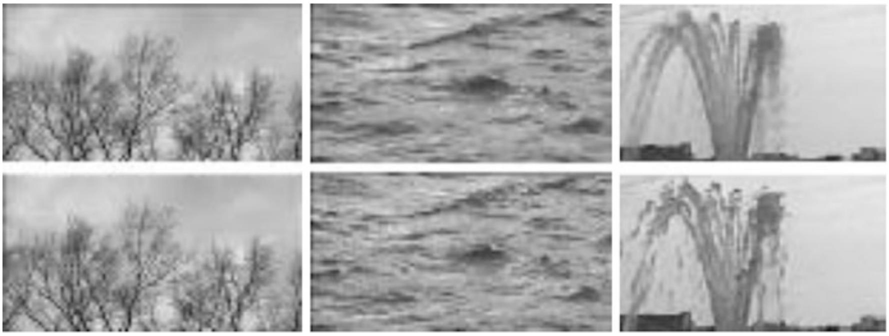


Fig. 8 Three images modeled by reciprocal processes (original at *bottom*)

Solving the corresponding circulant maximum-entropy problem (with $\mathbf{P} = \mathbf{I}$), $n = 1$, $m = 125$ and $N = 88$, Carli et al. [12] derived a bilateral model of the images at the bottom row of Fig. 8 to compress the images in the top row, thereby achieving a compression of 5:1.

While the compression ratio falls short of competing with current jpeg standards (typically 10:1 for such quality), our approach suggests a new stochastic alternative to image encoding. Indeed the results in Fig. 8 apply just the maximum entropy solution of order $n = 1$. Simulations such as those in Fig. 4 suggest that much better compression can be made using bilateral ARMA modeling.

An alternative approach to image compression using multidimensional covariance extension can be found in the recent paper [39].

References

1. Blomqvist, A., Lindquist, A., Nagamune, R.: Matrix-valued Nevanlinna-Pick interpolation with complexity constraint: an optimization approach. IEEE Trans. Autom. Control **48**, 2172–2190 (2003)
2. Byrnes, C.I., Lindquist, A.: On the partial stochastic realization problem. IEEE Trans. Autom. Control **42**, 1049–1069 (1997)
3. Byrnes, C.I., Lindquist, A.: The generalized moment problem with complexity constraint. Integr. Equ. Oper. Theory **56**, 163–180 (2006)
4. Byrnes, C.I., Lindquist, A.: 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, vol. 190, pp. 157–169. Birkhäuser, Basel (2009)
5. Byrnes, C.I., Lindquist, A, Gusev, S.V., Matveev, A.V.: A complete parameterization of all positive rational extensions of a covariance sequence. IEEE Trans. Autom. Control **40**, 1841–1857 (1995)
6. Byrnes, C.I., Gusev, S.V., Lindquist, A.: A convex optimization approach to the rational covariance extension problem. SIAM J. Contr. Opt. **37**, 211–229 (1999)
7. Byrnes, C.I., Enqvist, P., Lindquist, A.: Cepstral coefficients, covariance lags and pole-zero models for finite datastrings. IEEE Trans. Signal Process **50**, 677–693 (2001)

8. Byrnes, C.I., Georgiou, T.T., Lindquist, A.: A generalized entropy criterion for Nevanlinna-Pick interpolation with degree constraint. IEEE Trans. Autom. Control **45**, 822–839 (2001)
9. Byrnes, C.I., Gusev, S.V., Lindquist, A.: From finite covariance windows to modeling filters: a convex optimization approach. SIAM Rev. **43**, 645–675 (2001)
10. Byrnes, C.I., Enqvist, P., Lindquist, A.: Identifiability and well-posedness of shaping-filter parameterizations: a global analysis approach. SIAM J. Control Optim. **41**, 23–59 (2002)
11. Picci, G.: A new approach to circulant band extension: Proc. 22nd International Symposium on the Mathematical Theory of Networks and Systems (MTNS), July 11–15, Minneapolis, MN, USA, pp 123–130 (2016)
12. Carli, F.P., Ferrante, A., Pavon, M., Picci, G.: A maximum entropy solution of the covariance extension problem for reciprocal processes. IEEE Trans. Autom. Control **56**, 1999–2012 (2011)
13. Chiuso, A., Ferrante, A., Picci, G.: Reciprocal realization and modeling of textured images. In: Proceedings of the 44rd IEEE Conference on Decision and Control (2005)
14. Davis, P.: Circulant Matrices. Wiley, New York (1979)
15. Dempster, A.P.: Covariance selection. Biometrics **28**(1), 157–175 (1972)
16. Dunford, N., Schwartz, J.T.: Linear Operators, Part I: General Theory. Wiley, New York (1958)
17. Enqvist, P.: Spectral estimation by Geometric, Topological and Optimization Methods, PhD thesis, Optimization and Systems Theory, KTH, Stockholm, (2001)
18. Enqvist, P.: A convex optimization approach to ARMA(n,m) model design from covariance and cepstrum data. SIAM J. Control Optim. **43**(3), 1011–1036 (2004)
19. Georgiou, T.T.: Partial Realization of Covariance Sequences, Ph.D. thesis, CMST, University of Florida, Gainesville (1983)
20. Georgiou, T.T.: Realization of power spectra from partial covariances, IEEE Trans. Acoust. Speech Signal Process. **35**, 438–449 (1987)
21. Georgiou, T.T.: Solution of the general moment problem via a one-parameter imbedding. IEEE Trans. Autom. Control **50**, 811–826 (2005)
22. Georgiou, T.T., Lindquist, A.: Kullback-Leibler approximation of spectral density functions. IEEE Trans. Inform. Theory **49**, 2910–2917 (2003)
23. Jamison, B.: Reciprocal processes. Zeitschrift. Wahrsch. Verw. Gebiete **30**, 65–86 (1974)
24. Kalman, R.E.: Realization of covariance sequences. In: Proceedings of the Toeplitz Memorial Conference, Tel Aviv (1981)
25. Krein, M.G., Nudelman,A.A.: The Markov Moment Problem and Extremal Problems. American Mathematical Society, Providence, RI (1977)
26. Krener, A.J.: Reciprocal processes and the stochastic realization problem for acausal systems, In: Byrnes, C.I., Lindquist, A. (eds.) Modeling Identification and Robust Control, pp. 197–211. North-Holland, Amsterdam (1986)
27. Krener, A.J., Frezza, R., Levy, B.C.: Gaussian reciprocal processes and self-adjoint differential equations of second order. Stoch. Stoch. Rep. **34**, 29–56 (1991)
28. Levy, B.C., Ferrante, A.: Characterization of stationary discrete-time Gaussian reciprocal processes over a finite interval, SIAM J. Matrix Anal. Appl. **24**, 334–355 (2002)
29. Levy, B.C., Frezza, R., Krener, A.J.: Modeling and estimation of discrete-time Gaussian reciprocal processes. IEEE Trans. Autom. Contr. **35**, 1013–1023 (1990)
30. Lindquist, A., Picci, G.: The circulant rational covariance extension problem: the complete solution. IEEE Trans. Autom. Control **58**, 2848–2861 (2013)
31. Lindquist, A. Picci, G.: Linear Stochastic Systems: A Geometric Approach to Modeling, Estimation and Identification. Springer, Heidelberg, New York, Dordrecht, London (2015)
32. Lindquist, A. Masiero, C., Picci, G.: On the multivariate circulant rational covariance extension problem. In: Proceeding of the 52st IEEE Conference on Decision and Control (2013)
33. Masani, P.: The prediction theory of multivariate stochastic processes, III. Acta Math. **104**, 141–162 (1960)
34. Musicus, B.R., Kabel, A.M.: Maximum entropy pole-zero estimation, Technical Report 510, MIT Research Lab. Electronics, Aug. 1985; now available on the internet at http://dspace.mit.edu/bitstream/handle/1721.1/4233/RLE-TR-510-17684936.pdf

35. Pavon, M., Ferrante, A.: On the geometry of maximum entropy problems. SIAM Rev. **55**(3), 415–439 (2013)
36. Picci, G., Carli, F.: Modelling and simulation of images by reciprocal processes. In: Proceedings of the Tenth International Conference on Computer Modeling and Simulation UKSIM, pp. 513–518 (2008)
37. Ringh, A., Karlsson, J.: A fast solver for the circulant rational covariance extension problem. In: European Control Conference (ECC), July 2015, pp. 727–733 (2015)
38. Ringh, A., Lindquist, A.: Spectral estimation of periodic and skew periodic random signals and approximation of spectral densities. In: 33rd Chinese Control Conference (CCC), pp. 5322–5327 (2014)
39. Ringh A., Karlsson, J., Lindquist, A: Multidimensional rational covariance extension with applications to spectral estimation and image compression. SIAM J. Control Optim. **54**(4), 1950–1982 (2016)
40. Soatto, S., Doretto, G., Wu, Y.: Dynamic textures. In: Proceedings of the International Conference Computer Vision, July 2001, pp. 439–446