ON THE FACTORIZATION APPROACH TO BAND EXTENSION OF BLOCK-CIRCULANT MATRICES

Francesca Carli and Giorgio Picci

***Abstract*— The following problem occurs in modeling and estimation of stationary processes defined on a finite interval $[1, N]$ of the integer line: complete a given finite sequence $\Sigma_0, \dots, \Sigma_n$; $n < N$, of covariance matrices to form a symmetric block-Toeplitz matrix which has a *banded symmetric block-circulant* inverse of bandwidth n. This is a matrix band extension problem which does not seem to have been considered before. It looks apparently alike the classical block-Toplitz band extension problem for covariance matrices on $\mathbb{Z}$. We show that it cannot be approached by the standard factorization techniques used in the literature to do band extension for Toeplitz matrices.**

I. Introduction: the block-Toeplitz band extension problem

In section IV of this paper we shall formulate and discuss a matrix band extension problem for block-circulant matrices. Although this problem does not seem to have been considered before, it looks very similar to a classical band extension problem for Toeplitz matrices which we shall review below. Our point will be to show that the standard factorization techniques used in the literature to solve it, unfortunately do not apply to the circulant case.

Consider then the following classical problem:

Problem 1: Given a finite sequence of $m \times m$ matrices

$$\Sigma_0, \dots, \Sigma_n \,, \tag{I.1}$$

such that

$$T_n := \begin{bmatrix} \Sigma_0 & \Sigma_1^\top & \dots & \Sigma_n^\top \\ \Sigma_1 & \Sigma_0 & \ddots & \ddots \\ \vdots & \ddots & \ddots & \\ \Sigma_n & \ddots & \ddots & \Sigma_0 \end{bmatrix} > 0 \tag{I.2}$$

complete it to form an infinite symmetric block-Toeplitz matrix

$$\boldsymbol{\Sigma} := \begin{bmatrix} \Sigma_0 & \Sigma_1^\top & \dots & \Sigma_n^\top & ? & ? & \dots \\ \Sigma_1 & \Sigma_0 & \Sigma_1^\top & \dots & \Sigma_n^\top & ? & \dots \\ \vdots & \Sigma_1 & \Sigma_0 & \Sigma_1^\top & \dots & \Sigma_n^\top & \dots \\ \Sigma_n & \dots & \ddots & \ddots & \ddots & & \dots \\ ? & \Sigma_n & \dots & \Sigma_1 & \Sigma_0 & \Sigma_1^\top & \dots \\ ? & ? & \Sigma_n & \dots & \Sigma_1 & \Sigma_0 & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \ddots \end{bmatrix} \tag{I.3}$$

which admits an inverse $\Lambda := \boldsymbol{\Sigma}^{-1}$ with a *block-banded structure*; i.e. such that

$$\Lambda_{i,j} = 0 \quad \text{for} \quad |i - j| > n \,. \tag{I.4}$$

This problem is essentially the well-known *block-Toeplitz band extension problems* which has a long history in the literature, see e.g. [3], [4, pp 892-979] and the references therein. Positive definite symmetric block-Toeplitz matrices have a natural interpretation as covariance matrices of stationary processes. Recall that the *covariance sequence* $\{\Sigma_k \,;\, k \in \mathbb{Z}\}$ of a real, stationary, zero-mean m-dimensional stochastic process $\mathbf{y}$ is defined as

$$\Sigma_k = \mathbb{E}\, \mathbf{y}(t+k)\mathbf{y}(t)^\top \,, \qquad k = \dots -1, 0, 1, \dots$$

where $\mathbb{E}$ denotes mathematical expectation. This sequence is symmetric; i.e. $\Sigma_{-k} = \Sigma_k^\top$ and of positive type; i.e. for arbitrary N and arbitrary m-vectors $a_1, a_2, \dots, a_N$ it must hold that

$$\sum_{k,j=1}^{N} a_k^\top \Sigma_{k-j}\, a_j \ge 0 \,. \tag{I.5}$$

Since the block-Toeplitz matrix T_n (I.2) is positive definite, it can be given the meaning of a *partial* (i.e. finite) covariance matrix of some zero-mean m-dimensional stationary process $\mathbf{y}$, namely

$$T_n = \mathbb{E} \begin{bmatrix} \mathbf{y}(t) \\ \vdots \\ \mathbf{y}(t+n) \end{bmatrix} \begin{bmatrix} \mathbf{y}(t) \\ \vdots \\ \mathbf{y}(t+n) \end{bmatrix}^\top , \tag{I.6}$$

Hence, in a probabilistic language, Problem 1 can be seen as the problem of "extending" (or "completing") the $n+1$ random variables $\begin{bmatrix}\mathbf{y}(t) & \dots & \mathbf{y}(t+n)\end{bmatrix}$ of given covariance (I.6), to form a stationary process on the whole integer line $\mathbb{Z}$, in such a way that the inverse of its covariance matrix has a banded structure. This banded structure has a particular statistical interpretation which will be clarified later on.

II. Block-Toeplitz band extension and the Schur-Levinson polynomials

We shall introduce the symbol $\mathbf{y}_{[t-n,t)} \equiv \mathbf{y}_{[t-n,\,t-1]}$ to denote the finite past history $\{\mathbf{y}(t-n), \dots, \mathbf{y}(t-1)\}$ of length n of $\mathbf{y}$ at time t and, similarly, $\mathbf{y}_{(t,t+n]} \equiv \mathbf{y}_{[t+1,t+n]}$ to denote the finite future string $\{\mathbf{y}(t{+}1), \dots, \mathbf{y}(t{+}n)\}$. The Hilbert space of scalar random variables linearly generated by the (scalar) components of the random vectors $\{\mathbf{y}(t-k), \dots, \mathbf{y}(t)\}$ is denoted $\mathbf{H}(\,\mathbf{y}_{[t-k,t]})$. $\hat{\mathbb{E}}\,[\mathbf{z} \mid \cdot]$ stands for the vector of orthogonal projections of the components of $\mathbf{z}$ onto

This work was partially supported by the ministery for higher education and scientific research of Italy (MIUR) under a PRIN grant.

Francesca Carli and Giorgio Picci are with the Department of Information Engineering – University of Padova, via Gradenigo 6/B, 35131 Padova, Italy; e-mail: carlifra@dei.unipd.it; picci@dei.unipd.it

a Hilbert space of random variables specified in the second argument. We write $\mathbf{z} \perp \mathbf{H}$ to mean $\mathbb{E}\,\mathbf{z}\mathbf{x} = 0$ for all $\mathbf{x} \in \mathbf{H}$.

The *forward and backward innovation processes of finite memory* n of $\mathbf{y}$ are defined as

$$\mathbf{e}_n(t) := \mathbf{y}(t) - \hat{\mathbb{E}}\,[\mathbf{y}(t) \mid \mathbf{y}_{[t-n,t)}], \tag{II.1}$$

$$\bar{\mathbf{e}}_n(t) := \mathbf{y}(t) - \hat{\mathbb{E}}\,[\mathbf{y}(t) \mid \mathbf{y}_{(t,t+n]}], \tag{II.2}$$

for $t \in \mathbb{Z}$. These are stationary processes; in particular (II.2) can equivalently be written as

$$\bar{\mathbf{e}}_n(t-n) = \mathbf{y}(t-n) - \hat{\mathbb{E}}\,[\mathbf{y}(t-n) \mid \mathbf{y}_{(t-n,t]}]\,. \tag{II.3}$$

The following well-known lemma describes the orthogonality properties of the finite memory innovation processes. For completeness we provide a quick proof.

Lemma 2.1: For every $t \in \mathbb{Z}$ the random vectors $\bar{\mathbf{e}}_0(t), \dots, \bar{\mathbf{e}}_k(t-k)$ form an orthogonal basis for the Hilbert space $\mathbf{H}(\mathbf{y}_{[t-k,t]})$, in other terms

$$\mathrm{span}\{\bar{\mathbf{e}}_0(t), \dots, \bar{\mathbf{e}}_k(t-k)\} = \mathbf{H}(\mathbf{y}_{[t-k,t]}) \tag{II.4}$$

$$\bar{\mathbf{e}}_{k+1}(t-k-1) \perp \mathbf{H}(\mathbf{y}_{[t-k,t]}) \tag{II.5}$$

for all t and $k \geq 0$.

Dually, the random vectors $\mathbf{e}_0(t), \mathbf{e}_1(t+1), \dots, \mathbf{e}_k(t+k), \dots$ form an orthogonal basis for the space $\mathbf{H}(\mathbf{y}_{[t,t+k]})$; i.e.

$$\mathrm{span}\,\{\mathbf{e}_0(t), \mathbf{e}_1(t+1), \dots, \mathbf{e}_k(t+k)\} = \mathbf{H}(\mathbf{y}_{[t,t+k]}) \tag{II.6}$$

$$\mathbf{e}_{k+1}(t+k+1) \perp \mathbf{H}(\mathbf{y}_{[t,t+k]}) \tag{II.7}$$

for all t and $k \geq 0$.

Proof: Indeed, for arbitrary k the vectors $\bar{\mathbf{e}}_0(t), \dots, \bar{\mathbf{e}}_k(t-k)$ are obtained by sequential orthogonalization of the random vectors $\{\mathbf{y}(t), \mathbf{y}(t-1), \dots \mathbf{y}(t-k)\}$ in that order, while $\mathbf{e}_0(t), \mathbf{e}_1(t+1), \dots, \mathbf{e}_k(t+k)$ are obtained by sequential orthogonalization of $\{\mathbf{y}(t), \mathbf{y}(t+1), \dots \mathbf{y}(t+k)\}$. ■

Since $\mathbf{e}_n(t)$ is a stationary linear function of the past $\mathbf{y}_{[t-n,t)}$ it can be represented as

$$\mathbf{e}_n(t) = \sum_{0}^{n} A_n(k)\,\mathbf{y}(t-k) \quad . \tag{II.8}$$

where the matrices $\{-A_n(k)\,;\, k = 1, 2, \dots, n\}$ are the coefficients of the (forward) linear predictor of $\mathbf{y}(t)$, based on the past $\mathbf{y}_{[t-n,t)}$ of finite length n and $A_n(0) = I$, the $m \times m$ identity matrix. Dually,

$$\bar{\mathbf{e}}_n(t) = \sum_{0}^{n} B_n(k)\,\mathbf{y}(t+k) \quad , \tag{II.9}$$

where $B_n(0) = I$ and the matrices $-B_n(k)$, $k = 1, \dots, n$ are the coefficients of the backward predictor of $\mathbf{y}(t)$, based on the future history $\mathbf{y}_{(t,t+n]}$ of finite length n.

In the following we shall represent the forward shift operator on an infinite sequence $\xi \equiv \{\xi(t)\,;\, t \in \mathbb{Z}\}$, as multiplication by the symbol z; namely

$$[z\xi](t) := \xi(t+1) \qquad t \in \mathbb{Z}\,.$$

With this convention the polynomials

$$A_n(z^{-1}) := \sum_{0}^{n} A_n(k)\, z^{-k} \qquad B_n(z) := \sum_{0}^{n} B_n(k)\, z^{k} \tag{II.10}$$

act on the process $\mathbf{y}$ to produce the forward and backward innovations of memory n. They are the well-known (matrix) Szegö-Levinson polynomials of degree n. $B_n(z)$ is the *dual* or *backward* of $A_n(z^{-1})$.

Consider now the $(n+1)m \times (n+1)m$ lower triangular matrix L_n, built by stacking the coefficients of the linear predictors of orders one up to n in the following way

$$L_n := \begin{bmatrix} I & 0 & \dots & 0 \\ A_1(1) & I & \ddots & \vdots \\ \vdots & & \ddots & 0 \\ A_n(n) & \dots & A_n(1) & I \end{bmatrix} \tag{II.11}$$

which is obviously invertible. We have

$$\begin{bmatrix} \mathbf{e}_0(t) \\ \mathbf{e}_1(t+1) \\ \vdots \\ \mathbf{e}_{n-1}(t+n-1) \\ \mathbf{e}_n(t+n) \end{bmatrix} = L_n \begin{bmatrix} \mathbf{y}(t) \\ \mathbf{y}(t+1) \\ \vdots \\ \mathbf{y}(t+n-1) \\ \mathbf{y}(t+n) \end{bmatrix}, \tag{II.12}$$

from which, computing the variance of both members and taking into account the orthogonality properties described in Lemma 2.1, we get the block-diagonalization[1]

$$L_n\, T_n\, L_n^{\top} = \mathrm{diag}\{\Delta_0, \dots, \Delta_{n-1}, \Delta_n\} := \mathbf{\Delta}_n\,. \tag{II.13}$$

Right-multiplying by $L_n^{-\top}$, we obtain

$$L_n\, T_n = \mathbf{\Delta}_n\, L_n^{-\top} \tag{II.14}$$

where $L_n^{-\top}$ is upper block-triangular with identities on the main diagonal, so that,

$$\begin{bmatrix} I & 0 & \dots & 0 \\ A_1(1) & I & \ddots & \vdots \\ \vdots & & \ddots & 0 \\ A_n(n) & \dots & A_n(1) & I \end{bmatrix} \begin{bmatrix} \Sigma_0 & \Sigma_1^{\top} & \dots & \Sigma_n^{\top} \\ \Sigma_1 & \Sigma_0 & \ddots & \ddots \\ \vdots & \ddots & \ddots & \\ \Sigma_n & \ddots & \ddots & \Sigma_0 \end{bmatrix} = \begin{bmatrix} \Delta_0 & * & \dots & * \\ 0 & \Delta_1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \dots & 0 & \Delta_n \end{bmatrix}. \tag{II.15}$$

This equation, which can also be found in [3], is just an equivalent form of the well-known *Yule-Walker equations* for the data set (I.1). In fact, expanding the last block row,

[1] Incidentally, using this factorization one obtains a "fast" algorithm for computing the inverse T_n^{-1}, by the Levinson algorithm; a generalization of a classical result of W. Trench [10].

one obtains the n-th order Yule-Walker equations for the coefficients $A_n(k)$, namely

$$\Sigma_k + \sum_{j=1}^{n} A_n(j)\,\Sigma_{k-j} \;=\; \Delta_n \delta_k\,, \qquad k = 0, 1, \ldots, n \tag{II.16}$$

where $\delta_k = 1$ for $k = 0$ and zero otherwise and $\Sigma_{-j} := \Sigma_j^\top$, according to the interpretation of Σ_k as a covariance matrix. Since T_n is symmetric and strictly positive definite, it has a unique normalized lower block-triangular factorization, which can be written as $T_n = L_n^{-1}\,\boldsymbol{\Delta}_n\,[L_n^{-1}]^\top$. Hence the matrix equation (II.14) (equivalently,) has a unique solution L_n with $\boldsymbol{\Delta}_n$ block-diagonal strictly positive definite. It follows in particular that the system (II.16) also has a unique solution $(\{A_n(j)\,;\, j = 1, 2, \ldots n\},\, \Delta_n)$ with $\Delta_n > 0$.
Due to the nested triangular structure of the equations, the same argument works for any $n = 0, 1, \ldots$ so that equations similar to (II.16) hold for the coefficients of the linear predictors of order one up to n, $\{A_\nu(k)\,;\, k = 1, \ldots, \nu;\, \nu \le n\}$, thereby uniquely determined together with the innovation variances Δ_ν for $\nu = 0, 1, \ldots, n$.

We can now exhibit the solution to the block-Toeplitz band extension problem.

Theorem 2.1: Let $N > n$ be an arbitrary integer. Form the lower triangular $m(N+1) \times m(N+1)$ banded matrix

$$L_N := \begin{bmatrix} I & 0 & \ldots & \ldots & \ldots & \ldots & 0 \\ A_1(1) & I & 0 & & & & \vdots \\ \vdots & & & \ddots & & & \vdots \\ A_n(n) & \ldots & A_n(1) & I & 0 & & \vdots \\ 0 & \ddots & & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & A_n(n) & \ldots & A_n(1) & I \end{bmatrix}, \tag{II.17}$$

define the block-diagonal matrix with N entries

$$\boldsymbol{\Delta}_N := \mathrm{diag}\{\Delta_0, \ldots, \Delta_{n-1}, \Delta_n, \ldots, \Delta_n\} \tag{II.18}$$

and let

$$T_N := L_N^{-1}\,\boldsymbol{\Delta}_N\,[L_N^{-1}]^\top\,. \tag{II.19}$$

Then,

1) *The upper left $m(n+1) \times m(n+1)$ principal submatrix of T_N is equal to T_n. In fact, T_N is a symmetric positive extension of T_n.*
2) *The inverse T_N^{-1} is symmetric banded of bandwidth n.*

Proof: It is apparent that the upper left $m(n+1) \times m(n+1)$ principal submatrix of T_N, say $\tilde{T}_n$, is obtained by extracting the upper left $m(n+1) \times m(n+1)$ principal submatrix from all three factors in the right member of (II.19) so that,

$$\tilde{T}_n = L_n^{-1}\,\boldsymbol{\Delta}_n\,[L_n^{-1}]^\top$$

which is equivalent to

$$L_n \tilde{T}_n L_n^\top = \boldsymbol{\Delta}_n\,.$$

Comparing with the factorization (II.13) we see that $\tilde{T}_n$ must actually be equal to T_n and the first statement of the theorem follows.

To prove the second statement, note that the inverse T_N^{-1} admits the factorization

$$T_N^{-1} = L_N^\top\,\boldsymbol{\Delta}_N^{-1}\,L_N \tag{II.20}$$

where $L_N^\top$ and hence also $L_N^\top\,\boldsymbol{\Delta}_N^{-1}$ is banded upper triangular of width n, while L_N is banded lower triangular also of width n. It is then easy to check that T_N^{-1} is a *block-banded matrix of bandwidth n.* ■

Since this argument holds for arbitrarily large N, it follows that formula (II.19) solves the Toeplitz band extension problem.
There is an efficient way of solving the Yule Walker equations by the celebrated *Levinson Algorithm*, whose matrix version is due to Whittle, [11], which computes recursively the pair $(A_n(z^{-1}),\, B_n(z))$.

III. Band extension and AR processes

In this section we shall show that the Toeplitz band extension problem is bi-uniquely related to auto-regressive interpolation.

Based on the covariance data (I.1), define a stationary auto-regressive (AR) process of order n [2] $\mathbf{y}_n$ by letting

$$\mathbf{y}_n(t) + \sum_{k=1}^{n} A_n(k)\mathbf{y}_n(t-k) = \mathbf{e}(t) \qquad t \in \mathbb{Z} \tag{III.1}$$

where the $A_n(k)$'s are the coefficients of the n-th Levinson Polynomial $A_n(z^{-1})$ of the sequence (I.1) and $\mathbf{e}$ is a m-dimensional *white noise process* having variance matrix Δ_n (positive definite). The process is well defined on the whole integer line since the polynomial $A_n(z^{-1})$ cannot have zeros on (or ouside of) the unit circle of the complex plane [11]. We shall write symbolically $A(z^{-1})\mathbf{y}_n(t) = \mathbf{e}(t)$. The following result relates to facts which are well-known in the stochastic system and signal-processing community. In the absence of a specific reference and for the sake of completeness we shall provide a quick proof anyway.

Proposition 3.1: The process $\mathbf{y}_n$ is well defined by (III.1) *for all $t \in \mathbb{Z}$ and $\mathbf{e}$ is its innovation (the minimum variance prediction error based on the whole past history). In fact we have $\mathbf{e} \equiv \mathbf{e}_n$. The variance matrix of the vector obtained by stacking $n+1$ successive random variables of the process $\mathbf{y}_n$, is exactly T_n. In other words, formula* (I.6) *holds with $\mathbf{y} \equiv \mathbf{y}_n$.*

Proof: Since the polynomial $A_n(z^{-1})$ cannot have zeros on (or ouside of) the unit circle of the complex plane the solution $\mathbf{y}_n(t)$ of the difference equation (III.1) is a well-defined linear functional of the (infinite) past history at time t of the white noise $\mathbf{e}$, for all times $t \in \mathbb{Z}$. Since $\mathbf{e}$ is white we have $\mathbf{e}(t) \perp \mathbf{y}_n(t-s)$ for all $s > 0$ so that $\mathbf{e}(t)$ is indeed the one step ahead prediction error of $\mathbf{y}_n(t)$. Clearly, because of the AR structure (III.1), the

[2] The *order* of an autoregressive model is the number of lags appearing in the difference equation.

predictor only depends on the n previous random variables $\{\mathbf{y}_n(t-k)\,;\,k=1,2,\ldots,n\}$ so that $\mathbf{e}\equiv\mathbf{e}_n$. On the other hand $\mathbb{E}\,\mathbf{y}_n(t)\mathbf{e}^\top(t)=\mathbb{E}\,\mathbf{e}_n(t)\mathbf{e}_n^\top(t)=\Delta_n$ since $\mathbf{e}(t)\perp-\sum_{k=1}^{n}A_n(k)\mathbf{y}_n(t-k)$. Therefore by multiplying (III.1) from the right by $\mathbf{y}(t-s)$ for $s\geq 0$ one obtains the Yule-Walker difference equation for the covariance sequence $\Sigma_0,\,\Sigma_1,\,\ldots,\Sigma_n$. ■

For this reason we may call a process $\mathbf{y}_n$ for which the last property of the proposition holds, an *extension*, or an *AR interpolant* of order n, of the original process $\mathbf{y}$ (which indeed was only "partially" defined).

Since the process $\mathbf{y}_n$ is stationary and defined on the whole line, it has a well-defined infinite covariance sequence, say $\{\hat\Sigma_k,\,k\geq 0\}$ which *a fortiori* satisfies the interpolation conditions

$$\hat\Sigma_k=\Sigma_k \qquad k\leq n\,.$$

Hence the (infinite) covariance sequence of $\mathbf{y}_n$ fills an infinite positive block-Toeplitz extension of T_n. *This extension is in fact the unique positive extension satisfying the banded inverse condition* (I.4). For assume that, based on this infinite covariance string, we keep on computing Levinson polynomials and finite memory innovations of arbitrary order. We easily end up discovering that,

Proposition 3.2: The forward (and backward) Levinson polynomials of index ν of the sequence $\{\hat\Sigma_k,\,k\geq 0\}$ remain the same for all $\nu\geq n$ and we have

$$A_\nu(z^{-1})\equiv A_n(z^{-1})\,,\qquad \Delta_\nu=\Delta_n=\mathbb{E}\,\mathbf{e}(t)\mathbf{e}(t)^\top\quad \nu\geq n\,. \tag{III.2}$$

The finite memory innovations $\{\mathbf{e}_{n+k}\,;\,k\geq 0\}$ are white process and all coincide with the true innovation $\mathbf{e}\equiv\mathbf{e}_n$ of the process $\mathbf{y}_n$. A completely symmetric formula holds for the backward polynomials.

The above implies that for all $N\geq n$ we have the representation

$$\begin{bmatrix}\mathbf{e}_0(t)\\ \mathbf{e}_1(t+1)\\ \vdots\\ \mathbf{e}_{n-1}(t+n-1)\\ \mathbf{e}(t+n)\\ \vdots\\ \mathbf{e}(t+N)\end{bmatrix}=L_N\begin{bmatrix}\mathbf{y}_n(t)\\ \mathbf{y}_n(t+1)\\ \vdots\\ \mathbf{y}_n(t+n-1)\\ \mathbf{y}_n(t+n)\\ \vdots\\ \mathbf{y}_n(t+N)\end{bmatrix} \tag{III.3}$$

which, computing the variance matrices of the two members, shows that the covariance matrix, T_N, of the first $N+1$ successive variables of the process $\mathbf{y}_n$, satisfies the equation,

$$\mathbf{\Delta}_N=L_N\,T_N\,L_N^\top\,,$$

where L_N is banded lower triangular of bandwidth n. Compare with (II.19). Extracting T_N^{-1} from this relation one obtains (II.20) which indeed shows that T_N^{-1} is banded. The argument holding for arbitrary N, we come to the following conclusion:

Theorem 3.1: The covariance matrix, T_N, of the first $N+1$ successive variables of a full rank AR process of order n on $\mathbb{Z}$ has a banded inverse of bandwidth n. In fact, it can be block-diagonalized as

$$L_NT_NL_N^\top=\mathrm{diag}\{\Delta_0,\ldots,\Delta_N\}:=\mathbf{\Delta}_N\,, \tag{III.4}$$

where the matrix L_N, lower triangular of bandwidth n, is formed from the Levinson polynomials of the process as in (II.17).

Conversely, let T_N be a positive definite block-Toeplitz covariance matrix with N blocks. If there exists a lower banded block-Toeplitz matrix L_N of bandwidth n which block-diagonalizes T_N; i.e. such that (III.4) *holds with $\mathbf{\Delta}_N$ positive definite, then the inverse T_N^{-1} is a symmetric bilaterally banded matrix of bandwidth n. In this case, T_N is the covariance matrix of an AR process of order n on $\mathbb{Z}$.*

Next, we shall rephrase the discussion above in terms of spectral densities and Fourier Transform. Let $\Psi(z)$ be an $m\times m$ spectral density matrix which together with its inverse $\Psi^{-1}(z)$ belongs to L^∞ of the unit circle and let $T(\Psi)$ denote the infinite block-Toeplitz matrix with symbol Ψ, namely the symmetric block-Toeplitz matrix constructed with the Fourier coefficients (covariances) of Ψ or, equivalently, with those of the Laurent expansion of $\Psi(z)$ in case Ψ admits an analytic extension about the unit circle. Under this assumption $T(\Psi)$ is a bounded invertible operator on the Hilbert space of (real) square-summable sequences, see e.g. [5].

It follows by well-known results in harmonic analysis that the spectral density of the AR process $\mathbf{y}_n$ is a rational matrix function having the expression

$$\Phi_n(z):=A_n(z^{-1})^{-1}\Delta_n\,[A_n(z)^{-1}]^\top \tag{III.5}$$

and since by stationarity the polynomial $A_n(z^{-1})$ cannot have zeros on (or ouside of) the unit circle of the complex plane [11], the rational matrix $\Phi_n(z)$ admits a Laurent expansion

$$\Phi_n(z):=\sum_{k=-\infty}^{+\infty}\hat\Sigma_k\,z^{-k}\,, \tag{III.6}$$

where $\hat\Sigma_{-k}=\hat\Sigma_k^\top$, converging in some anulus containig the unit circle. Since the covariance sequence of $\mathbf{y}_n$ coincides with the sequence of matrix coefficients in the expansion (III.6), the infinite Toeplitz matrix built from the matrix sequence $\{\hat\Sigma_k\}$ is just the infinite block-Toeplitz matrix $T(\Phi_n)$ with symbol Φ_n.

It is trivial to check that the inverse of the spectrum (III.5) is a symmetric matrix dipolynomial

$$\Phi_n(z)^{-1}=A_n(z)^\top\Delta_n^{-1}A_n(z^{-1}):=\sum_{k=-n}^{n}\Lambda_n(k)\,z^{-k}\,,$$

where $\Lambda_n(-k)=\Lambda_n(k)^\top$, which is everywhere positive definite on the unit circle and whose Laurent expansion consists of a *finite number of nonzero terms* say, $\{\Lambda_n(k)\,;\,|k|\leq n\}$. Hence $\Phi_n(z)^{-1}$ is a *bona-fide* spectral density which has only a *finite string of nonzero Fourier coefficients*. A stationary process having spectral density equal to the inverse $\Phi_n(z)^{-1}$ has only a finite sequence of nonzero covariances.

Such a process is commonly called a *moving average (M.A.)* process. Hence, $T(\Phi_n^{-1})$ (is well-defined and) is a *banded symmetric block-Toeplitz matrix of bandwidth* n.

As we have seen, the autoregressive covariance extension defined by formula (II.19) is well-known. It is often called the *maximum entropy extension* [6] of (I.1). We may therefore be led to the conclusion that the band extension problems is just the same as maximum entropy extension. This is actually true only to a point. In fact, our Toeplitz band extension problem 1 would be immediately solved in terms of maximum entropy extension provided a *spectral mapping* property of the following kind

$$T(\Psi)^{-1} \stackrel{?}{=} T(\Psi^{-1}) \tag{III.7}$$

would hold (of course assuming that the inverse of the Toeplitz matrix is well defined). For, equation (III.7) just means that the Fourier coefficients of $\Psi^{-1}(z)$ are stationary covariances forming the entries of the inverse Toeplitz matrix $T(\Psi)^{-1}$.

Unfortunately the spectral mapping property (III.7) **does not hold**. It does only in rather trivial cases, e.g. for analytic symbols in H^∞ (see [5]), but these cases do not include symbols which are nontrivial spectral densities. Note that (III.7) being true, would in particular imply that the inverse of a stationary covariance matrix $T(\Psi)$ should itself be a stationary (i.e. block-Toeplitz) covariance, which is well-known to be generally false. There are "transient" phenomena which should be accounted for in the inversion process, which induce non-stationarity. In fact, the inverse of (II.19) is non-Toeplitz. It is only asymptotically so.

IV. Band extension of block-circulant matrices

In the rest of this paper we shall consider stationary processes defined on a finite interval $[1,\, N]$ of the integer line $\mathbb{Z}$. Such a process is just a finite collection of say m-dimensional random vectors whose covariance matrix has a block-Toeplitz structure and will be denoted just as a mN-dimensional column vector. It can be shown that a stationary process $\mathbf{y}$ on the interval $[\,1,\, N\,]$ is the restriction to $[\,1,\, N\,]$ of a stationary process on $\mathbb{Z}$ which is periodic of period N, if and only if its covariance matrix is a symmetric block-circulant matrix. Processes of this kind can naturally be seen as being defined on the group $\mathbb{Z}_N$ of the integers modulo N and in this setting stationarity can be seen as propagation in time of random variables under the action of a (discrete) unitary group. Then most of the harmonic analysis of stationary processes on $\mathbb{Z}$ carries over naturally provided the Fourier transform is understood as a mapping from functions on $\mathbb{Z}_N$ onto complex-valued functions on the discrete unit circle $\mathbb{C}_N$ in the complex plane. We shall only consider stationary process of *full rank* whose covariance matrix is positive definite.

Definition 1: An m-dimensional stationary process $\mathbf{y}$ *on the discrete group $\mathbb{Z}_N$ is said to be* reciprocal of order n *if the random variables of the process in the interval $(t_1,\, t_2\,)$ are conditionally orthogonal to the random variables in the exterior, $(t_1,\, t_2)^c$, given the $2n$ boundary values $\mathbf{y}_{(t_1-n,\, t_1\,]}$ and $\mathbf{y}_{[t_2,\, t_2+n\,)}$, that is*

$$\begin{aligned}&\hat{\mathbb{E}}\,[\,\mathbf{y}_{(t_1,\, t_2)} \mid \mathbf{y}(s),\, s \in (t_1,\, t_2)^c\,] = \\ &= \hat{\mathbb{E}}\,[\,\mathbf{y}_{(t_1,\, t_2)} \mid \mathbf{y}_{(t_1-n,\, t_1\,]} \vee \mathbf{y}_{[t_2,\, t_2+n\,)}\,]\,, \quad t_1,\, t_2 \in [1,\, N]\,.\end{aligned} \tag{IV.1}$$

These processes are discussed in [1] and we shall refer the reader to this paper for details. The following is a fundamental characterization of (full-rank) reciprocal processes of order n.

Theorem 4.1 ([1]): A nonsingular $mN \times mN$-dimensional matrix $\mathbf{\Sigma}_N$ is the covariance matrix of a reciprocal process of order n on the discrete group $\mathbb{Z}_N$ if and only if its inverse is a positive-definite symmetric block-circulant matrix which is banded of bandwidth n.

An m-dimensional *white noise process* on $\mathbb{Z}_N$ is a process with covariance $\mathbf{I}_N$ (the $mN \times mN$ identity matrix partitioned in N diagonal blocks). It is clearly reciprocal of any order.

The circulant band extension problem on $\mathbb{Z}_N$ is motivated by stochastic realization and identification of reciprocal processes. Given the initial data (I.1), we want to find a block-circulant symmetric positive extension of T_n, say $\mathbf{\Sigma}_N$, which has a banded inverse of bandwidth n. We know from [1] that if N is large enough, such a block-circulant band extension of T_n of bandwidth (of the inverse) exactly equal to n exists and is in fact unique. We want to see if we can compute $\mathbf{\Sigma}_N$, or better $\mathbf{\Sigma}_N^{-1}$, by a factorization procedure based on the Yule-Walker/Schur-Levinson idea exposed in the previous sections. As we have seen before, computing the Toeplitz band extension of T_n is essentially the same thing as computing the AR model of order n which extends the process to the whole line $\mathbb{Z}$ [3]. In the next section we show that, similarly to the Toepltz case, block-circulant covariances which admit a banded inverse are precisely covariances of AR type reciprocal processes (which will be defined in the next section).

V. Factorization and unilateral AR representations

A *lower banded block-circulant* matrix of bandwidth n, has the representation $\mathbf{L}_N = \mathrm{Circ}\{L_0, L_1\,, \ldots, L_n\,0\,\ldots, 0\}$ where the $n+1$-st block in the list is nonzero; i.e. has the following structure

$$\begin{bmatrix} L_0 & 0 & \ldots & 0 & L_n & L_{n-1} & \ldots & L_1 \\ L_1 & L_0 & 0 & \ldots & 0 & L_n & \ldots & L_2 \\ \vdots & & \ddots & & & \ddots & & \vdots \\ L_{n-1} & L_{n-2} & \ldots & L_0 & 0 & \ldots & 0 & L_n \\ L_n & \ldots & \ldots & L_1 & L_0 & 0 & \ldots & 0 \\ 0 & L_n & \ldots & \ldots & L_1 & L_0 & 0 & \ldots \\ \vdots & & \ddots & & & \ddots & \ddots & \\ 0 & \ldots & 0 & L_n & \ldots & \ldots & L_1 & L_0 \end{bmatrix} \tag{V.1}$$

[3] "Extension of the process" means that the covariance of this AR process is a band extension of the partial covariance T_n.

Note that lower banded block-circulant matrices of bandwidth $N-1$; i.e. with $L_{N-1} \neq 0$, are just full block-circulant matrices (with the first upper diagonal different from zero). In this case the concept of lower banded block-circulant degenerates. $\mathbf{L}_N$ is said to be *normalized* whenever $L_0 = I$.

Lemma 5.1: The inverse of a lower banded block-circulant matrix $\mathbf{L}_N$ can itself be lower banded only when $\mathbf{L}_N$ is block-diagonal.

Proof: The symbol of $\mathbf{L}_N$ is a polynomial of degree $n < N/2$ in the discrete Fourier variable ζ (or ζ^{-1}). Since the symbol of the inverse is the inverse of the symbol, it is clear that it will in general be a polynomial of full degree N only if the symbol was a constant. ■

Hence, while the inverse of a (block-) lower triangular matrix is still (block-) lower triangular, the inverse of a lower banded block-circulant matrix is in general neither lower nor upper banded.

Theorem 5.1: Let $\mathbf{\Sigma}_N$ be a block-circulant covariance matrix with N blocks. If there exist a lower banded block-circulant matrix $\mathbf{L}_N$ of bandwidth $n < N/2$ which block-diagonalizes $\mathbf{\Sigma}_N$; i.e.

$$\mathbf{L}_N \mathbf{\Sigma}_N \mathbf{L}_N^\top = \operatorname{diag}\{D, \dots, D\} := \mathbf{D}_N \tag{V.2}$$

with $D = D^\top$ positive definite $m \times m$, then the inverse $\mathbf{\Sigma}_N^{-1}$ is a symmetric banded block-circulant matrix of bandwidth n and $\mathbf{\Sigma}_N$ is the covariance matrix of an order n reciprocal process on $\mathbb{Z}_N$.

Conversely, the covariance matrix of an order n full rank reciprocal process on $\mathbb{Z}_N$ admits lower banded block-circulant matrices $\mathbf{L}_N$ of bandwidth $n < N/2$ for which a relation of the type (V.2) *holds.*

Proof: Inverting the relation (V.2), we get the factorization

$$\mathbf{\Sigma}_N^{-1} = \mathbf{L}_N^\top \mathbf{D}_N^{-1} \mathbf{L}_N \tag{V.3}$$

where, given that $\mathbf{L}_N^\top$ is banded upper triangular, $\mathbf{D}_N^{-1}$ is block-diagonal and $\mathbf{L}_N$ is banded lower triangular, the product on the right hand side is symmetric, banded of (bilateral) bandwidth n. Since $n < N/2$, in $\mathbf{L}_N^\top \mathbf{D}_N^{-1} \mathbf{L}_N$ there is at least one upper (and lower) block-diagonal which is zero and the matrix $\mathbf{\Sigma}_N^{-1}$ is effectively banded. The conclusion follows from Theorem 4.1.

Conversely, for a reciprocal process of order n, the matrix $\mathbf{M}_N = \mathbf{\Sigma}_N^{-1}$ is (bilaterally) banded of bandwith $n < N/2$. Therefore, by the spectral factorization result in [2], it admits a spectral factor which can be chosen upper triangular of bandwith $n < N/2$. This factor would produce a factorization of $\mathbf{\Sigma}_N^{-1}$ of the type (V.3). Inverting this relation one obtains (V.2). ■

We shall call a lower banded, respectively upper banded, block-circulant matrix for which (V.2) holds a lower banded (upper banded) *left divisor* of $\mathbf{\Sigma}_N$.

In probabilistic terms, any block-circulant left divisor $\mathbf{L}_N$ of $\mathbf{\Sigma}_N$, not necessarily lower or upper banded, defines a *whitening filter*, since the process $\mathbf{w}$ defined by $\mathbf{w} := \mathbf{L}_N^{-1}\, \mathbf{y}$, must, because of the relation (V.2), have a block-diagonal covariance matrix, namely

$$\mathbb{E}\{\mathbf{w}\mathbf{w}^\top\} = \mathbf{D}_N\,.$$

In particular, we shall say that a lower banded $\mathbf{L}_N$ of bandwidth n, defines a *(causal) whitening filter of memory* n which can be written

$$\mathbf{w}(t) = \sum_{k=0}^{n} L_k \mathbf{y}(t-k)\,, \qquad t \in \mathbb{Z}_N\,. \tag{V.4}$$

Now, assuming $\mathbf{w}$ is a given white noise process, we can look at this equation as a recursive model *for the process* $\mathbf{y}$. A process $\mathbf{y}$ described by a model of the type (V.4) will be called *Auto-Regressive* and the model (V.4) of the *(causal) AR-type* on $\mathbb{Z}_N$. Note that the model must be associated to *cyclic initial conditions* $\mathbf{y}(0) = \mathbf{y}(N)\,, \dots, \mathbf{y}(-n+1) = \mathbf{y}(N-n+1)$. The qualification *causal* is added to distinguish it from *anticausal* models involving instead "future lags" $\mathbf{y}(t+k)$.

Lemma 5.2: Causal (resp. anticausal) Auto-Regressive models associated with initial (or terminal) cyclic boundary conditions generate periodic processes.

Proof: In fact the model (V.4) associated with the cyclic boundary conditions $\mathbf{y}(0) = \mathbf{y}(N)\,, \dots, \mathbf{y}(-n+1) = \mathbf{y}(N-n+1)$ can be written in matrix form as $\mathbf{L}_N \mathbf{y} = \mathbf{w}$ where $\mathbf{L}_N = \operatorname{Circ}\{L_0, L_1, \dots, L_n, 0, \dots, 0\}$ is an invertible block-circulant matrix. Then

$$\mathbf{\Sigma}_N = \mathbb{E}\,\mathbf{y}\mathbf{y}^\top = \mathbf{L}_N^{-1} \mathbf{I}_N \mathbf{L}_N^{-\top}$$

which is symmetric block-circulant. Therefore by Proposition 2.1 in [1], $\mathbf{y}$ is a periodic process in $\mathbb{Z}_N$. ■

The following theorem establishes the equivalence of reciprocal processes of order n with periodic autoregressive processes on $\mathbb{Z}_N$. Due to space limitations the proof will be skipped.

Theorem 5.2: Let $n < N/2$. Every full rank reciprocal process of order n can be described by a n-th order AR-type model (either causal or anticausal) with cyclic initial (or terminal) conditions. Conversely, every stationary full-rank process on $\mathbb{Z}_N$, described by an AR-type model of order n, is a reciprocal process of order n. Hence the class of full-rank AR-type processes of order n on $\mathbb{Z}_N$ coincides with the class of reciprocal processes of order n.

This equivalence has been very briefly addressed in [8, Sect. V] for general (non stationary) reciprocal processes of order one and can be seen as a limit case of the result of the paper [7], valid for stationary reciprocal processes on a doubly infinite interval. We have generalized these representation results to reciprocal processes of arbitrary order n.

It seems clear that a generic block-circulant covariance matrix $\mathbf{\Sigma}_N$ will in general not admit lower (or upper) banded divisors of bandwidth less than $N-1$. Equivalently, a generic stationary process $\mathbf{y}$ on $\mathbb{Z}_N$ cannot in general be whitened by a filter of memory $n < N/2$ (equivalently cannot be represented as an AR-process of order n). However the covariance of a full-rank reciprocal process of order n does (Theorem 5.1) and the question is to understand if the lower banded divisors can be computed by some analog of the Levinson polynomials recursion.

VI. Circulant Schur-Levinson Polynomials

Given a stationary periodic process $\mathbf{y}$ on $\mathbb{Z}_N$, inspired by the procedure of Section II, we introduce the *forward and backward innovation processes of finite memory* n of $\mathbf{y}$, defined as

$$\mathbf{e}_n(t) := \mathbf{y}(t) - \hat{\mathbb{E}}\left[\mathbf{y}(t) \mid \mathbf{y}_{[t-n,t)}\right], \qquad \text{(VI.1)}$$

$$\bar{\mathbf{e}}_n(t) := \mathbf{y}(t) - \hat{\mathbb{E}}\left[\mathbf{y}(t) \mid \mathbf{y}_{(t,t+n]}\right], \qquad \text{(VI.2)}$$

where the time interval is $\mathbb{Z}_N$ with modular arithmetic mod N.

For $n < t$, (VI.1) is just the same forward innovation of memory n introduced in section II. Note however that, when $n \geq t$, the subinterval of $[t-n, t)$ with nonpositive times gets "folded around" by the arithmetics mod N, so that we have

$$\begin{aligned}
\mathbf{e}_n(1) &:= \mathbf{y}(1) - \hat{\mathbb{E}}\left[\mathbf{y}(1) \mid \mathbf{y}_{[1-n,0]}\right] \\
&= \mathbf{y}(1) - \hat{\mathbb{E}}\left[\mathbf{y}(1) \mid \mathbf{y}(0), \mathbf{y}(-1), \ldots, \mathbf{y}(-n+1)\right] \\
&= \mathbf{y}(1) - \hat{\mathbb{E}}\left[\mathbf{y}(1) \mid \mathbf{y}(N), \mathbf{y}(N-1), \ldots, \mathbf{y}(N-n+1)\right] \\
\mathbf{e}_n(2) &:= \mathbf{y}(2) - \hat{\mathbb{E}}\left[\mathbf{y}(2) \mid \mathbf{y}_{[2-n,1]}\right] \\
&= \mathbf{y}(2) - \hat{\mathbb{E}}\left[\mathbf{y}(2) \mid \mathbf{y}(1), \mathbf{y}(0), \ldots, \mathbf{y}(-n+2)\right] \\
&= \mathbf{y}(2) - \hat{\mathbb{E}}\left[\mathbf{y}(2) \mid \mathbf{y}(1), \mathbf{y}(N), \ldots, \mathbf{y}(N-n+2)\right] \\
&\ldots\ldots \\
\mathbf{e}_n(n) &:= \mathbf{y}(n) - \hat{\mathbb{E}}\left[\mathbf{y}(n) \mid \mathbf{y}_{[0,n-1]}\right] \\
&= \mathbf{y}(n) - \hat{\mathbb{E}}\left[\mathbf{y}(n) \mid \mathbf{y}(0), \mathbf{y}(1), \ldots, \mathbf{y}(n-1)\right] \\
&= \mathbf{y}(n) - \hat{\mathbb{E}}\left[\mathbf{y}(n) \mid \mathbf{y}(N), \mathbf{y}(1), \ldots, \mathbf{y}(n-1)\right]
\end{aligned}$$

and so on, from $t = 1, 2, \ldots, n$ to all $t \in \mathbb{Z}_N$. The expressions above can be rewritten in matrix form.

$$\begin{aligned}
\mathbf{e}_n(1) &:= \begin{bmatrix} I & 0 & \ldots & 0 & A(n) & A(n-1) & \ldots & A(1) \end{bmatrix} \mathbf{y} \\
\mathbf{e}_n(2) &:= \begin{bmatrix} A(1) & I & 0 & \ldots & 0 & A(n) & A(n-1) & \ldots & A(2) \end{bmatrix} \mathbf{y} \\
&\ldots\ldots \\
\mathbf{e}_n(n) &:= \begin{bmatrix} A(n-1) & A(n-2) & \ldots & I & 0 & \ldots & 0 & A(n) \end{bmatrix} \mathbf{y}
\end{aligned} \qquad \text{(VI.3)}$$

$$\ldots\ldots\ .$$

and hence as a circulant convolution filter

$$\mathbf{e}_n(t) = \sum_{k=0}^{n} A(k)\, \mathbf{y}(t-k), \qquad t \in \mathbb{Z}_N \qquad \text{(VI.4)}$$

That the coefficients $\{A(k)\}$ in these expressions are actually independent of t, is proven in the next proposition.

Proposition 6.1: The coefficients $A(k)$ do not depend on t and are uniquely determined by the initial data (I.1). *They coincide with the coefficients $A_n(k)$ of the n-th Schur-Levinson polynomial $A_n(z^{-1})$ introduced in Section II computed from the same initial covariance data* (I.1). *In other words, $\hat{A}(\zeta^{-1}) := \sum_{k=0}^{n} A(k)\zeta^{-k}\,; \zeta \in \mathbb{C}_N$ is obtained by frequency sampling*

$$\hat{A}(\zeta^{-1}) = \left[A_n(z^{-1})\right]_{|z=\zeta} \qquad \text{(VI.5)}$$

where $\zeta \equiv e^{i\Delta}$.

Proof: Time invariance follows from the joint stationarity of $\mathbf{e}_n(t)$ and $\{\mathbf{y}(t), \mathbf{y}(t-1), \mathbf{y}(t-2) \ldots, \mathbf{y}(t-n)\}$ for all $t \in \mathbb{Z}_N$ which implies that all cross covariances of these random elements do not depend on t. The coefficients $A(k)$ are determined by the orthogonality condition $\mathbf{e}_n(t) \perp \{\mathbf{y}(t-1), \mathbf{y}(t-2) \ldots, \mathbf{y}(t-n)\}$ which must hold for all $t \in \mathbb{Z}_N$. For $t = 1$ for example, rearranging the first equation in (VI.3), we obtain

$$\begin{bmatrix} I & A(1) & A(n-1) & \ldots & A(n) \end{bmatrix} \times \mathbb{E} \begin{bmatrix} \mathbf{y}(1) \\ \mathbf{y}(N) \\ \mathbf{y}(N-1) \\ \vdots \\ \mathbf{y}(N-n+1) \end{bmatrix} \begin{bmatrix} \mathbf{y}(N) \\ \mathbf{y}(N-1) \\ \vdots \\ \mathbf{y}(N-n+1) \end{bmatrix}^{\top} = 0 \qquad \text{(VI.6)}$$

which yields

$$\begin{bmatrix} I & A(1) & A(n-1) & \ldots & A(n) \end{bmatrix} \begin{bmatrix} \Sigma_{N-1}^{\top} & \Sigma_{N-2}^{\top} & \ldots & \Sigma_{N-n}^{\top} \\ \Sigma_0 & \Sigma_1 & \ldots & \Sigma_{n-1} \\ \Sigma_1^{\top} & \Sigma_0 & \Sigma_1 & \ldots \\ \ldots & & & \ldots \\ \Sigma_{n-1}^{\top} & \ldots & \Sigma_1^{\top} & \Sigma_0 \end{bmatrix} = 0$$

Substituting the known boundary values $\Sigma_{N-k}^{\top} = \Sigma_k\,;\ k = 1, 2, \ldots, n$ imposed by circulant symmetry, one obtains the "circulant" Yule-Walker equation

$$\begin{bmatrix} I & A(1) & A(n-1) & \ldots & A(n) \end{bmatrix} \begin{bmatrix} \Sigma_1 & \Sigma_2 & \ldots & \Sigma_n \\ \Sigma_0 & \Sigma_1 & \ldots & \Sigma_{n-1} \\ \Sigma_1^{\top} & \Sigma_0 & \Sigma_1 & \ldots \\ \ldots & & & \ldots \\ \Sigma_{n-1}^{\top} & \ldots & \Sigma_1^{\top} & \Sigma_0 \end{bmatrix} = 0$$

which can easily be seen to be exactly the same as equations (II.16) for $k = 1, 2, \ldots, n$, determining the n-th Schur-Levinson polynomial $A_n(z^{-1})$. Note that the $A(k)$'s depend only on the initial data (I.1) of the problem. ■

Naturally, introducing the lower banded block-circulant matrix $\mathbf{A}_N$, defined as

$$\mathbf{A}_N := \mathrm{Circ}\{I,\, A(1),\, A(2),\, \ldots, A(n),\, 0 \ldots, 0\} \qquad \text{(VI.7)}$$

the N ordered samples of the memory n circulant innovation process $\mathbf{e}_n$, can be expressed as

$$\begin{bmatrix} \mathbf{e}_n(1) \\ \mathbf{e}_n(2) \\ \vdots \\ \mathbf{e}_n(n) \\ \vdots \\ \mathbf{e}_n(N) \end{bmatrix} = \mathbf{A}_N \begin{bmatrix} \mathbf{y}(1) \\ \mathbf{y}(2) \\ \vdots \\ \mathbf{y}(n) \\ \vdots \\ \mathbf{y}(N) \end{bmatrix} \qquad \text{(VI.8)}$$

Of course a symmetric reasoning leads to similar representations for the backward innovations (VI.2). However we shall not need to report them here.

As we have seen, even for a general stationary process $\mathbf{y}$ on $\mathbb{Z}$, the memory n innovation will in general not be white, unless the process has an autoregressive structure of order n. In this case the memory n innovation of $\mathbf{y}$ coincides with the actual innovation process (Proposition 3.2). Likewise, for a general stationary process on $\mathbb{Z}_N$, the circulant innovation of memory n will in general not be white. However, in analogy to the construction of Section II, one may hope that this would be true for an AR process of order n; i.e. for a reciprocal process of order n. Unfortunately the proposition below shows that this can be the case only in trivial situations.

Proposition 6.2: Let $\mathbf{y}$ *be a reciprocal process of order* n *on* $\mathbb{Z}_N$ *with initial covariances* (I.1). *Then its memory* n *innovation* $\mathbf{e}_n$ *is in general not a white noise process.*

Proof: By Theorem 5.2 the process $\mathbf{y}$ can be described by a causal AR model of order n, say

$$\mathbf{y}(t) + \sum_{k=1}^{n} L_k \mathbf{y}(t-k) = \mathbf{w}(t)\,, \qquad t \in \mathbb{Z}_N$$

where $\mathbf{w}$ is white noise. This model can be written in vector notation as

$$\mathbf{L}_N \mathbf{y} = \mathbf{w} \tag{VI.9}$$

where $\mathbf{L}_N$ is a lowerbanded block-circulant matrix with entries the coefficients L_k.

If and only if $\mathbf{w}(t) \perp \{\mathbf{y}(t-1), \mathbf{y}(t-2) \ldots, \mathbf{y}(t-n)\}$, then $\mathbf{w}(t)$ would coincide with the memory n predction error; i.e. $\mathbf{w}(t) \equiv \mathbf{e}_n(t)$ and therefore $\mathbf{e}_n(t)$ would be a white process. However since the inverse of the lower banded matrix $\mathbf{L}_N$ is in general not lower banded, the solution $\mathbf{y}$ of the AR equation will in general not be a causal function of $\mathbf{w}$, that is in general each random variable in the set $\{\mathbf{y}(t-1), \mathbf{y}(t-2) \ldots, \mathbf{y}(t-n)\}$ will *depend on the future values of the noise* $\mathbf{w}$, so that $\mathbf{w}(t)$ cannot be uncorrelated with $\{\mathbf{y}(t-1), \mathbf{y}(t-2) \ldots, \mathbf{y}(t-n)\}$. ■

We conjecture that $\mathbf{e}_n$ can be white noise only when $\mathbf{y}$ is itself white. If this is true, then the Levinson matrix $\mathbf{A}_N$ can be a lower banded divisor of its covariance matrix only in the trivial case when $\mathbf{y}$ itself is white.

VII. Conclusions

In this paper we have drawn a parallel between block-Toeplitz and block-circulant band extension problems. Although the two settings present some close similarities, we have shown that the factorization idea on which the solution of the Toeplitz band extension problem is based, does not work in the circulant case.

References

[1] F. Carli, A. Ferrante, M. Pavon, and G. Picci. A maximum entropy solution of the covariance extension problem for reciprocal processes. *IEEE Transactions Automatic Control*, To appear:..., 2010.

[2] F. Carli and G. Picci. Covariance extension of reciprocal processes. Technical report, DEI University of Padova, 2010.

[3] H. Dym and I. Gohberg. Extension of band matrices with band inverses. *Linear Algebra and Applications*, 36:1–24, 1981.

[4] I. Gohberg, Goldberg, and M. Kaashoek. *Classes of Linear Operators vol II*. Birkhauser, Boston, 1994.

[5] P. R. Halmos. *A Hilbert space problem book*, volume 19 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1982. Encyclopedia of Mathematics and its Applications, 17.

[6] E. T. Jaynes. On the rationale of maximum-entropy methods. *Proc. IEEE*, 70:939–952, 1982.

[7] B. C. Levy. Regular and reciprocal multivariate stationary Gaussian reciprocal processes over $\mathbf{Z}$ are necessarily Markov. *J. Math. Systems, Estimation and Control*, 2:133–154, 1992.

[8] B. C. Levy, R. Frezza, and A.J. Krener. Modeling and estimation of discrete-time Gaussian reciprocal processes. *IEEE Trans. Automatic Control*, AC-35(9):1013–1023, 1990.

[9] W. F. Trench. An algorithm for the inversion of finite toeplitz matrices. *SIAM J. Appl. Math*, 12:515–522, 1964.

[10] P. Whittle. On the fitting of multivariate autoregressions and the approximate canonical factorization of a spectral density matrix. *Biometrica*, 50:129–134, 1963.