An Efficient Algorithm for Solving General Periodic Toeplitz Systems

Mrityunjoy Chakraborty

***Abstract*— An efficient algorithm is presented for inverting matrices which are periodically Toeplitz, i.e., whose diagonal and subdiagonal entries exhibit periodic repetitions. Such matrices are not per symmetric and thus cannot be inverted by Trench's method. An alternative approach based on appropriate matrix factorization and partitioning is suggested. The algorithm provides certain insight on the formation of the inverse matrix, is implementable on a set of circularly pipelined processors and, as a special case, can be used for inverting a set of block Toeplitz matrices without requiring any matrix operation.**

I. Introduction

A linear system of equations of the form $\boldsymbol{Rw} = \boldsymbol{p}$ is said to be a periodic Toeplitz system with period d, if the matrix $\boldsymbol{R}$ satisfies the periodicity condition: $\boldsymbol{R}_{i,j} = \boldsymbol{R}_{i\pm d,\, j\pm d}$. Such matrices typically represent the correlation structure of cyclostationary processes and have been studied in [1] and [2] in the context of periodic time series analysis, where the periodicity of the matrix elements along each diagonal and subdiagonal has been exploited to obtain efficient parameter estimation algorithms. In addition, such matrices also play a vital role in efficient realization of multichannel modeling and filtering algorithms, as shown in [1] and [2] and indirectly in [3] and [4]. An example of a set of such matrices with period, say, 3, can be constructed as

$$\boldsymbol{A} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 1 \end{bmatrix}, \quad \boldsymbol{B} = \begin{bmatrix} 6 & 7 & 8 & 18 \\ 10 & 11 & 12 & 17 \\ 14 & 15 & 1 & 2 \\ 19 & 16 & 5 & 6 \end{bmatrix}$$

$$\boldsymbol{C} = \begin{bmatrix} 11 & 12 & 17 & 20 \\ 15 & 1 & 2 & 3 \\ 16 & 5 & 6 & 7 \\ 21 & 9 & 10 & 11 \end{bmatrix}$$

where each matrix shown is periodic Toeplitz although, in each case, only the (1, 1)th element is seen to repeat itself since the period chosen is 3, whereas the dimension is 4×4. The matrices $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$ are, in fact, chosen carefully to represent the cross-correlation matrices of a cyclostationary process with period $= 3$, measured at three successive indices of time. Such matrices overlap in a certain manner, as explained later in Section II. The algorithms presented in [1]–[4], however, apply only to certain specific sets of equations involving such matrices, namely, Yule–Walker (YW) equations [1] and modified Yule–Walker (MYW) equations [2]. This correspondence takes up the more general problem of solving arbitrary systems of the form $\boldsymbol{Rw} = \boldsymbol{p}$, where $\boldsymbol{R}$ is non-Hermitian periodic Toeplitz. An efficient algorithm to compute $\boldsymbol{R}^{-1}$ has been presented, which, apart from providing insight on the structure of $\boldsymbol{R}^{-1}$, also has an interesting pipelined implementation. Since a $p \times p$ periodic Toeplitz matrix acquires the so-called block Toeplitz structure when $p = Nd$ for an integer N, the proposed algorithm can also be used for efficient inversion of block Toeplitz matrices with the additional advantage of not requiring any matrix computation.

Manuscript received April 10, 1996; revised June 10, 1997. The associate editor coordinating the review of this paper and approving it for publication was Prof. Pierre Comon.

The author is with the Department of Electronics and Electrical Communications Engineering, Indian Institute of Technology, Kharagpur, India.

Publisher Item Identifier S 1053-587X(98)01334-8.

II. Review of Periodic Toeplitz Systems

Consider a process $y(n)$, which is given to be periodically WSS with period d, implying that $r(n,m) = r(n \pm d, m \pm d)$, where $r(n,m)$ denotes the autocorrelation function of $y(n)$, i.e., $r(n,m) = E[y(n)y(m)]$. The general $(p+1)\times(p+1)$ correlation matrix $\boldsymbol{R}_n(p)$ for the nth index is then given by $\boldsymbol{R}_n(p) = E[\boldsymbol{y}_{n-q}(p)\boldsymbol{y}_n^t(p)]$, where $\boldsymbol{y}_n(p) = [y(n), y(n-1), \cdots, y(n-p)]^t, p, q = 0, 1, 2, \cdots$.It is easy to verify that

$$\boldsymbol{R}_n(p) = \left[\begin{array}{c|c} \boldsymbol{R}_n(p-1) & \\ \hline & \end{array}\right] = \left[\begin{array}{c|c} & \\ \hline & \boldsymbol{R}_{n-1}(p-1) \end{array}\right] \tag{1a}$$

and

$$\boldsymbol{R}_n(p) = \boldsymbol{R}_{n\pm d}(p). \tag{1b}$$

Thus, for each order p, there is a total of d matrices: one each for the indices $n, n-1, \cdots, n-d+1$. The objective is to compute $\boldsymbol{R}_n^{-1}(p)$'s for each of these indices in an efficient manner. It will be useful in this context to consider the MYW equations for $\boldsymbol{R}_n(p)$, solved in [2] and given, for the nth index and the pth order, by

$$\boldsymbol{R}_n(p)[1, \boldsymbol{a}_{n,p}^t]^t = [\alpha_{n,p}, \boldsymbol{0}_{1\times p}]^t \tag{2a}$$

$$\boldsymbol{R}_n(p)[\boldsymbol{b}_{n,p}^t, 1]^t = [\boldsymbol{0}_{1\times p}, \beta_{n,p}]^t \tag{2b}$$

where $\boldsymbol{a}_{n,p} = [a_{n,p}(1), a_{n,p}(2), \cdots, a_{n,p}(p)]^t$, and $\boldsymbol{b}_{n,p} = [b_{n,p}(p), b_{n,p}(p-1), \cdots, b_{n,p}(1)]^t$. The two unknowns $\alpha_{n,p}$ and $\beta_{n,p}$ are given by the inner products of the first and the last rows of $\boldsymbol{R}_n(p)$ with the vectors $[1, \boldsymbol{a}_{n,p}^t]^t$ and $[\boldsymbol{b}_{n,p}^t, 1]^t$, respectively. In [2], an order-recursive procedure for solving (2a) and (2b) has been presented that exploits the periodicity along the diagonals of $\boldsymbol{R}_n(p)$ to obtain a pipelined realization, involving d processors and $O(p^2d^2)$ computations. Interestingly, $\boldsymbol{R}_n^t(p)$, like $\boldsymbol{R}_n(p)$, is also periodically Toeplitz with period d, and thus, the same algorithm can be used to solve the MYW equations for $\boldsymbol{R}_n^t(p)$ as well. We denote the corresponing solutions for (2a) and (2b), respectively, by $\tilde{\boldsymbol{a}}_{n,p}$ and $\tilde{\boldsymbol{b}}_{n,p}$ with $\tilde{\alpha}^{n,p}$ and $\tilde{\beta}_{n,p}$ substituting for $\alpha_{n,p}$ and $\beta_{n,p}$.

The inversion algorithm proposed in this paper is based on the so-called "$\boldsymbol{LDU}$" and "$\boldsymbol{UDL}$" factorizations of $\boldsymbol{R}_n(p)$, where the "$\boldsymbol{L}$" and the "$\boldsymbol{U}$" matrices are constructed using both $\boldsymbol{a}_{n,p}, \boldsymbol{b}_{n,p}$ and $\tilde{\boldsymbol{a}}_{n,p}, \tilde{\boldsymbol{b}}_{n,p}$. For the Hermitian case, however, only half of these solutions will be required as, in that case, $\boldsymbol{R}_n(p) \equiv \boldsymbol{R}_n^t(p)$, implying that $\boldsymbol{a}_{n,p} \equiv \tilde{\boldsymbol{a}}_{n,p}$ and $\boldsymbol{b}_{n,p} \equiv \tilde{\boldsymbol{b}}_{n,p}$.

III. The Inversion Algorithm

A. Derivation of the Inversion Formula

We begin by considering the $\boldsymbol{LDU}$ factorization of $\boldsymbol{R}_p(p)$. Construct

$$\tilde{\boldsymbol{L}}_n(p) = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ \tilde{b}_{n,1}(1) & 1 & 0 & \cdots & 0 \\ . & . & . & \cdots & . \\ . & . & . & \cdots & . \\ . & . & . & \cdots & . \\ \tilde{b}_{n,p}(p) & \tilde{b}_{n,p}(p-1) & . & \cdots & 1 \end{bmatrix}$$

$$\boldsymbol{U}_n(p) = \begin{bmatrix} 1 & b_{n,1}(1) & \cdots & b_{n,p}(p) \\ 0 & 1 & \cdots & b_{n,p}(p-1) \\ . & . & \cdots & . \\ . & . & \cdots & . \\ . & . & \cdots & . \\ 0 & 0 & \cdots & 1 \end{bmatrix}. \tag{3}$$

Then, we have $\boldsymbol{R}_n(p) = (\tilde{\boldsymbol{L}}_n(p))^{-1}\overline{\boldsymbol{D}_n}(p)\,(\boldsymbol{U}_n(p))^{-1}$, where $\overline{\boldsymbol{D}_n}(p)$ is a diagonal matrix with $(\overline{\boldsymbol{D}_n}(p))_{i,i} = \beta_{n,i-1}, i = 1, 2, \cdots, p+1$. This result can be proved in several ways. A proof based on Hilbert space treatment of random variables has been suggested in the Appendix. Equivalently, $\boldsymbol{R}_n^{-1}(p) = \boldsymbol{U}_n(p)\,(\overline{\boldsymbol{D}_n}(p))^{-1}\tilde{\boldsymbol{L}}_n(p)$. Partitioning the matrices $\boldsymbol{U}_n(p)$, $\tilde{\boldsymbol{L}}_n(p)$, and $\overline{\boldsymbol{D}_n}(p)$ blockwise, one can write

$$\boldsymbol{R}_n^{-1}(p) = \left[\begin{array}{c|c} \boldsymbol{U}_n(p-1) & \boldsymbol{b}_{n,p} \\ \hline \boldsymbol{0}_{1\times p} & 1 \end{array}\right]\left[\begin{array}{c|c} (\boldsymbol{D}_n(p-1))^{-1} & \boldsymbol{0}_{p\times 1} \\ \hline \boldsymbol{0}_{1\times p} & \beta_{n,p}^{-1} \end{array}\right] \cdot \left[\begin{array}{c|c} \tilde{\boldsymbol{L}}_n(p-1) & \boldsymbol{0}_{p\times 1} \\ \hline \tilde{\beta}_{n,p}^{t} & 1 \end{array}\right]. \tag{4}$$

Carrying out the matrix products, we then obtain

$$\boldsymbol{R}_n^{-1}(p) = \left[\begin{array}{c|c} \boldsymbol{R}_n^{-1}(p-1) + \beta_{n,p}^{-1}\boldsymbol{b}_{n,p}\tilde{\boldsymbol{b}}_{n,p}^{t} & \beta_{n,p}^{-1}\boldsymbol{b}_{n,p} \\ \hline \beta_{n,p}^{-1}\tilde{\boldsymbol{b}}_{n,p}^{t} & \beta_{n,p}^{-1} \end{array}\right]. \tag{5}$$

It may be recalled at this point that for a WSS $y(n)$, $\boldsymbol{R}_n(p)$ is purely Toeplitz and is inverted by the well-known Trench's algorithm [5], which makes use of the persymmetry[1] of $\boldsymbol{R}_n(p)$ or, equivalently, of $\boldsymbol{R}_n^{-1}(p)$, and obtains an alternative expression for $\boldsymbol{R}_n^{-1}(p)$ by premultiplying and postmultiplying the left- and right-hand sides of (5) by the exchange matrix $\boldsymbol{J}$. Such an approach, whose final goal is to obtain a recurrence relation between the (i,j)th and the $(i-1, j-1)$th elements of $\boldsymbol{R}_n^{-1}(p)$, cannot, however, be extended to periodic Toeplitz matrices, as the persymmetry condition is not satisfied by such matrices. As an alternative to this approach, we consider $\boldsymbol{UDL}$ the factorization of $\boldsymbol{R}_n(p)$ first. Defining

$$\boldsymbol{L}_n(p) = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ a_{n,p}(1) & 1 & 0 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ a_{n,p}(p) & a_{n-1,p-1}(p-1) & \cdot & \cdots & 1 \end{bmatrix}$$

$$\tilde{\boldsymbol{U}}_n(p) = \begin{bmatrix} 1 & \tilde{a}_{n\,p}(1) & \cdots & \tilde{a}_{n,p}(p) \\ 0 & 1 & \cdots & \tilde{a}_{n-1,p-1}(p-1) \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & \cdots & 1 \end{bmatrix} \tag{6}$$

we can write, as shown in the Appendix, $\boldsymbol{R}_n(p) = (\tilde{\boldsymbol{U}}_n(p))^{-1}\boldsymbol{D}_n(p)\,(\boldsymbol{L}_n(p))^{-1}$, where $\boldsymbol{D}_n(p)$ is a diagonal matrix with $(\boldsymbol{D}_n(p))_{i,i} = \alpha_{n-i+1,p-i+1}, i = 1, 2, \cdots, p+1$. Equivalently, $\boldsymbol{R}_n^{-1}(p) = \boldsymbol{L}_n(p)\boldsymbol{D}_n^{-1}(p)\tilde{\boldsymbol{U}}_n(p)$. Once again, partitioning $\tilde{\boldsymbol{U}}_n(p), \boldsymbol{L}_n(p)$ and $\boldsymbol{D}_n(p)$ blockwise, $\boldsymbol{R}_n^{-1}(p)$ can be written as

$$\boldsymbol{R}_n^{-1}(p) = \left[\begin{array}{c|c} 1 & \boldsymbol{0}_{1\times p} \\ \hline \boldsymbol{a}_{n,p} & \boldsymbol{L}_{n-1}(p-1) \end{array}\right]\left[\begin{array}{c|c} \alpha_{n,p}^{-1} & \boldsymbol{0}_{1\times p} \\ \hline \boldsymbol{0}_{p\times 1} & \boldsymbol{D}_{n-1}^{-1}(p-1) \end{array}\right] \cdot \left[\begin{array}{c|c} 1 & \tilde{a}_{n,p}^{t} \\ \hline \boldsymbol{0}_{p\times 1} & \tilde{\boldsymbol{U}}_{n-1}(p-1) \end{array}\right]. \tag{7}$$

Carrying out the matrix products

$$\boldsymbol{R}_n^{-1}(p) = \left[\begin{array}{c|c} \alpha_{n,p}^{-1} & \alpha_{n,p}^{-1}\tilde{\boldsymbol{a}}_{n,p}^{t} \\ \hline \alpha_{n,p}^{-1}\boldsymbol{a}_{n,p} & \boldsymbol{R}_{n-1}^{-1}(p-1) + \alpha_{n,p}^{-1}\boldsymbol{a}_{n,p}\tilde{\boldsymbol{a}}_{n,p}^{t} \end{array}\right]. \tag{8}$$

Changing n to $n+1$ in (8), we get $[\boldsymbol{R}_{n+1}^{-1}(p)]_{i+1,j+1} = [\boldsymbol{R}_n^{-1}(p-1)]_{i,j} + \alpha_{n+1,p}^{-1}a_{n+1,p}(i)\tilde{a}_{n+1,p}(j), p \geq i, j \geq 1$. On the other hand, from (5), we have $[\boldsymbol{R}_n^{-1}(p)]_{i,j} = [\boldsymbol{R}_n^{-1}(p-1)]_{i,j} + \beta_{n,p}^{-1}b_{n,p}(p-i+1)\tilde{b}_{n,p}(p-j+1), p \geq i, j \geq 1$. Eliminating $[\boldsymbol{R}_n^{-1}(p-1)]_{i,j}$, one then obtains

$$\begin{aligned} &[\boldsymbol{R}_{n+1}^{-1}(p)]_{i+1,j+1} \\ &\quad = [\boldsymbol{R}_n^{-1}(p)]_{i,j} + \alpha_{n+1,p}^{-1}a_{n+1,p}(i)\tilde{a}_{n+1,p}(j) \\ &\qquad - \beta_{n,p}^{-1}b_{n,p}(p-i+1)\tilde{b}_{n,p}(p-j+1) \\ &\qquad\quad p \geq i, j \geq 1. \end{aligned} \tag{9}$$

In addition, from (8), we have $[\boldsymbol{R}_{n+1}^{-1}(p)]_{1,1} = \alpha_{n+1,p}^{-1}$, $[\boldsymbol{R}_{n+1}^{-1}(p)]_{1,j+1} = \alpha_{n+1,p}^{-1}\tilde{a}_{n+1,p}(j)$, and $[\boldsymbol{R}_{n+1}^{-1}(p)]_{i+1,1} = \alpha_{n+1,p}^{-1}a_{n+1,p}(i)$, $p \geq i, j \geq 1$.

Equation (9) reveals the process by which the diagonals and subdiagonals of $\boldsymbol{R}_{n+1}^{-1}(p)$ are formed from the respective diagonals and subdiagonals of $\boldsymbol{R}_n^{-1}(p)$ and can be recognized as the periodic Toeplitz generalization of Trench's inversion formula [5]. Since the underlying process is periodically stationary, it is sufficient to consider $\boldsymbol{R}_n^{-1}(p)$ for $n = 0, 1, \cdots, d-1$.

A sequential algorithm to compute $\boldsymbol{R}_n^{-1}(p)$ based on (9) is listed in Table I. The algorithm constructs $\boldsymbol{R}_n^{-1}(p)$ by forming the rows iteratively. Alternatively, one can also iterate over the columns and thus construct $\boldsymbol{R}_n^{-1}(p)$.

TABLE I
A SEQUENTIAL ALGORITHM FOR COMPUTING $\boldsymbol{R}_n^{-1}(p)$, $n \in S = \{0, 1, \cdots, d-1\}$. THE SYMBOL '$\ominus$' DENOTES MODULO d SUBTRACTION

Initialization $(n = 0, 1, \ldots, d-1)$

$$[\boldsymbol{R}_n^{-1}(p)]_{1,1} = \alpha_{n,p}^{-1}, \quad [\boldsymbol{R}_n^{-1}(p)]_{1,j+1} = \alpha_{n,p}^{-1}\tilde{a}_{n,p}(j), \quad j = 1, 2, \ldots, p$$

$$[\boldsymbol{R}_n^{-1}(p)]_{j+1,1} = \alpha_{n,p}^{-1}a_{n,p}(j), \quad j = 1, 2, \ldots, p.$$

For $i = 1$ **to** p **begin**
 For $n = 0, 1, \ldots, d-1$ **begin**
  For $j = 1, 2, \ldots, p$ **begin**

$$[\boldsymbol{R}_n^{-1}(p)]_{i+1,j+1} = [\boldsymbol{R}_{n\ominus 1}^{-1}(p)]_{i,j} + \alpha_{n,p}^{-1}a_{n,p}(i)\tilde{a}_{n,p}(j) - \beta_{n\ominus 1,p}^{-1}b_{n\ominus 1,p}(p-i+1)\tilde{b}_{n\ominus 1,p}(p-j+1)$$

  end
 end
end

B. Parallel Implementation

A parallel implementation of the algorithm is possible by engaging a set of d processors and assigning the task of computing $\boldsymbol{R}_n^{-1}(p)$ to the nth processor $n \in S = \{0, 1, \cdots, d-1\}$. In Table I, this would imply that i) the "begin–end" statement for $n = 0, 1, \cdots, d-1$ be replaced with a "do par-end par" statement, and ii) in the ith iterative step, $i = 1, 2, \cdots, p$, $[\boldsymbol{R}_n^{-1}(p)]_{i,j}$, $j = 1, 2, \cdots, p$ be transmitted from the nth processor to the $(n+1)$th processor for $n = 0, 1, \cdots, d-2$ and from the $(d-1)$th processor to the zeroth processor for $n = d-1$. The communication amongst the processors is *circular* and thus local, making the procedure suitable for VLSI implementation.

The algorithm requires a total of $2p^2 + 4p$ multiplications and $2p^2$ additions per processor. If the columns (rows) are also evaluated parallelly, then a further p-fold increase in speedup is possible, leading to a so-called superfast algorithm, with multiplication and addition counts reducing to $2p+4$ and $2p$, respectively.

[1] A matrix $\boldsymbol{R}$ is persymmetric if $\boldsymbol{JRJ} = \boldsymbol{R}^t$, where $\boldsymbol{J}$ is the exchange matrix.

C. Closed-Form Expression

A closed-form expression for $\boldsymbol{R}_n^{-1}(p)$ can be worked out using (9). Defining for each n

a) $[\boldsymbol{R}_n^{-1}(p)]_{i,j} = 0$ whenever $i = 0$, $j = 0$, or $i = j = 0$;
b) $a_{n,p}(0) = \tilde{a}_{n,p}(0) = 1$;
c) $b_{n,p}(p+1) = \tilde{b}_{n,p}(p+1) = 0$;

we can write

$$[\boldsymbol{R}_{n+1}^{-1}(p)]_{i+1,j+1} = [\boldsymbol{R}_n^{-1}(p)]_{i,j} + T_{n+1}(i,j), \qquad 0 \le i,j \le p \tag{10}$$

where

$$T_{n+1}(i,j) = \alpha_{n+1,p}^{-1} a_{n+1,p}(i)\tilde{a}_{n+1,p}(j) - \beta_{n,p}^{-1} b_{n,p}(p-i+1)\tilde{b}_{n,p}(p-j+1). \tag{11}$$

Carrying out the iteration in (10) further until i or j or both become zero, one gets

$$[\boldsymbol{R}_{n+1}^{-1}(p)]_{i+1,j+1} = \sum_{l=0}^{\min(i,j)} T_{n+1-l}(i-1, j-l). \tag{12}$$

In matrix form, (12) leads to

$$\boldsymbol{R}_{n+1}^{-1}(p) = \boldsymbol{L1}_{n+1}(p)\boldsymbol{D1}_{n+1}(p)\boldsymbol{U\tilde{1}}_{n+1}(p) - \boldsymbol{U1}_n(p)^t\overline{\boldsymbol{D1}}_n(p)\boldsymbol{L\tilde{1}}_n(p)^t \tag{13}$$

where

$$\boldsymbol{L1}_{n+1}(p) = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ a_{n+1,p}(1) & 1 & 0 & \cdots & 0 \\ a_{n+1,p}(2) & a_{n,p}(1) & \cdot & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ a_{n+1,p}(p) & a_{n,p}(p-1) & \cdot & \cdots & 1 \end{bmatrix}$$

$$\boldsymbol{U\tilde{1}}_{n+1}(p) = \begin{bmatrix} 1 & \tilde{a}_{n+1,p}(1) & \tilde{a}_{n+1,p}(2) & \cdots & \tilde{a}_{n+1,p}(p) \\ 0 & 1 & \tilde{a}_{n,p}(1) & \cdots & \tilde{a}_{n,p}(p-1) \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}.$$

$\boldsymbol{D1}_n(p)$ and $\overline{\boldsymbol{D1}}_n(p)$ are two diagonal matrices with $[\boldsymbol{D1}_n(p)]_{i,i} = \alpha_{n-i+1,p}^{-1}$ and $[\overline{\boldsymbol{D1}}_n(p)]_{i,i} = \beta_{n-i+1,p}^{-1}$, $1 \le i \le p+1$, and

$$\boldsymbol{L\tilde{1}}_n(p) = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ \tilde{b}_{n,p}(p) & 0 & 0 & \cdots & 0 \\ \tilde{b}_{n,p}(p-1) & \tilde{b}_{n-1,p}(p) & 0 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \tilde{b}_{n,p}(1) & \tilde{b}_{n-1,p}(2) & \cdot & \cdots & 0 \end{bmatrix}$$

$$\boldsymbol{U1}_n(p) = \begin{bmatrix} 0 & b_{n,p}(p) & b_{n,p}(p-1) & \cdots & b_{n,p}(1) \\ 0 & 0 & b_{n-1,p}(p) & \cdots & b_{n-1,p}(2) \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}.$$

It is interesting to note that the matrices $\boldsymbol{L1}_{n+1,p}, \boldsymbol{U\tilde{1}}_{n+1,p}, \boldsymbol{L\tilde{1}}_n(p)$, and $\boldsymbol{U1}_n(p)$ also constitute periodic Toeplitz matrices with period d.

Equation (13) gives a closed-form expression for $\boldsymbol{R}_n^{-1}(p)$ and can be recognized as the periodic Toeplitz extension of the Gohberg–Semencul formula [6] for Toeplitz matrices.

IV. Discussion and Conclusion

An explicit inversion formula for a set of periodic Toeplitz matrices has been derived, and an efficient algorithm to implement the formula has been presented. The algorithm may be viewed as a non-Hermitian generalization of [7]. Further, for $d = 1$, i.e., for the Toeplitz case, the proposed recursions boil down to those of Trench and Zohar [8]. To see this, first observe that for $d = 1$, $\boldsymbol{R}_n(p) \equiv \boldsymbol{R}_{n+1}(p)$ for all n, implying that the variables used in (9) are independent of n. For convenience, we then take out the index n from these variables. Second, for the Toeplitz case, $\tilde{\boldsymbol{a}}_p = \hat{\boldsymbol{b}}_p$ and $\tilde{\boldsymbol{b}}_p = \hat{\boldsymbol{a}}_p$, where $\hat{\boldsymbol{a}}_p$ and $\hat{\boldsymbol{b}}_p$ are the reversed versions of $\boldsymbol{a}_p$ and $\boldsymbol{b}_p$, respectively, i.e., $\hat{\boldsymbol{a}}_p = \boldsymbol{J}\boldsymbol{a}_p$ and $\hat{\boldsymbol{b}}_p = \boldsymbol{J}\boldsymbol{b}_p$, $\boldsymbol{J}$ being the exchange matrix. This is easily seen by noting that $\boldsymbol{R}(p)$ and $\boldsymbol{R}^t(p)$ can be written as

$$\boldsymbol{R}(p) = \left[\begin{array}{c|c} \boldsymbol{R}(p-1) & \boldsymbol{x} \\ \hline \boldsymbol{y}^t & \end{array}\right] = \left[\begin{array}{c|c} & \hat{\boldsymbol{x}}^t \\ \hline \hat{\boldsymbol{y}}^t & \boldsymbol{R}(p-1) \end{array}\right]$$

$$\boldsymbol{R}^t(p) = \left[\begin{array}{c|c} \boldsymbol{R}^t(p-1) & \boldsymbol{y} \\ \hline \boldsymbol{x}^t & \end{array}\right] = \left[\begin{array}{c|c} & \hat{\boldsymbol{y}}^t \\ \hline \hat{\boldsymbol{x}} & \boldsymbol{R}^t(p-1) \end{array}\right]$$

where $\hat{\boldsymbol{x}} = \boldsymbol{J}\boldsymbol{x}$ and $\hat{\boldsymbol{y}} = \boldsymbol{J}\boldsymbol{y}$. Noting that $\boldsymbol{a}_p = -\boldsymbol{R}^{-1}(p-1)\hat{\boldsymbol{y}}$, $\tilde{\boldsymbol{b}}_p = -[\boldsymbol{R}^t(p-1)]^{-1}\boldsymbol{y}$, and the matrix $\boldsymbol{R}(p-1)$ is persymmetric, i.e., $\boldsymbol{R}^t(p-1) = \boldsymbol{J}\boldsymbol{R}(p-1)\boldsymbol{J}$, it is easy to verify that $\tilde{\boldsymbol{b}}_p = \hat{\boldsymbol{a}}_p$. In a similar manner, it can be checked that $\tilde{\boldsymbol{a}}_p = \hat{\boldsymbol{b}}_p$. In other words, one can directly obtain $\tilde{\boldsymbol{a}}_p$ and $\tilde{\boldsymbol{b}}_p$ from $\boldsymbol{b}_p$ and $\boldsymbol{a}_p$ respectively, simply by reversing the order of the elements. This is, however, not possible when the matrix is periodically Toeplitz, and one has to actually solve the MYW equations for $\boldsymbol{R}_n^t(p)$ to obtain $\tilde{\boldsymbol{a}}_{n,p}$ and $\tilde{\boldsymbol{b}}_{n,p}$.

This correspondence also gives a Gohberg–Semencul-type closed-form expression for $\boldsymbol{R}_n^{-1}(p)$. For Toeplitz matrices, the Gohberg–Semencul expression is computationally equivalent to Trench's inversion formula. This is because the triangular matrices appearing in the former, in this case, are Toeplitz matrices, whose products can be computed efficiently via FFT [9]. It is, however, not clear at this stage and thus remains to be investigated whether similar fast multiplication is feasible when the triangular factors are given to be periodic Toeplitz.

Appendix
$\boldsymbol{UDL}$ and $\boldsymbol{LDU}$ Factorizations of $\boldsymbol{R}_n(p)$

For the Hermitian case, i.e., when $\boldsymbol{R}_n(p) = E[\boldsymbol{y}_n(p)\boldsymbol{y}_n^t(p)]$, the $\boldsymbol{UDL}$ and $\boldsymbol{LDU}$ factorizations of $\boldsymbol{R}_n(p)$ are straightforward and follow the standard procedure of orthogonalization of the set $\{y(n), y(n-1), \cdots, y(n-p)\}$, either by premultiplying $\boldsymbol{y}_n(p)$ with a lower triangular matrix to generate a vector of backward prediction errors for the index n and order varying from 0 to p or by premultiplying $\boldsymbol{y}_n(p)$ with an upper triangular matrix to obtain a vector of forward prediction errors for the indices $n, n-1, \cdots, n-p$ with orders $p, p-1, \cdots, 0$, respectively. For the non-Hermitian case, i.e., for $\boldsymbol{R}_n(p) = E[\boldsymbol{y}_{n-q}(p)\boldsymbol{y}_n^t(p)]$ with $q \neq 0$, this approach, however, does not result in the desired factorizations since in this case, the prediction errors obtained by orthogonalizing the two sets $\{y(n), y(n-1), \cdots, y(n-p)\}$, and $\{y(n-q), y(n-q-1), \cdots, y(n-q-p)\}$ are not mutually orthogonal. We show here how the factorizations of a non-Hermitian $\boldsymbol{R}_n(p)$ can be carried out by generalizing the concept of forward and backward predictions.

Consider the problem of finding out the coefficients $a_{n-j,p-j}(1), \cdots, a_{n-j,p-j}(p-j)$ so that $e_{n,j} = y(n-j) + a_{n-j,p-j}(1)y(n-j-1) + \cdots + a_{n-j,p-j}(p-j)y(n-p)$ is orthogonal to $y(n-q-j-1), y(n-q-j-2), \cdots, y(n-q-p)$ for $0 \le j \le p-1$. Defining $\alpha_{n-j,p-j} = E[y(n-q-j)e_{n,j}]$, one can then write

$$E\left[\boldsymbol{y}_{n-q-j}(p-j)\boldsymbol{y}^t_{n-j}(p-j)\begin{bmatrix}1\\ \boldsymbol{a}_{n-j,\,p-j}\end{bmatrix}\right]=\begin{bmatrix}\alpha_{n-j,\,p-j}\\ \boldsymbol{o}_{(p-j)\times 1}\end{bmatrix} \tag{A.1}$$

where $\boldsymbol{a}_{n-j,\,p-j}=[a_{n-j,\,p-j}(1),a_{n-j,\,p-j}(2),\cdots,a_{n-j,\,p-j}(p-j)]^t$. Equation (A.1) can be identified as the MYW equation (2a) for the matrix $\boldsymbol{R}_{n-j}(p-j)$. Assumption of nonsingularity of $\boldsymbol{R}_{n-j}(p-j)$ guarantees unique choice of the coefficients $a_{n-j,\,p-j}(1),\cdots,a_{n-j,\,p-j}(p-j)$. Similarly, consider the set of coefficients $\tilde{a}_{n-i,\,p-i}(1),\cdots,\tilde{a}_{n-i,\,p-i}(p-i)$ so that $\tilde{e}_{n,\,i}=y(n-q-i)+\tilde{a}_{n-i,\,p-i}(1)y(n-q-i-1)+\cdots+\tilde{a}_{n-i,\,p-i}(p-i)y(n-q-p)$ is orthogonal to $y(n-i-1),y(n-i-2),\cdots,y(n-p)$ for $0\le i\le p-1$. Defining $\tilde{\alpha}_{n-i,\,p-i}=E[y(n-i)\tilde{e}_{n,\,i}]$, one can then write

$$E\left[\boldsymbol{y}_{n-i}(p-i)\boldsymbol{y}^t_{n-q-i}(p-i)\begin{bmatrix}1\\ \tilde{\boldsymbol{a}}_{n-i,\,p-i}\end{bmatrix}\right]=\begin{bmatrix}\tilde{\alpha}_{n-i,\,p-i}\\ \boldsymbol{o}_{(p-i)\times 1}\end{bmatrix} \tag{A.2}$$

where $\tilde{\boldsymbol{a}}_{n-i,\,p-i}=[\tilde{a}_{n-i,\,p-i}(1),\tilde{a}_{n-i,\,p-i}(2),\cdots,\tilde{a}_{n-i,\,p-i}(p-i)]^t$. Equation (A.2) can be recognized as the MYW equation (2a) for the matrix $\boldsymbol{R}^t_{n-i}(p-i)$. It is easy to see that $E[\tilde{e}_{n,\,i}e_{n,\,j}]=0$ for $i\neq j$, since for $i>j$, $e_{n,\,j}$ is orthogonal to $y(n-q-i),\cdots,y(n-q-p)$ and for $i<j$, $\tilde{e}_{n,\,i}$ is orthogonal to $y(n-j),\cdots,y(n-p)$. For $i=j$, $E[\tilde{e}_{n,\,i}e_{n,\,i}]=E[y(n-q-i)e_{n,\,i}]=E[\tilde{e}_{n,\,i}y(n-i)]$, implying that $E[e_{n,\,i}\tilde{e}_{n,\,i}]=\alpha_{n-i,\,p-i}=\tilde{\alpha}_{n-i,\,p-i}$. Using this, constructing the matrices $\tilde{\boldsymbol{U}}_n(p)$ and $\boldsymbol{L}_n(p)$ as given by (6), and defining $e_{n,\,p}=y(n-p)$, $\tilde{e}_{n,\,p}=y(n-q-p)$, we can then write

$$\begin{aligned}&[\tilde{\boldsymbol{U}}_n(p)\boldsymbol{R}_n(p)\boldsymbol{L}_n(p)]_{i,\,j}\\ &\quad=E[\tilde{\boldsymbol{U}}_n(p)\boldsymbol{y}_{n-q}(p)\boldsymbol{y}^t_n(p)\boldsymbol{L}_n(p)]_{i,\,j}\\ &\quad=E[\tilde{e}_{n,\,i}e_{n,\,j}]=\alpha_{n-i,\,p-i}\delta(i-j)\end{aligned}$$

where $\delta(l)$ is the unit sample function. Thus, $\tilde{\boldsymbol{U}}_n(p)\boldsymbol{R}_n(p)\boldsymbol{L}_n(p)=\boldsymbol{D}_n(p)$, or $\boldsymbol{R}_n(p)=(\tilde{\boldsymbol{U}}_n(p))^{-1}\boldsymbol{D}_n(p)(\boldsymbol{L}_n(p))^{-1}$, which gives the factorization of $\boldsymbol{R}_n(p)$.

For $\boldsymbol{LDU}$ factorization, we consider the coefficients $b_{n,\,j}(1),\cdots,b_{n,\,j}(j)$, so that $f_{n,\,j}=b_{n,\,j}(j)y(n)+\cdots+b_{n,\,j}(1)y(n-j+1)+y(n-j)$ is orthogonal to $y(n-q-r)$, $0\le r\le j-1$ for $1\le j\le p$. In other words

$$E\left[\boldsymbol{y}_{n-q}(j)\boldsymbol{y}^t_n(j)\begin{bmatrix}\boldsymbol{b}_{n,\,j}\\ 1\end{bmatrix}\right]=\begin{bmatrix}\boldsymbol{o}_{j\times 1}\\ \beta_{n,\,j}\end{bmatrix} \tag{A.3}$$

where $\boldsymbol{b}_{n,\,j}=[b_{n,\,j}(j),b_{n,\,j}(j-1),\cdots,b_{n,\,j}(1)]^t$, and $\beta_{n,\,j}=E[f_{n,\,j}y(n-q-j)]$. Equation (A.3) is clearly the MYW equation (2b) for the matrix $\boldsymbol{R}_n(j)$. Similarly, let $\tilde{b}_{n,\,i}(1),\cdots,\tilde{b}_{n,\,i}(i)$ denote a set of coefficients for which $\tilde{f}_{n,\,i}=\tilde{b}_{n,\,i}(i)y(n-q)+\cdots+\tilde{b}_{n,\,i}(1)y(n-q-i+1)+y(n-q-i)$ is orthogonal to $y(n-r)$, $0\le r\le i-1$ for $1\le i\le p$. Defining $\tilde{\beta}_{n,\,j}=E[\tilde{f}_{n,\,j}y(n-j)]$, we can then write

$$E\left[\boldsymbol{y}_n(j)\boldsymbol{y}^t_{n-q}(j)\begin{bmatrix}\tilde{\boldsymbol{b}}_{n,\,j}\\ 1\end{bmatrix}\right]=\begin{bmatrix}\boldsymbol{o}_{j\times 1}\\ \tilde{\beta}_{n,\,j}\end{bmatrix} \tag{A.4}$$

where $\tilde{\boldsymbol{b}}_{n,\,j}=[\tilde{b}_{n,\,j}(j),\tilde{b}_{n,\,j}(j-1),\cdots,\tilde{b}_{n,\,j}(1)]^t$. Equation (A.4) gives the MYW equation (2b) for $\boldsymbol{R}^t_n(j)$. Using the stated orthogonality conditions for $f_{n,\,j}$ and $\tilde{f}_{n,\,i}$, it is easy to verify that $E[\tilde{f}_{n,\,i}f_{n,\,j}]=0$ for $i\neq j$. For $i=j$, $E[\tilde{f}_{n,\,i}f_{n,\,i}]=E[y(n-q-i)f_{n,\,i}]=E[\tilde{f}_{n,\,i}y(n-i)]$, meaning that $E[f_{n,\,i}\tilde{f}_{n,\,i}]=\beta_{n,\,i}=\tilde{\beta}_{n,\,i}$. Constructing the matrices $\boldsymbol{U}_n(p)$ and $\tilde{\boldsymbol{L}}_n(p)$ as defined in (3) and defining $f_{n,\,o}=y(n)$, $\tilde{f}_{n,\,0}=y(n-q)$, it then follows that

$$\begin{aligned}&[\tilde{\boldsymbol{L}}_n(p)\boldsymbol{R}_n(p)U_n(p)]_{i,\,j}\\ &\quad=E[\tilde{\boldsymbol{L}}_n(p)\boldsymbol{y}_{n-q}(p)\boldsymbol{y}^t_n(p)\boldsymbol{U}_n(p)]_{i,\,j}\\ &\quad=E[\tilde{f}_{n,\,i}f_{n,\,j}]=\beta_{n,\,i}\delta(i-j)\end{aligned}$$

implying that $\tilde{\boldsymbol{L}}_n(p)\boldsymbol{R}_n(p)\boldsymbol{U}_n(p)=\overline{\boldsymbol{D}}_n(p)$, or $\boldsymbol{R}_n(p)=(\tilde{\boldsymbol{L}}_n(p))^{-1}\overline{\boldsymbol{D}}_n(p)\,(\boldsymbol{U}_n(p))^{-1}$, which gives the $\boldsymbol{LDU}$ factorization of $\boldsymbol{R}_n(p)$.

ACKNOWLEDGMENT

The author is grateful for the constructive suggestions by the reviewers, which have helped to improve this paper.

REFERENCES

[1] H. Sakai, "Circular lattice filtering using Pagano's method," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-30, pp. 279–287, Apr. 1982.
[2] M. Chakraborty and S. Prasad, "Multivariate ARMA modeling by acalar Algorithms," *IEEE Trans. Signal Processing*, vol. 41, pp. 1692–1697, Apr. 1993.
[3] T. Kawase, H. Sakai, and H. Tokumaru, "Recursive least squares circular lattice and escalator estimation algorithms," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-31, pp. 228–231, Feb. 1983.
[4] M. Chakraborty and S. Prasad, "Multichannel ARMA modeling by least squares circular lattice filtering," *IEEE Trans. Signal Processing*, vol. 42, pp. 2304–2318, Sept. 1994.
[5] W. F. Trench, "An algorithm for the inversion of finite Toeplitz matrices," *J. Soc. Indus. Appl. Math.*, vol. 12, no. 3, pp. 515–522, Sept. 1964.
[6] I. C. Gohberg and A. A. Semencul, "On the inversion of finite Toeplitz matrices and their continuous analog," *Mat. Issled.*, vol. 7, pp. 201–223, 1972.
[7] H. Sakai, "Further results on the circular Levinson algorithm," *IEICE Trans. Fundamentals Elect. Commun. Comput. Sci.*, vol. E74, pp. 3962–3967, Dec. 1991.
[8] S. Zohar, "Toeplitz matrix inversion: The algorithm of W. F. Trench," *J. ACM*, vol. 16, no. 4, pp. 592–601, Oct. 1969.
[9] J. Jain, "An efficient algorithm for a large Toeplitz set of linear equations," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-27, pp. 612–615, Dec. 1979.