Approximation of Vector Processes by Covariance Matching With Applications to Smoothing

Giorgio Picci and Bin Zhu

***Abstract*—We propose a new algorithm for the partial stochastic realization of vector-valued periodic processes from finite covariance data, based on a nonlinear generalization of the classical Yule-Walker equations. We discuss an application to finite-interval smoothing of a linear time-invariant system.**

***Index Terms*—Stochastic systems, identification, modeling.**

I. Introduction

THE MOTIVATION of this letter originates from the *rational covariance extension problem* also called *partial stochastic realization* which has a long history, see [1], [3], [11]–[13] and has wide applications in the fields of signal processing, system identification and control; see [2], [5], [6], [10], [12] and references therein. For modeling stochastic processes defined on a finite-interval, such covariance extension leads to periodic ARMA models and to a *circulant* matrix completion problem. As a generalization of earlier work on reciprocal processes [7], [8], a complete parameterization of finite-interval rational solutions is given in [17] and extended to a special multivariate case in [16] and [18]. In these papers the circulant completion problem is reformulated as optimization of a *generalized entropy* functional, the main tools being convex optimization and calculus of variations.

In this and in the companion paper [25], we consider a variant of the circulant rational covariance extension problem stated in terms of periodic ARMA models defined on a finite interval. We have found that the problem can be reformulated in terms of a generalization of the Yule-Walker covariance equations without involving the solution of a variational problem as in [16] and [17]. Although the resulting equations turn out to be nonlinear, a natural iterative solution is apparent from their structure which has been implemented and shown to work well in several examples.

Manuscript received February 27, 2017; revised May 7, 2017; accepted May 22, 2017. Date of publication June 1, 2017; date of current version June 21, 2017. This work was supported by the China Scholarship Council under Grant 201506230140. Recommended by Senior Editor G. Yin. *(Corresponding author: Giorgio Picci.)*

The authors are with the Department of Information Engineering, University of Padova, 35131 Padua, Italy (e-mail: picci@dei.unipd.it; zhubin@dei.unipd.it).

Digital Object Identifier 10.1109/LCSYS.2017.2710944

This original contribution leads to a new algorithm valid for vector processes which avoids the variational formulation and seems to provide a viable alternative to the earlier optimization framework of [16] and [17]. In the companion paper [25], a proof of the local convergence of the algorithm is presented for the scalar problem. Generalization of this proof to the vector case is under way.

As we shall see later, periodic ARMA models can provide a useful finite-interval approximation of a stationary state space model, e.g., a Gauss-Markov model defined on the whole time line $\mathbb{Z}$, obtained by matching a certain finite number of covariances of the process. The finite-interval model can then be used to derive an easy-to-implement constant-coefficients algorithms for linear smoothing of data of finite duration.

The outline of this letter is as follows: Some preliminaries on periodic ARMA models and their spectral representation are given in Section II. In Section III, the covariance extension problem for vector-valued periodic processes is formulated. A set of nonlinear Yule-Walker equations is introduced and an iterative algorithm to compute the solution is described. Application of the theory to the finite-interval smoothing problem is presented in Section IV exploiting model approximation by covariance matching. Finally, a simplified numerical example is given for illustration.

II. Bilateral and Unilateral ARMA Models

The dynamics of a periodic process can be defined in terms of relations among its random variables in just one particular period, namely $[-N+1, N]$ which will be identified with the finite modular group $\mathbb{Z}_{2N}$. In this setting, the most general finitely parametrized analog of a stationary finite-dimensional linear model for an m-dimensional periodic process turns out to be a *bilateral ARMA model* of finite order n [7], [17],

$$\sum_{k=-n}^{n} Q_k\, y(t-k) = \sum_{k=0}^{n} P_k\, e(t-k), \quad t \in \mathbb{Z}_{2N}, \tag{1}$$

with $\{Q_k,\ P_k \in \mathbb{R}^{m\times m}\}$ matrix parameters satisfying the symmetry $P_{-k} = P_k^\top$ and $Q_{-k} = Q_k^\top$ and $\{e(t)\}$ an m-dimensional periodic process, called the *conjugate process* of $\{y(t)\}$ (also called the *double-sided innovation* [20]). The conjugate process is *delta-correlated with* y in the sense that $\mathbb{E}y(t)e(s)^\top = \delta_{t,s}I_m$ where $\delta_{t,s} = 1$ for $t = s$ and zero otherwise. By periodicity, the model is associated to periodic

boundary conditions at the end points

$$y(-N) = y(N), \ldots, \; y(-N-n+1) = y(N-n+1). \quad (2)$$

which induce a block-circulant structure on the model (1). To explicit this structure we shall use boldface notations for the vectors obtained by stacking the components of a periodic process

$$\mathbf{y} = \begin{bmatrix} y(-N+1) \\ \vdots \\ y(N) \end{bmatrix}, \quad \mathbf{e} = \begin{bmatrix} e(-N+1) \\ \vdots \\ e(N) \end{bmatrix} \in \mathbb{R}^{2mN}. \quad (3)$$

The model (1) can then be rewritten as an equivalent circulant matrix equation

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

where $\mathbf{Q}$ and $\mathbf{P}$ are symmetric positive semidefinite *block-circulants* with elements the coefficients $\{Q_k\}$ and $\{P_k\}$ defined above. Therefore $\mathbf{Q}$ and $\mathbf{P}$ are symmetric *banded* block-circulants of (block-)bandwidth n. Because of the orthogonality of $\mathbf{y}$ to its conjugate process, assuming invertibility of $\mathbf{Q}$, the covariance $\mathbf{\Sigma} \coloneqq \mathbb{E}\mathbf{y}\mathbf{y}^\top$ has the expression

$$\Sigma = \mathbf{Q}^{-1}\mathbf{P} \in \mathbb{R}^{2Nm \times 2Nm}, \quad (4)$$

which is also a block-circulant positive semidefidefinite symmetric matrix. Note that since $\mathbf{\Sigma}$ is symmetric the matrices $\mathbf{Q}$ and $\mathbf{P}$ must commute. The special structure of the resulting covariance matrix expressed as a ratio of two block-circulant banded matrices, will later be exploited to solve the finite-interval smoothing problem in a stationary setting.

The expression (4) has an analog in terms of spectral representation which we shall describe right after recalling some facts about the Discrete Fourier Transform (DFT). For a more extensive review of the DFT see [8], [17].

Let $\zeta_1 \coloneqq e^{i\Delta}$ where $\Delta = \pi/N$, be the primitive $2N$-th root of unity and define the discrete variable ζ taking the $2N$ values $\zeta_k \coloneqq \zeta_1^k = e^{i\Delta k}$; $k = -N+1, \ldots, 0, \ldots, N$ running counterclockwise on the unit circle $\mathbb{T}$. The set of $2N$ points $\{\zeta_k\}$ will be called the *discrete unit circle*, denoted by $\mathbb{T}_{2N}$. In particular, we have $\zeta_{-k} = \overline{\zeta_k}$ (complex conjugate). The DFT $\mathcal{F}$ maps a, possibly random, finite support signal $g = \{g(t);\, t = -N+1, \ldots, N\}$, into a complex sequence:

$$\hat{g}(\zeta_j) \coloneqq \sum_{t=-N+1}^{N} g(t)\zeta_j^{-t}, \quad j = -N+1, -N+2, \ldots, N. \quad (5)$$

and the signal g can be recovered from its DFT $\hat{g}$ by the formula

$$g(t) = \sum_{j=-N+1}^{N} \zeta_j^t \hat{g}(\zeta_j)\frac{\Delta}{2\pi}, \quad t = -N+1, -N+2, \ldots, N, \quad (6)$$

where $\frac{\Delta}{2\pi} = \frac{1}{2N}$ plays the role of a uniform discrete measure dv with total mass one supported on $\mathbb{T}_{2N}$. The map $\mathcal{F}$ is in fact unitary (the *Parseval Identity* hods also for the DFT). For a random process y the inverse formula (6) yields a spectral representation of $y(t)$

$$y(t) = \sum_{k=-N+1}^{N} \zeta_k^t \hat{y}(\zeta_k)\frac{1}{2N} = \int_{-\pi}^{\pi} e^{ik\theta} d\hat{y}(\theta), \quad (7)$$

similar to the usual one valid for stationary processes on $\mathbb{Z}$ [19]. Here the random variables $\{\hat{y}(\zeta_k)\}$ are uncorrelated and define an orthogonal random measure $d\hat{y}(\theta) \coloneqq \hat{y}(e^{i\theta})dv(\theta)$. supported on $\mathbb{T}_{2N}$. Therefore the *spectral density* of the process $\{y(t)\}$, defined by

$$\Phi(\zeta) = \frac{1}{2N}\,\mathbb{E}\big[\,\hat{y}(\zeta)\hat{y}(\zeta)^*\big] \quad (8)$$

is an Hermitian positive semidefinite matrix function on $\mathbb{T}_{2N}$. One can in general define the *symbol* of a block-circulant just as the DFT of its first (block) column; see [17] for details. Then $\Phi(\zeta)$ is just the *symbol* of the covariance matrix $\mathbf{\Sigma}$.

An important fact is that the DFT is an *algebra homomorphism* mapping the circulant matrices of the same dimension to their symbols [17, p. 2851]. It follows that the representation (4) is equivalent to

$$\Phi(\zeta) = Q(\zeta)^{-1}P(\zeta) \quad (9)$$

where $Q(\zeta)$ and $P(\zeta)$ are the symbols of the circulants $\mathbf{Q}$ and $\mathbf{P}$; they are Hermitian matrix pseudo-polynomials (with both positive and negative powers of the indeterminate) of order n and positive semidefinite on $\mathbb{T}_{2N}$. This is why we say that $\Phi(\zeta)$ is a *rational* matrix function. We shall require that

$$\det Q(\zeta) \neq 0, \; \forall\, \zeta \in \mathbb{T}_{2N}, \quad (10)$$

i.e., the matrix $Q(\zeta)$ is invertible on the discrete unit circle. This is in turn equivalent to nonsingularity of $\mathbf{Q}$ [8].

Periodic processes can also be conveniently described by *unilateral ARMA models*, which are useful for the recursive implementation of algorithms. Essentially they can be obtained by a spectral factorization of the spectral density (9). Since $\Phi(\zeta)$ is a positive definite rational function of ζ it must admit square spectral factors, that is

$$\Phi(\zeta) = W(\zeta)W(\zeta^{-1})^\top, \qquad \zeta \in \mathbb{T}_{2N} \quad (11)$$

where $W(\zeta)$ is also a rational $m \times m$ matrix function having a matrix fraction representation $W(\zeta) = A(\zeta)^{-1}B(\zeta)$ where

$$A(\zeta) \coloneqq \sum_{k=0}^{n} A_k\,\zeta^{-k}, \quad B(\zeta) \coloneqq \sum_{k=0}^{n} B_k\,\zeta^{-k}. \quad (12)$$

We have used negative powers of ζ corresponding to delay operators in the Fourier domain. An m-dimensional stationary periodic process $\{y(t)\}$ can then be described by a *forward unilateral ARMA representation*

$$\sum_{k=0}^{n} A_k y(t-k) = \sum_{k=0}^{n} B_k\, w(t-k), \quad t \in \mathbb{Z}_{2N}, \quad (13)$$

where $\{w(t)\}$ is an m-dimensional periodic normalized white noise, i.e., $\mathbb{E}[w(t)w(t)^\top] = I_m$. With this convention the model is also called *normalized*. Naturally it must also be associated to the periodic boundary conditions (2) at the end points.

The details of circulant spectral factorization cannot be fully reviewed in this letter and will be referred to [17] and [21].

In particular, the factorization (11) has an isomorphic counterpart in terms of circulants

$$\mathbf{\Sigma} = \mathbf{W}\mathbf{W}^\top \tag{14}$$

where $\mathbf{W}$ is block circulant having symbol $W(\zeta)$

$$\mathbf{W} = \mathrm{Circ}\{W_0, W_1, \dots, W_N, W_{-N+1}, \dots, W_{-1}\}. \tag{15}$$

With the vector notation (3), model (13) can be written compactly as a linear equation

$$\mathbf{A}\mathbf{y} = \mathbf{B}\mathbf{w}, \tag{16}$$

in which $\mathbf{A}$ and $\mathbf{B}$ are $2mN \times 2mN$ having symbols (12). They are represented as *lower-triangular block-circulant* matrices of bandwidth n, that is

$$\begin{aligned}\mathbf{A} &= \begin{bmatrix} A_0 & 0 & \cdots & 0 & A_n & \cdots & A_1 \\ A_1 & A_0 & 0 & \cdots & 0 & \cdots & A_2 \\ \vdots & \vdots & \ddots & & \vdots & & \vdots \\ A_n & A_{n-1} & \cdots & A_0 & 0 & \cdots & 0 \\ 0 & A_n & A_{n-1} & & & & \\ \vdots & \vdots & \ddots & \ddots & & \ddots & \vdots \\ 0 & 0 & \cdots & A_n & \cdots & A_1 & A_0 \end{bmatrix} \\ &:= \mathrm{Circ}\{A_0, A_1, \dots, A_n, 0, \dots, 0\}\end{aligned} \tag{17}$$

and similarly

$$\mathbf{B} = \mathrm{Circ}\{B_0, B_1, \dots, B_n, 0, \dots, 0\}. \tag{18}$$

We require condition (10) also holds for $A(\zeta)$ so that $\mathbf{A}$ is invertible. The solution of (16) can formally be written in the Fourier and time domain as

$$\hat{y}(\zeta) = A(\zeta)^{-1}B(\zeta)\hat{w}(\zeta) \quad \Leftrightarrow \quad \mathbf{y} = \mathbf{A}^{-1}\mathbf{B}\mathbf{w} \tag{19}$$

so that the block-circulant $\mathbf{W}$ has as symbol the discrete transfer function $W(\zeta) := A(\zeta)^{-1}B(\zeta)$. The inverse Fourier transform

$$W_t := \sum_{k=-N+1}^{N} \zeta_k^t A(\zeta_k)^{-1}B(\zeta_k)\frac{1}{2N}, \quad t \in \mathbb{Z}_{2N} \tag{20}$$

called the *impulse response* of the system, yields a convolution representation of the process

$$y(t) = \sum_{s=-N+1}^{N} W_{t-s}\, w(s), \quad t \in \mathbb{Z}_{2N} \tag{21}$$

which corresponds to the matrix-vector product

$$\mathbf{y} = \mathbf{W}\mathbf{w}. \tag{22}$$

From unilateral to bilateral ARMA models. Later on we shall need to express the covariance matrix $\mathbf{\Sigma} = \mathbf{W}\mathbf{W}^\top$ as a bilateral matrix fraction of the type (4). This operation is easy when $B(z)$ is a scalar polynomial times the identity matrix, i.e., $B(z) = b(z)I_m$, in which case $W(\zeta)$ admits a factorization of the type $A(\zeta)^{-1}b(\zeta)I_m$ so that

$$\Phi(\zeta) = \Big[A(\zeta^{-1})^\top A(\zeta)\Big]^{-1} b(\zeta)b(\zeta^{-1})$$

and one may take $Q(\zeta) = A(\zeta^{-1})^\top A(\zeta)$ and $P(\zeta) = b(\zeta)b(\zeta^{-1})I_m$. In general, letting $U(\zeta)$ be the $m \times m$ unitary matrix of eigenvectors of $\Phi(\zeta)$, a sufficient condition is that its scalar positive Hermitian eigenvalues $\hat{\Phi}_k(\zeta)$ should each be a ratio of positive scalar symmetric polynomials of degree smaller than N, i.e.,

$$\hat{\Phi}_k = \frac{\hat{P}_k(\zeta)}{\hat{Q}_k(\zeta)}, \qquad k = 1, 2, \dots, m.$$

Then with $Q(\zeta) := U(\zeta)^{-1}\,\mathrm{diag}\{\hat{Q}_1(\zeta), \dots, \hat{Q}_m(\zeta)\}U(\zeta)^{-*}$ and $P(\zeta) := U(\zeta)^*\,\mathrm{diag}\{\hat{P}_1(\zeta), \dots, \hat{P}_m(\zeta)\}U(\zeta)$ one obtains precisely the factorization (9).

III. Covariance Matching

The solution of the following *ARMA Covariance Matching Problem* will be a central issue in this letter.

Problem 1: Suppose we are given $n+1$ matrix MA parameters $\{B_k\}$ in the unilateral ARMA model (13) and $n+1$ real $m \times m$ matrices $\Sigma_0, \Sigma_1, \dots, \Sigma_n$, such that the block-Toeplitz matrix

$$\mathbf{T}_n = \begin{bmatrix} \Sigma_0 & \Sigma_1^\top & \Sigma_2^\top & \cdots & \Sigma_n^\top \\ \Sigma_1 & \Sigma_0 & \Sigma_1^\top & \cdots & \Sigma_{n-1}^\top \\ \Sigma_2 & \Sigma_1 & \Sigma_0 & \cdots & \Sigma_{n-2}^\top \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \Sigma_n & \Sigma_{n-1} & \Sigma_{n-2} & \cdots & \Sigma_0 \end{bmatrix}, \qquad n \in \mathbb{Z}_+ \tag{23}$$

is positive definte. We want to determine the matrix parameters $\{A_k,\ k = 0, 1, \dots, n\}$ such that the first $n+1$ covariance matrices of the periodic process $\{y(t)\}$ match the sequence $\{\Sigma_k\}$.

When $B_1 = B_2 = \dots = B_n = 0$, that is the MA part is trivial, this is essentially the *Covariance Extension Problem for n-Reciprocal Processes* discussed and solved in [7]. The solution of the scalar version of this problem, stated in terms of the symmetric polynomials $Q(\zeta)$, $P(\zeta)$ and its solution in terms of generalized maximum entropy is discussed in [16]. Here for reasons that will be apparent in a moment our unknowns will instead be the coefficients of some distinguished spectral factor of the denominator $Q(\zeta)$. An important point is that to get a unique solution $Q(\zeta)$ one needs to fix the numerator polynomial $P(\zeta)$ in (9). It has in fact been shown in [16, Th. 2] that there is a 1:1 continuous relation between positive Hermitian numerator polynomials $P(\zeta)$ and solutions $Q(\zeta)$. Clearly to fix the MA parameters, is equivalent to fixing $P(\zeta)$ and this explains the formulation of Problem 1 with fixed MA parameters $\{B_k\}$.

Observe that the ARMA covariance matching problem is just asking that the submatrix made of the upper-left $(n+1) \times (n+1)$ blocks extracted from the circulant covariance $\mathbf{\Sigma}$ in (4) should match the Toeplitz data covariance (23). To make this precise we need to fix some notations. Let $A, B \in \mathbb{R}^{m \times m(n+1)}$ denote the AR and MA matrix coefficients of the ARMA model (13); i.e.,

$$A = \begin{bmatrix} A_0 & A_1 & \dots & A_n \end{bmatrix}, \quad B = \begin{bmatrix} B_0 & \dots & B_n \end{bmatrix}$$

and let us use the symbol $\mathbf{\Sigma}_n(A)$ to denote the upper-left $(n+1) \times (n+1)$ block-submatrix of $\mathbf{\Sigma}$ written as a function of the unknown AR parameters. Since B is fixed, the covariance matching condition leads to the equation

$$\mathbf{T}_n = \mathbf{\Sigma}_n(A). \tag{24}$$

Proposition 1: The covariance matching condition (24) is equivalent to

$$A\mathbf{T}_n = B\mathbf{W}_n^\top(A), \tag{25}$$

where $\mathbf{W}_n^\top(A)$ is the block-Toeplitz matrix extracted from the block-circulant matrix $\mathbf{W}$

$$\mathbf{W}_n^\top(A) := \begin{bmatrix} W_0^\top & W_{-1}^\top & \cdots & W_{-n}^\top \\ W_1^\top & W_0^\top & \cdots & W_{-n+1}^\top \\ \vdots & \vdots & \ddots & \vdots \\ W_n^\top & W_{n-1}^\top & \cdots & W_0^\top \end{bmatrix}.$$

Proof: Equation (25) can be proven directly by a Yule-Walker type calculation. Specifically, combining the model equation (13) with the convolution representation (21) and using the relation $\mathbb{E}w(t-k)y(t-j)^\top = W_{k-j}^\top$, we easily obtain

$$\sum_{k=0}^{n} A_k \Sigma_{j-k} = \sum_{k=0}^{n} B_k W_{k-j}^\top, \quad j = 0, 1, \ldots, n, \tag{26}$$

which is equivalent to (25). The proof that (24) follows from (25) is indicated in [25]. ∎

With the MA coefficients B and the covariance data $\mathbf{T}_n$ fixed, (25) is a nonlinear equation for the unknown A. Before attempting a solution, we notice from the circulant identity

$$\mathbf{A}\boldsymbol{\Sigma}\mathbf{A}^\top = \mathbf{B}\mathbf{B}^\top \tag{27}$$

that the AR coefficient matrix A solving the matching problem must satisfy the constraint

$$A\mathbf{T}_n A^\top = \sum_{k=0}^{n} B_k B_k^\top = \frac{1}{2N} \sum_{k=-N+1}^{N} P(\zeta_k) := M_P, \tag{28}$$

where the second equality comes from a matrix DFT analog of Parseval formula. M_P is a constant matrix. This constraint is satisfied whenever the model (13) is *normalized*.

The structure of (25) suggests a natural iterative scheme for the solution, namely:

$$A^{(k+1)} = B\mathbf{W}_n^\top(A^{(k)})\mathbf{T}_n^{-1}, \tag{29}$$

with $A^{(0)}$ initialized, e.g., with the output of the Levinson-Whittle algorithm [24] for the data $\{\Sigma_k\,;\, k = 0, 1, \ldots, n\}$. After each iteration, the new $A^{(k+1)}$ does not necessarily satisfy the normalization constraint (28) and thus needs to be scaled by a suitable map, $S(A)$: $A \mapsto K^{-1}A$, for some $m \times m$ nonsingular matrix K. This amounts to solving for K the matrix equation

$$A\mathbf{T}_n A^\top = KM_P K^\top, \tag{30}$$

which can be done by various methods such as Cholesky factorization. To sum up, we propose the following iterative algorithm:

Algorithm

1) Initialize $A^{(0)}$ and set a threshold δ to decide convergence
2) Iterate $A^{(k+1)} = B\mathbf{W}_n(A^{(k)})\mathbf{T}_n^{-1}$
3) Solve (30) with $A \equiv A^{(k+1)}$ to compute the scaling transformation K
4) Set $A^{(k+1)} = K^{-1}A^{(k+1)}$
5) If $\|A^{(k+1)} - A^{(k)}\| > \delta$, go to step 2.

Local convergence of the algorithm is established in the companion paper [25] for scalar processes ($m = 1$). This local convergence result can very likely be extended to the vector case and we have plenty of simulation data confirming this conjecture. However so far some details of the extension still need to be filled in and will be left to a future publication.

IV. Smoothing of Stationary Linear Systems With Boundary Constraints

Consider the following problem. We have a wide-sense stationary zero-mean vector signal $\{x(t)\}$ observed on the finite interval $[-N+1, N]$, the observation channel being described by the linear equation

$$y(t) = Cx(t) + v(t), \qquad t \in [-N+1,\ N] \tag{31}$$

where $\{v(t)\}$ is a stationary white noise with a known covariance matrix $R = R^\top > 0$, independent of $\{x(t)\}$. We want to compute the smoothed estimate $\{\hat{x}(t)\}$ given a finite chunk of observations,

$$\hat{x}(t) := \mathbb{E}\{x(t) \mid y(s),\ s \in [-N+1, N]\},\ t \in [-N+1,\ N]. \tag{32}$$

The right-hand side of (32) is the orthogonal projection onto the Hilbert space of random variables spanned by the components of $\{y(s),\ s \in [-N+1, N]\}$. We shall assume that the process $\{x(t)\}$ has a (stationary) *periodic extension* to the whole integer line $\mathbb{Z}$. Equivalently $\{x(t)\}$ can be imagined to be the restriction to the interval $[-N+1, N]$ of a periodic stationary process defined on $\mathbb{Z}$. There are estimates in [9] for how large should this N be. For short, we shall call $\{x(t)\}$ a *periodic process* and think of it as being defined on the finite modular group $\mathbb{Z}_{2N}$. Even if we do not care about the extension, which we are never going to see, this apparently innocent assumption (which is obviously always legitimate for deterministic signals on finite intervals) has important consequences. In particular, the covariance matrix of the finite string $\{x(t)\,;\, t \in [-N+1, N]\}$ must be *block-circulant* and automatically, $\{x(t)\}$ is associated to *periodic boundary conditions* at the extremes.

We shall first discuss the problem of finding a periodic model (13) defined on the finite discrete interval $[-N+1, N]$ which approximates in a suitable sense a given Gauss-Markov stationary model defined on $\mathbb{Z}$. Consider a stationary signal $\{x(t)\}$ given as the output of the state-space model

$$\begin{cases} \xi(t+1) = F\xi(t) + Hu(t) \\ x(t) = G\xi(t) + Ju(t) \end{cases} \tag{33}$$

where $u(t)$ is a normalized white noise. We assume without loss of generality that the Lyapunov equation $\Pi = F\Pi F^\top + HH^\top$ for the variance matrix of $\xi(t)$ has a unique positive definite solution. Let $D := F\Pi G^\top + HJ^\top$ and let

$$\begin{cases} \Sigma_0 := G\Pi G^\top + JJ^\top \\ \Sigma_k := GF^{k-1}D\,; \qquad k = 1, \ldots, n \end{cases} \tag{34}$$

be the string of the first $n+1$ state covariance matrices. One needs to provide a set of $n+1$ MA coefficients, or equivalently a positive polynomial $P(\zeta)$ to fix the zero dynamics of the system. One possible way to do this is to approximate in

the DFT domain the numerator polynomial of the model (33) or use estimates of its *cepstral coefficients* see [4], [10], [17]. Form with the data $\{\Sigma_k\}$, the block-Toeplitz matrix (23); then the periodic ARMA approximation can be computed by running the covariance matching algorithm described in the previous section.

With the identified approximate model (13) at hand, one can now proceed to compute the solution of the smoothing problem. The procedure is inspired to that for reciprocal processes described in [15, Sec. VI].

Write the observation equation (31) in vector notation as

$$\mathbf{y} = \mathbf{C}\mathbf{x} + \mathbf{v},$$

where $\mathbf{C} = \text{diag}\{C, \dots, C\}$. Then use the standard one-shot solution for the minimum variance Bayesian estimate $\hat{\mathbf{x}}$ [19, p. 29] to get the relation

$$(\Sigma^{-1} + \mathbf{C}^\top \mathbf{R}^{-1}\mathbf{C})\hat{\mathbf{x}} = \mathbf{C}^\top \mathbf{R}^{-1}\mathbf{y}, \tag{35}$$

Substituting (4) into the above equation, the matrix on the left-hand side becomes

$$\mathbf{P}^{-1}\mathbf{Q} + \mathbf{C}^\top \mathbf{R}^{-1}\mathbf{C} = \mathbf{P}^{-1}(\mathbf{Q} + \mathbf{P}\mathbf{C}^\top \mathbf{R}^{-1}\mathbf{C}). \tag{36}$$

Then define

$$\hat{\mathbf{Q}} := \mathbf{Q} + \mathbf{P}\mathbf{C}^\top \mathbf{R}^{-1}\mathbf{C}\,, \tag{37}$$

which is a positive-definite block-circulant since $\mathbf{C}^\top \mathbf{R}^{-1}\mathbf{C}$ is a block-diagonal matrix with positive-semidefinite blocks. In fact, $\hat{\mathbf{Q}}$ is bilaterally banded of bandwidth n since such are both summands on the right-hand side of (37). Then (35) is equivalent to

$$\hat{\mathbf{Q}}\hat{\mathbf{x}} = \mathbf{P}\mathbf{C}^\top \mathbf{R}^{-1}\mathbf{y} := \hat{\mathbf{y}}. \tag{38}$$

In order to carry out a two-sweep smoothing procedure in the style of the Rauch-Striebel-Tung smoother [22], we first perform a banded matrix factorization $\hat{\mathbf{Q}} = \hat{\mathbf{A}}\hat{\mathbf{A}}^\top$, where

$$\hat{\mathbf{A}} = \text{Circ}\{\hat{A}_0, \hat{A}_1, \dots, \hat{A}_n, 0, \dots, 0\}. \tag{39}$$

As discussed in [25], such a factorization is possible if N is taken large enough; it can be computed in the spectral domain by standard matrix polynomial factorization algorithms, see [23]. Then, given $\hat{\mathbf{A}}$ and $\hat{\mathbf{y}}$, to compute the solution to (38) we first perform a forward sweep described by

$$\hat{\mathbf{A}}\mathbf{z} = \hat{\mathbf{y}}, \tag{40}$$

and then a backward sweep

$$\hat{\mathbf{A}}^\top \hat{\mathbf{x}} = \mathbf{z}\,. \tag{41}$$

The two sweeps can be implemented by a forward and a backward recursive algorithm described by unilateral AR models. To this end we need to attach to them explicit boundary values $\hat{x}(-N+1), \hat{x}(-N+2), \dots, \hat{x}(-N+n)$ and $\hat{x}(N-n+1), \dots, \hat{x}(N)$ extracted from the process $\{\hat{x}(t)\}$, which we assume are given. Due to the lower block-triangular circulant structure of $\hat{\mathbf{A}}$ exactly like (17), the first equation of the forward sweep can be written as

$$\hat{A}_0 z(-N+1) = -\sum_{i=1}^{n} \hat{A}_i z(N-i+1) + \hat{y}(-N+1)\,, \tag{42}$$

which needs to be initialized with the boundary values $z(N-n+1),\ z(N-n+2),\ \dots, z(N)$. These values can be obtained by solving for $\mathbf{z}$ the last n block equations in the backward sweep (41) since only the boundary values at two ends of $\hat{\mathbf{x}}$ are involved there due to the banded upper-triangular block-circulant structure of $\hat{\mathbf{A}}^\top$.

The forward sweep starts by computing the boundary values $z(N-n+1), \dots, z(N)$. After these n endpoint boundary values are available, the recursion for $\mathbf{z}$ can be implemented by the scheme

$$z(t) = \hat{A}_0^{-1}\left[\hat{y}(t) - \sum_{i=1}^{n} \hat{A}_i z(t-i)\right], \quad t \in [-N+1, N-n]. \tag{43}$$

One should notice that in this notation, we impose implicitly that $z(-N) = z(N), \dots, z(-N-n+1) = z(N-n+1)$. The backward sweep then proceeds by using

$$\hat{x}(t) = \hat{A}_0^{-\top}\left[z(t) - \sum_{i=1}^{n} \hat{A}_i^\top \hat{x}(t+i)\right], \tag{44}$$

$$t \in [-N+n+1, N-n], \tag{45}$$

which is initialized with the known terminal boundary values $\hat{x}(N-n+1), \hat{x}(N-n+2), \dots, \hat{x}(N)$.

There is also a dual factorization which would lead to a backward-forward sequence of sweeps but we shall not insist on this point.

A numerical example: Just to show the feasibility of the method, we shall discuss a toy example. For this particular example we have chosen scalar MA coefficients resulting in obvious computational advantage for the smoothing algorithm. We should stress that this example is not meant to reflect any realistic situation. Referring to model (33), fix the matrices

$$F = \begin{bmatrix} 0.9 & -0.3 \\ 0.3 & 0.9 \end{bmatrix}, \quad G = \begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix} \tag{46}$$

and $H,\ J$ equal to identity. For the observation process (31), we take

$$C = \begin{bmatrix} 1 & 1 \end{bmatrix}.$$

The eigenvalues of A are $0.9 \pm 0.3i$ with a modulus 0.9487.

To compute the smoothed process (32), we first build a periodic ARMA model of order $n = 1$ to approximately describe the process $\{x(t)\}$ on a finite interval by matching the first two steady-state covariances

$$\Sigma_0 = \begin{bmatrix} 51 & 10 \\ 10 & 11 \end{bmatrix}, \quad \Sigma_1 = \begin{bmatrix} 46 & 17 \\ 4 & 9 \end{bmatrix}$$

computed with the formulae (34). The period of interest is set as $2N = 50$ and the MA parameters are chosen (quite arbitrarily) as $b_0 = 0.4893,\ b_1 = 0.3377$. The unilateral ARMA model looks like

$$A_0 x(t) + A_1 x(t-1) = b_0 w(t) + b_1 w(t-1), \tag{47}$$

and the AR parameters are computed with the algorithm of Section III

$$A_0 = \begin{bmatrix} 0.3725 & 0 \\ 0.1324 & 0.3341 \end{bmatrix},\ A_1 = \begin{bmatrix} -0.2571 & -0.3579 \\ -0.0739 & -0.3659 \end{bmatrix}.$$

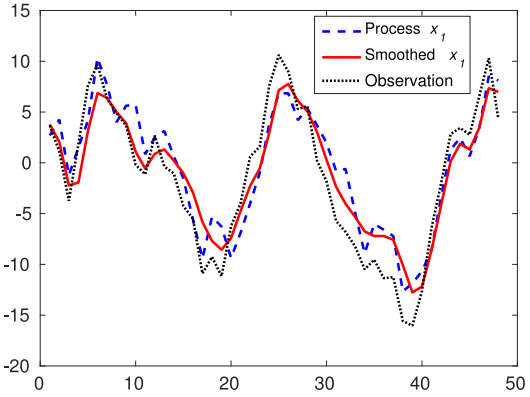

Fig. 1. Result of smoothing for x_1.

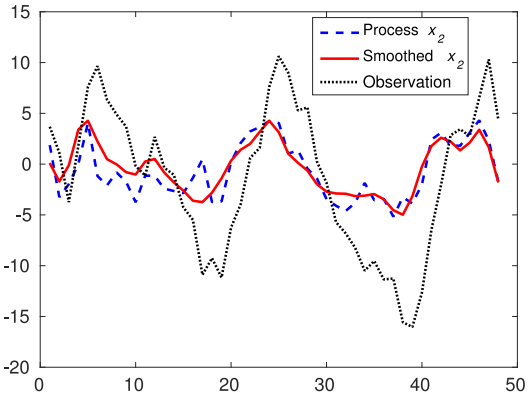

Fig. 2. Result of smoothing for x_2.

The resulted poles of (47), i.e., roots of the equation $\det A(z) = 0$ are $0.7022 \pm 0.2240i$ of modulus 0.7371.

Given the approximate model (47) and the observation process (31), the two-sweep smoothing algorithm described in the previous part can be implemented. The two components of the smoothed process $\{\hat{x}(t)\}$ are shown in Figures 1 and 2. The effect of smoothing is appreciable.

V. Conclusion

We have presented a new approach to the rational covariance extension problem for vector processes on a finite interval which is not based on the minimization of a generalized entropy functional. The approach leads to the solution of a nonlinear system of equations by an iterative algorithm which has been proven to converge at least in the scalar case [25]. The proof for the vector case will be presented elsewhere. We have indicated how to solve a class of finite-interval smoothing problems subjected to boundary constraints. Using this algorithm an approximate stationary periodic model for the original underlying process is constructed based on matching a finite number of the original covariance lags.

References

[1] 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. Autom. Control*, vol. 40, no. 11, pp. 1841–1857, Nov. 1995.

[2] A. Blomqvist, A. Lindquist, and R. Nagamune, "Matrix-valued Nevanlinna—Pick interpolation with complexity constraint: An optimization approach," *IEEE Trans. Autom. Control*, vol. 48, no. 12, pp. 2172–2190, Dec. 2003.

[3] C. I. Byrnes, S. V. Gusev, and A. Lindquist, "A convex optimization approach to the rational covariance extension problem," *SIAM J. Control Optim.*, vol. 37, no. 1, pp. 211–229, 1999.

[4] C. I. Byrnes, P. Enqvist, and A. Lindquist, "Cepstral coefficients, covariance lags, and pole-zero models for finite data strings," *IEEE Trans. Signal Process.*, vol. 49, no. 4, pp. 677–693, Apr. 2001.

[5] C. I. Byrnes, S. V. Gusev, and A. Lindquist, "From finite covariance windows to modeling filters: A convex optimization approach," *SIAM Rev.*, vol. 43, no. 4, pp. 645–675, 2001.

[6] C. I. Byrnes, P. Enqvist, and A. Lindquist, "Identifiability and well-posedness of shaping-filter parameterizations: A global analysis approach," *SIAM J. Control Optim.*, vol. 41, no. 1, pp. 23–59, 2002.

[7] F. P. Carli, A. Ferrante, M. Pavon, and G. Picci, "A maximum entropy solution of the covariance extension problem for reciprocal processes," *IEEE Trans. Autom. Control*, vol. AC-56, no. 9, pp. 1999–2012, Sep. 2011.

[8] F. P. Carli and T. T. Georgiou, "On the covariance completion problem under a circulant structure," *IEEE Trans. Autom. Control*, vol. 56, no. 4, pp. 918–922, Apr. 2011.

[9] A. Dembo, C. L. Mallows, and L. A. Shepp, "Embedding nonnegative definite Toeplitz matrices in nonnegative definite circulant matrices, with application to covariance estimation," *IEEE Trans. Inf. Theory*, vol. IT-35, no. 6, pp. 1206–1212, Nov. 1989.

[10] P. Enqvist, "A convex optimization approach to ARMA(n,m) model design from covariance and cepstral data," *SIAM J. Control Optim.*, vol. 43, no. 3, pp. 1011–1036, 2004.

[11] T. T. Georgiou, "Partial realization of covariance sequences," Ph.D. dissertation, CMST, Univ. Florida, Gainesville, FL, USA, 1983.

[12] T. T. Georgiou, "Realization of power spectra from partial covariances," *IEEE Trans. Acoust. Speech Signal Process.*, vol. ASSP-35, no. 4, pp. 438–449, Apr. 1987.

[13] R. E. Kalman, "Realization of covariance sequences," in *Proc. Toeplitz Memorial Conf.*, Tel Aviv, Israel, 1981, pp. 331–342.

[14] A. J. Krener, "Filtering boundary value discrete time linear systems," in *Proc. IEEE Conf. Decis. Control*, Florence, Italy, 2013, pp. 7684–7689.

[15] B. C. Levy, R. Frezza, and A. J. Krener, "Modeling and estimation of discrete-time Gaussian reciprocal processes," *IEEE Trans. Autom. Control*, vol. AC-35, no. 9, pp. 1013–1023, Sep. 1990.

[16] A. Lindquist, C. Masiero, and G. Picci, "On the multivariate circulant rational covariance extension problem," in *Proc. IEEE Conf. Decis. Control*, Florence, Italy, 2013, pp. 7155–7161.

[17] A. G. Lindquist and G. Picci, "The circulant rational covariance extension problem: The complete solution," *IEEE Trans. Autom. control*, vol. 58, no. 11, pp. 2848–2861, Nov. 2013.

[18] A. Lindquist and G. Picci, "Modeling of stationary periodic time series by ARMA representations," in *Optimization and Its Applications in Control and Data Sciences* (Springer Optimization and Its Applications), vol. 115. Cham, Switzerland: Springer, pp. 281–314.

[19] A. Lindquist and G. Picci, *Linear Stochastic Systems: A Geometric Approach to Modeling, Estimation and Identification* (Springer Series in Contemporary Mathematics). Heidelberg, Germany: Springer Verlag, 2015.

[20] P. Masani, "The prediction theory of multivariate stochastic processes, III," *Acta Math.*, vol. 104, nos. 1–2, pp. 141–162, 1960.

[21] F. Carli and G. Picci, "On the factorization approach to band extension of block-circulant matrices," in *Proc. 19th Int. Symp. Math. Theory Netw. Syst. (MTNS)*, Budapest, Hungary, 2010, pp. 907–914.

[22] H. E. Rauch, C. T. Striebel, and F. Tung, "Maximum likelihood estimates of linear dynamic systems," *AIAA J.*, vol. 3, no. 8, pp. 1445–1450, 1965.

[23] J. Rissanen, "Algorithms for triangular decomposition of block-Hankel and Toeplitz matrices with applications to factoring positive matrix polynomials," *Math. Comput.*, vol. 27, no. 121, pp. 147–154, 1973.

[24] P. Whittle, "On the fitting of multivariate autoregressions, and the approximate canonical factorization of a spectral density matrix," *Biometrika*, vol. 50, nos. 1–2, pp. 129–134, 1963.

[25] B. Zhu and G. Picci, "Proof of local convergence of a new algorithm for covariance matching of periodic ARMA models," *IEEE Control Syst. Lett.*, vol. 1, no. 1, pp. 206–211, Jul. 2017.