Probability Theory and Related Fields (2023) 185:513–552
https://doi.org/10.1007/s00440-022-01162-9

Explicit formulas for the inverses of Toeplitz matrices, with applications

Akihiko Inoue[1]

Received: 9 August 2021 / Revised: 5 June 2022 / Accepted: 13 August 2022 /
Published online: 2 September 2022

Abstract
We derive novel explicit formulas for the inverses of truncated block Toeplitz matrices that correspond to a multivariate minimal stationary process. The main ingredients of the formulas are the Fourier coefficients of the phase function attached to the spectral density of the process. The derivation of the formulas is based on a recently developed finite prediction theory applied to the dual process of the stationary process. We illustrate the usefulness of the formulas by two applications. The first one is a strong convergence result for solutions of general block Toeplitz systems for a multivariate short-memory process. The second application is closed-form formulas for the inverses of truncated block Toeplitz matrices corresponding to a multivariate ARMA process. The significance of the latter is that they provide us with a linear-time algorithm to compute the solutions of corresponding block Toeplitz systems.

Keywords Toeplitz matrix · Finite prediction · Dual process · Toeplitz system · Linear-time algorithm

Mathematics Subject Classification 60G10 · 60G25 · 15B05 · 65F05

1 Introduction

Let $\mathbb{T} := \{z \in \mathbb{C} : |z| = 1\}$ be the unit circle in $\mathbb{C}$. We write σ for the normalized Lebesgue measure $d\theta/(2\pi)$ on $([-\pi, \pi), \mathcal{B}([-\pi, \pi)))$, where $\mathcal{B}([-\pi, \pi))$ is the Borel σ-algebra on $[-\pi, \pi)$; thus we have $\sigma([-\pi, \pi)) = 1$. For $p \in [1, \infty)$, we write $L_p(\mathbb{T})$ for the Lebesgue space of measurable functions $f : \mathbb{T} \to \mathbb{C}$ such that $\|f\|_p < \infty$, where $\|f\|_p := \{\int_{-\pi}^{\pi} |f(e^{i\theta})|^p \sigma(d\theta)\}^{1/p}$. Let $L_p^{m\times n}(\mathbb{T})$ be the space of $\mathbb{C}^{m\times n}$-valued functions on $\mathbb{T}$ whose entries belong to $L_p(\mathbb{T})$.

✉ Akihiko Inoue
inoue100@hiroshima-u.ac.jp

1 Department of Mathematics, Hiroshima University, Higashi-Hiroshima 739-8526, Japan

Let $d \in \mathbb{N}$. For $n \in \mathbb{N}$, we consider the block Toeplitz matrix

$$T_n(w) := \begin{pmatrix} \gamma(0) & \gamma(-1) & \cdots & \gamma(-n+1) \\ \gamma(1) & \gamma(0) & \cdots & \gamma(-n+2) \\ \vdots & \vdots & \ddots & \vdots \\ \gamma(n-1) & \gamma(n-2) & \cdots & \gamma(0) \end{pmatrix} \in \mathbb{C}^{dn\times dn},$$

where

$$\gamma(k) := \int_{-\pi}^{\pi} e^{-ik\theta} w(e^{i\theta}) \frac{d\theta}{2\pi} \in \mathbb{C}^{d\times d}, \quad k \in \mathbb{Z}, \tag{1.1}$$

and the symbol w satisfies the following two conditions:

$$w \in L_1^{d\times d}(\mathbb{T}) \text{ and } w(e^{i\theta}) \text{ is a positive Hermitian matrix } \sigma\text{-a.e.}, \tag{1.2}$$

$$w^{-1} \in L_1^{d\times d}(\mathbb{T}). \tag{1.3}$$

Let $\{X_k : k \in \mathbb{Z}\}$ be a $\mathbb{C}^d$-valued, centered, weakly stationary process that has spectral density w, hence autocovariance function γ. Then the conditions (1.2) and (1.3) imply that $\{X_k\}$ is *minimal* (see Sect. 10 of [21, Chapter II]).

In this paper, we show novel explicit formulas for $T_n(w)^{-1}$ (Theorem 2.1), which are especially useful for large n (see [2]). The formulas are new even for $d = 1$. The main ingredients of the formulas are the Fourier coefficients of $h^* h_\sharp^{-1} = h^{-1} h_\sharp^*$, where h and $h_\sharp$ are $\mathbb{C}^{d\times d}$-valued outer functions on $\mathbb{T}$ such that

$$w(e^{i\theta}) = h(e^{i\theta}) h(e^{i\theta})^* = h_\sharp(e^{i\theta})^* h_\sharp(e^{i\theta}), \qquad \sigma\text{-a.e.} \tag{1.4}$$

(see [10]; see also Sect. 2). We note that the unitary matrix valued function $h^* h_\sharp^{-1} = h^{-1} h_\sharp^*$ on $\mathbb{T}$ attached to w is called the *phase function* of w (see page 428 in [20]).

Let $\{X_k\}$ be as above, and let $\{X'_k : k \in \mathbb{Z}\}$ be the dual process of $\{X_k\}$ (see [19]; see also Sect. 2 below). In the proof of the above explicit formulas for $T_n(w)^{-1}$, the dual process $\{X'_k\}$ plays an important role. In fact, the key to the proof of the explicit formulas for $T_n(w)^{-1}$ is the following equality (Theorem 3.1):

$$\left(T_n(w)^{-1}\right)^{s,t} = \langle X'_s, P_{[1,n]} X'_t \rangle, \quad s, t \in \{1, \ldots, n\}. \tag{1.5}$$

Here, $\langle \cdot, \cdot \rangle$ stands for the Gram matrix (see Sect. 3) and $P_{[1,n]} X'_t$ denotes the best linear predictor of X'_t based on the observations $X_1, \ldots, X_n$ (see Sect. 2 for the precise definition). Moreover, for $n \in \mathbb{N}$, $A \in \mathbb{C}^{dn\times dn}$ and $s, t \in \{1, \ldots, n\}$, we write $A^{s,t} \in \mathbb{C}^{d\times d}$ for the (s, t) block of A; thus $A = (A^{s,t})_{1\le s,t\le n}$. The equality (1.5) enables us to apply the $P_{[1,n]}$-related methods developed in [11, 12, 14–16] and others to derive the explicit formulas for $T_n(w)^{-1}$.

We illustrate the usefulness of the explicit formulas for $T_n(w)^{-1}$ by two applications. The first one is a strong convergence result for solutions of block Toeplitz systems. For this application, we assume (1.2) as well as the following condition:

$$\sum_{k=-\infty}^{\infty} \|\gamma(k)\| < \infty \text{ and } \min_{z\in\mathbb{T}} \det w(z) > 0. \tag{1.6}$$

Here, for $a \in \mathbb{C}^{d\times d}$, $\|a\|$ denotes the operator norm of a. The condition (1.6) implies that $\{X_k\}$ with spectral density w is a *short-memory* process. We note that (1.3) follows from (1.2) and (1.6) (see Sect. 4). Under (1.2) and (1.6), for $n \in \mathbb{N}$ and a $\mathbb{C}^{d\times d}$-valued sequence $\{y_k\}_{k=1}^{\infty}$ such that $\sum_{k=1}^{\infty} \|y_k\| < \infty$, let

$$Z_n = (z_{n,1}^\top, \dots, z_{n,n}^\top)^\top \in \mathbb{C}^{dn\times d} \text{ with } z_{n,k} \in \mathbb{C}^{d\times d},\ k \in \{1, \dots, n\}, \tag{1.7}$$

be the solution to the block Toeplitz system

$$T_n(w) Z_n = Y_n, \tag{1.8}$$

where

$$Y_n := (y_1^\top, \dots, y_n^\top)^\top \in \mathbb{C}^{dn\times d}. \tag{1.9}$$

Also, let

$$Z_\infty = (z_1^\top, z_2^\top, \dots)^\top \text{ with } z_k \in \mathbb{C}^{d\times d},\ k \in \mathbb{N}, \tag{1.10}$$

be the solution to the corresponding infinite block Toeplitz system

$$T_\infty(w) Z_\infty = Y_\infty, \tag{1.11}$$

where

$$T_\infty(w) := \begin{pmatrix} \gamma(0) & \gamma(-1) & \gamma(-2) & \cdots \\ \gamma(1) & \gamma(0) & \gamma(-1) & \cdots \\ \gamma(2) & \gamma(1) & \gamma(0) & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \tag{1.12}$$

and

$$Y_\infty := (y_1^\top, y_2^\top, \dots)^\top. \tag{1.13}$$

Then, our result (Theorem 4.1) reads as follows:

$$\lim_{n\to\infty} \sum_{k=1}^{n} \|z_{n,k} - z_k\| = 0. \tag{1.14}$$

We explain the background of the result (1.14). As above, let $\{X_k : k \in \mathbb{Z}\}$ be a $\mathbb{C}^d$-valued, centered, weakly stationary process that has spectral density w. For $n \in \mathbb{N}$, the finite and infinite predictor coefficients $\phi_{n,k} \in \mathbb{C}^{d\times d}$, $k \in \{1, \ldots, n\}$, and ϕ_k, $k \in \mathbb{N}$, of $\{X_k\}$ are defined by

$$P_{[1,n]}X_{n+1} = \sum_{k=1}^{n} \phi_{n,k} X_{n+1-k} \quad \text{and} \quad P_{(-\infty,n]}X_{n+1} = \sum_{k=1}^{\infty} \phi_k X_{n+1-k},$$

respectively; see Sect. 3 for the precise definitions of $P_{[1,n]}$ and $P_{(-\infty,n]}$. We note that $\sum_{k=1}^{\infty} \|\phi_k\| < \infty$ holds under (1.2) and (1.6) (see Sect. 4 below and (2.16) in [16]). *Baxter's inequality* in [1, 5, 9] states that, under (1.2) and (1.6), there exists $K \in (0, \infty)$ such that

$$\sum_{k=1}^{n} \|\phi_{n,k} - \phi_k\| \le K \sum_{k=n+1}^{\infty} \|\phi_k\|, \quad n \in \mathbb{N}. \tag{1.15}$$

In particular, we have

$$\lim_{n\to\infty} \sum_{k=1}^{n} \|\phi_{n,k} - \phi_k\| = 0. \tag{1.16}$$

If we put $\tilde{w}(e^{i\theta}) := w(e^{-i\theta})$, then, $(\phi_{n,1}, \ldots, \phi_{n,n})$ is the solution to the block Toeplitz system

$$T_n(\tilde{w})(\phi_{n,1}, \ldots, \phi_{n,n})^* = (\gamma(1), \ldots, \gamma(n))^*,$$

called the *Yule–Walker equation*, while $(\phi_1, \phi_2, \ldots)$ is the solution to the corresponding infinite block Toeplitz system

$$T_\infty(\tilde{w})(\phi_1, \phi_2, \ldots)^* = (\gamma(1), \gamma(2), \ldots)^*.$$

Clearly, $\tilde{w}$ satisfies (1.2) and (1.6) since so does w. Therefore, our result (1.14) can be viewed as an extension to (1.16). It should be noted, however, that we prove (1.14) directly, without proving an analogue of Baxter's inequality (1.15).

The convergence result (1.16) has various applications in time series analysis, such as the autoregressive sieve bootstrap (see, e.g., [16] and the references therein), whille Toeplitz systems of the form (1.8) appear in various fields, such as filtering of signals. Therefore the extension (1.14), as well as the other results explained below, may potentially be useful in such fields. We note that Baxter's inequality (1.15), hence (1.16), is also proved for univariate and multivariate FARIMA (fractional autoregressive integrated moving-average) processes, which are long-memory processes, in [14] and [16], respectively. The FARIMA processes have singular spectral densities w but our explicit formulas for $T_n(w)^{-1}$ above also cover them since we only assume minimality in the formulas. Applications of the explicit formulas to univariate and multivariate FARIMA processes will be discussed elsewhere. However, the problem of proving results of the type (1.14) for FARIMA processes remains unsolved so far.

The second application of the explicit formulas for $T_n(w)^{-1}$ is closed-form formulas for $T_n(w)^{-1}$ with rational w that corresponds to a univariate ($d = 1$) or multivariate ($d \geq 2$) ARMA (autoregressive moving-average) process (Theorem 5.2). More precisely, we assume that w is of the form

$$w(e^{i\theta}) = h(e^{i\theta})h(e^{i\theta})^*, \quad \theta \in [-\pi, \pi), \tag{1.17}$$

where $h : \mathbb{T} \to \mathbb{C}^{d\times d}$ satisfies the following condition:

$$\begin{aligned}&\text{the entries of } h(z) \text{ are rational functions in } z \text{ that have}\\ &\text{no poles in } \overline{\mathbb{D}}, \text{ and } \det h(z) \text{ has no zeros in } \overline{\mathbb{D}}.\end{aligned} \tag{1.18}$$

Here $\overline{\mathbb{D}} := \{z \in \mathbb{C} : |z| \leq 1\}$ is the closed unit disk in $\mathbb{C}$. The closed-form formulas for $T_n(w)^{-1}$ consist of several building block matrices that are of fixed sizes independent of n. The significance of the formulas for $T_n(w)^{-1}$ is that they provide us with a linear-time, or $O(n)$, algorithm to compute the solution $Z \in \mathbb{C}^{dn\times d}$ to the block Toeplitz system

$$T_n(w)Z = Y \tag{1.19}$$

for $Y \in \mathbb{C}^{dn\times d}$ (see Sect. 6). The famous Durbin–Levinson algorithm solves the Eq. (1.19) for more general w in $O(n^2)$ time. Algorithms for Toeplitz linear systems that run faster than $O(n^2)$ are called *superfast*. While our algorithm is restricted to the class of w corresponding to ARMA processes, the class is important in applications, and the linear-time algorithm is ideally superfast in the sense that there is no algorithm faster than $O(n)$.

Toeplitz matrices appear in a variety of fields, including operator theory, orthogonal polynomials on the unit circle, time series analysis, engineering, and physics. Therefore, there is a vast amount of literature on Toeplitz matrices. Here, we refer to [2, 3, 6, 8, 22, 23] and [24] as textbook treatments. For example, in [6, III], the Gohberg-Semencul formulas in [7], which express the inverse of a Toeplitz matrix as a difference of products of lower and upper triangular Toeplitz matrices, are explained.

After this work was completed, the author learned of [25] by Subba Rao and Yang, where they also provide an explicit series expansion for $T_n(w)^{-1}$ that corresponds to a univariate stationary process satisfying some conditions (see [25], Sect. 3.2). The main aim of [25] is to reconcile the Gaussian and Whittle likelihood, and the series expansion in [25] is tailored to this purpose, using the *complete DFT* (discrete Fourier transform) introduced in [25]. It should be noticed that $T_n(w)^{-1}$ appears in the Gaussian likelihood, while the Whittle likelihood is based on the ordinary DFT. Since most results of the present paper directly concern $T_n(w)^{-1}$, some of them may also be useful for studies related to the Gaussian likelihood.

This paper is organized as follows. We state the explicit formulas for $T_n(w)^{-1}$ in Sect. 2. In Sect. 3, we first prove (1.5) and then use it to prove the explicit formulas for $T_n(w)^{-1}$. In Sect. 4, we prove (1.14) for w satisfying (1.2) and (1.6), using the explicit formulas for $T_n(w)^{-1}$. In Sect. 5, we prove the closed-form formulas for $T_n(w)^{-1}$ with w satisfying (1.18), using the explicit formulas for $T_n(w)^{-1}$. In Sect. 6, we explain how the results in Sect. 5 give a linear-time algorithm to compute the solution to (1.19). Finally, the Appendix contains the omitted proofs of two lemmas.

2 Explicit formulas

Let $\mathbb{C}^{m\times n}$ be the set of all complex $m \times n$ matrices; we write $\mathbb{C}^d$ for $\mathbb{C}^{d\times 1}$. Let I_n be the $n \times n$ unit matrix. For $a \in \mathbb{C}^{m\times n}$, $a^\top$ denotes the transpose of a, and $\overline{a}$ and a^* the complex and Hermitian conjugates of a, respectively; thus, in particular, $a^* := \overline{a}^\top$. For $a \in \mathbb{C}^{d\times d}$, we write $\|a\|$ for the operator norm of a:

$$\|a\| := \sup_{u\in\mathbb{C}^d, |u|\le 1} |au|.$$

Here $|u| := (\sum_{i=1}^d |u^i|^2)^{1/2}$ denotes the Euclidean norm of $u = (u^1, \dots, u^d)^\top \in \mathbb{C}^d$. For $p \in [1,\infty)$ and $K \subset \mathbb{Z}$, $\ell_p^{d\times d}(K)$ denotes the space of $\mathbb{C}^{d\times d}$-valued sequences $\{a_k\}_{k\in K}$ such that $\sum_{k\in K} \|a_k\|^p < \infty$. We write $\ell_{p+}^{d\times d}$ for $\ell_p^{d\times d}(\mathbb{N}\cup\{0\})$ and ℓ_{p+} for $\ell_{p+}^{1\times 1} = \ell_p^{1\times 1}(\mathbb{N}\cup\{0\})$.

Recall σ from Sect. 1. The Hardy class $H_2(\mathbb{T})$ on $\mathbb{T}$ is the closed subspace of $L_2(\mathbb{T})$ consisting of $f \in L_2(\mathbb{T})$ such that $\int_{-\pi}^{\pi} e^{im\theta} f(e^{i\theta})\sigma(d\theta) = 0$ for $m = 1, 2, \dots$. Let $H_2^{m\times n}(\mathbb{T})$ be the space of $\mathbb{C}^{m\times n}$-valued functions on $\mathbb{T}$ whose entries belong to $H_2(\mathbb{T})$. Let $\mathbb{D} := \{z \in \mathbb{C} : |z|<1\}$ be the open unit disk in $\mathbb{C}$. We write $H_2(\mathbb{D})$ for the Hardy class on $\mathbb{D}$, consisting of holomorphic functions f on $\mathbb{D}$ such that $\sup_{r\in[0,1)} \int_{-\pi}^{\pi} |f(re^{i\theta})|^2\sigma(d\theta) < \infty$. As usual, we identify each function f in $H_2(\mathbb{D})$ with its boundary function $f(e^{i\theta}) := \lim_{r\uparrow 1} f(re^{i\theta})$, σ-a.e., in $H_2(\mathbb{T})$. A function h in $H_2^{d\times d}(\mathbb{T})$ is called *outer* if $\det h$ is a $\mathbb{C}$-valued outer function, that is, $\det h$ satisfies $\log|\det h(0)| = \int_{-\pi}^{\pi} \log|\det h(e^{i\theta})|\sigma(d\theta)$ (see Definition 3.1 in [18]).

We assume that w satisfies (1.2) and (1.3). Then $\log\det w$ is in $L_1(\mathbb{T})$ (see Sect. 3 in [16]). Therefore w has the decompositions (1.4) for two outer functions h and $h_\sharp$ belonging to $H_2^{d\times d}(\mathbb{T})$, and h and $h_\sharp$ are unique up to constant unitary factors (see Chapter II in [21] and Theorem 11 in [10]; see also Sect. 3 in [16]). We may take $h_\sharp = h$ for the case $d = 1$ but there is no such simple relation between h and $h_\sharp$ for $d \ge 2$. We define the outer function $\tilde{h}$ in $H_2^{d\times d}(\mathbb{T})$ by

$$\tilde{h}(z) := \{h_\sharp(\overline{z})\}^*. \tag{2.1}$$

All of h^{-1}, $h_\sharp^{-1}$ and $\tilde{h}^{-1}$ also belong to $H_2^{d\times d}(\mathbb{T})$ since we have assumed (1.3).

We define four $\mathbb{C}^{d\times d}$-valued sequences $\{c_k\}$, $\{a_k\}$, $\{\tilde{c}_k\}$ and $\{\tilde{a}_k\}$ by

$$h(z) = \sum_{k=0}^{\infty} z^k c_k, \quad z \in \mathbb{D}, \tag{2.2}$$

$$-h(z)^{-1} = \sum_{k=0}^{\infty} z^k a_k, \quad z \in \mathbb{D}, \tag{2.3}$$

$$\tilde{h}(z) = \sum_{k=0}^{\infty} z^k \tilde{c}_k, \quad z \in \mathbb{D}, \tag{2.4}$$

and

$$-\tilde{h}(z)^{-1} = \sum_{k=0}^{\infty} z^k \tilde{a}_k, \quad z \in \mathbb{D}, \tag{2.5}$$

respectively. By (1.3), all of $\{c_k\}$, $\{a_k\}$, $\{\tilde{c}_k\}$ and $\{\tilde{a}_k\}$ belong to $\ell_{2+}^{d\times d}$.

We define a $\mathbb{C}^{d\times d}$-valued sequence $\{\beta_k\}_{k=-\infty}^{\infty}$ as the (minus of the) Fourier coefficients of the phase function $h^* h_\sharp^{-1} = h^{-1} h_\sharp^*$:

$$\begin{aligned}\beta_k &= -\int_{-\pi}^{\pi} e^{-ik\theta} h(e^{i\theta})^* h_\sharp(e^{i\theta})^{-1} \frac{d\theta}{2\pi} \\ &= -\int_{-\pi}^{\pi} e^{-ik\theta} h(e^{i\theta})^{-1} h_\sharp(e^{i\theta})^* \frac{d\theta}{2\pi}, \quad k \in \mathbb{Z}.\end{aligned} \tag{2.6}$$

For $n \in \mathbb{N}$, $u \in \{1, \ldots, n\}$ and $k \in \mathbb{N}$, we can define the sequences $\{b_{n,u,\ell}^k\}_{\ell=0}^{\infty} \in \ell_{2+}^{d\times d}$ by the recursion

$$\begin{cases} b_{n,u,\ell}^1 = \beta_{u+\ell}, \\ b_{n,u,\ell}^{2k} = \sum_{m=0}^{\infty} b_{n,u,m}^{2k-1} \beta_{n+1+m+\ell}^*, \quad b_{n,u,\ell}^{2k+1} = \sum_{m=0}^{\infty} b_{n,u,m}^{2k} \beta_{n+1+m+\ell} \end{cases} \tag{2.7}$$

(see Sect. 3 below). Similarly, for $n \in \mathbb{N}$, $u \in \{1, \ldots, n\}$ and $k \in \mathbb{N}$, we can define the sequences $\{\tilde{b}_{n,u,\ell}^k\}_{\ell=0}^{\infty} \in \ell_{2+}^{d\times d}$ by the recursion

$$\begin{cases} \tilde{b}_{n,u,\ell}^1 = \beta_{n+1-u+\ell}^*, \\ \tilde{b}_{n,u,\ell}^{2k} = \sum_{m=0}^{\infty} \tilde{b}_{n,u,m}^{2k-1} \beta_{n+1+m+\ell}, \quad \tilde{b}_{n,u,\ell}^{2k+1} = \sum_{m=0}^{\infty} \tilde{b}_{n,u,m}^{2k} \beta_{n+1+m+\ell}^*. \end{cases} \tag{2.8}$$

Recall from Sect. 1 that $(T_n(w)^{-1})^{s,t}$ denotes the (s,t) block of $T_n(w)^{-1}$. Since $T_n(w)$, hence $T_n(w)^{-1}$, is self-adjoint, we have

$$(T_n(w)^{-1})^{s,t} = ((T_n(w)^{-1})^{t,s})^*, \quad s, t \in \{1, \ldots, n\}. \tag{2.9}$$

We use the following notation:

$$s \vee t := \max(s,t), \quad s \wedge t := \min(s,t).$$

We are ready to state the explicit formulas for $(T_n(w))^{-1}$.

Theorem 2.1 *We assume* (1.2) *and* (1.3). *Then the following two assertions hold.*

(i) For $n \in \mathbb{N}$ *and* $s, t \in \{1, \dots, n\}$, *we have*

$$\left(T_n(w)^{-1}\right)^{s,t} = \sum_{\ell=1}^{s\wedge t} \tilde{a}^*_{s-\ell}\tilde{a}_{t-\ell} + \sum_{u=1}^{t}\sum_{k=1}^{\infty}\left\{\sum_{\ell=0}^{\infty} \tilde{b}^{2k-1}_{n,u,\ell} a_{n+1-s+\ell} + \sum_{\ell=0}^{\infty} \tilde{b}^{2k}_{n,u,\ell}\tilde{a}_{s+\ell}\right\}^* \tilde{a}_{t-u}. \tag{2.10}$$

(ii) For $n \in \mathbb{N}$ *and* $s, t \in \{1, \dots, n\}$, *we have*

$$\left(T_n(w)^{-1}\right)^{s,t} = \sum_{\ell=s\vee t}^{n} a^*_{\ell-s}a_{\ell-t} + \sum_{u=t}^{n}\sum_{k=1}^{\infty}\left\{\sum_{\ell=0}^{\infty} b^{2k-1}_{n,u,\ell} \tilde{a}_{s+\ell} + \sum_{\ell=0}^{\infty} b^{2k}_{n,u,\ell} a_{n+1-s+\ell}\right\}^* a_{u-t}. \tag{2.11}$$

The proof of Theorem 2.1 will be given in Sect. 3.

Corollary 2.1 *We assume* (1.2) *and* (1.3). *Then the following two assertions hold.*

(i) For $n \in \mathbb{N}$ *and* $s, t \in \{1, \dots, n\}$, *we have*

$$\left(T_n(w)^{-1}\right)^{s,t} = \sum_{\ell=1}^{s\wedge t} \tilde{a}^*_{s-\ell}\tilde{a}_{t-\ell} + \sum_{u=1}^{s} \tilde{a}^*_{s-u}\sum_{k=1}^{\infty}\left\{\sum_{\ell=0}^{\infty} \tilde{b}^{2k-1}_{n,u,\ell} a_{n+1-t+\ell} + \sum_{\ell=0}^{\infty} \tilde{b}^{2k}_{n,u,\ell}\tilde{a}_{t+\ell}\right\}. \tag{2.12}$$

(ii) For $n \in \mathbb{N}$ *and* $s, t \in \{1, \dots, n\}$, *we have*

$$\left(T_n(w)^{-1}\right)^{s,t} = \sum_{\ell=s\vee t}^{n} a^*_{\ell-s}a_{\ell-t} + \sum_{u=s}^{n} a^*_{u-s}\sum_{k=1}^{\infty}\left\{\sum_{\ell=0}^{\infty} b^{2k-1}_{n,u,\ell} \tilde{a}_{t+\ell} + \sum_{\ell=0}^{\infty} b^{2k}_{n,u,\ell} a_{n+1-t+\ell}\right\}. \tag{2.13}$$

Proof Thanks to (2.9), we obtain (2.12) and (2.13) from (2.10) and (2.11), respectively. □

Remark 2.1 Recall $T_\infty(w)$ from (1.12). For $n \in \mathbb{N}\cup\{0\}$, we have $\gamma(n) = \sum_{k=0}^{\infty} \tilde{c}_k \tilde{c}_{n+k}^*$ and $\gamma(-n) = \sum_{k=0}^{\infty} \tilde{c}_{n+k}\tilde{c}_k^*$ (see (2.13) in [16]), hence $T_\infty(w) = \tilde{C}_\infty(\tilde{C}_\infty)^*$, where

$$\tilde{C}_\infty := \begin{pmatrix} \tilde{c}_0 & \tilde{c}_1 & \tilde{c}_2 & \cdots \\ & \tilde{c}_0 & \tilde{c}_1 & \cdots \\ & & \tilde{c}_0 & \cdots \\ 0 & & & \ddots \end{pmatrix}.$$

On the other hand, it follows from $\tilde{h}(z)\tilde{h}(z)^{-1} = I_d$ that $\sum_{k=0}^{n} \tilde{c}_k \tilde{a}_{n-k} = -\delta_{n0} I_d$ for $n \in \mathbb{N}\cup\{0\}$, hence $\tilde{C}_\infty \tilde{A}_\infty = -I_\infty$, where

$$\tilde{A}_\infty := \begin{pmatrix} \tilde{a}_0 & \tilde{a}_1 & \tilde{a}_2 & \cdots \\ & \tilde{a}_0 & \tilde{a}_1 & \cdots \\ & & \tilde{a}_0 & \cdots \\ 0 & & & \ddots \end{pmatrix}, \quad I_\infty := \begin{pmatrix} 1 & 0 & 0 & \cdots \\ 0 & 1 & 0 & \cdots \\ 0 & 0 & 1 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

Combining, we have $T_\infty(w)^{-1} = (\tilde{A}_\infty)^* \tilde{A}_\infty$. Thus, we find that the first term $\sum_{\ell=1}^{s\wedge t} \tilde{a}_{s-\ell}^* \tilde{a}_{t-\ell}$ in (2.10) or (2.12) coincides with the (s,t) block of $T_\infty(w)^{-1}$.

For $n \in \mathbb{N}$, we define

$$\tilde{A}_n := \begin{pmatrix} \tilde{a}_0 & \tilde{a}_1 & \tilde{a}_2 & \cdots & \tilde{a}_{n-1} \\ & \tilde{a}_0 & \tilde{a}_1 & \cdots & \tilde{a}_{n-2} \\ & & \ddots & \ddots & \vdots \\ & & & \ddots & \tilde{a}_1 \\ 0 & & & & \tilde{a}_0 \end{pmatrix} \in \mathbb{C}^{dn\times dn} \tag{2.14}$$

and

$$A_n := \begin{pmatrix} a_0 & & & & 0 \\ a_1 & a_0 & & & \\ a_2 & a_1 & \ddots & & \\ \vdots & \vdots & \ddots & \ddots & \\ a_{n-1} & a_{n-2} & \cdots & a_1 & a_0 \end{pmatrix} \in \mathbb{C}^{dn\times dn}. \tag{2.15}$$

The next lemma will turn out to be useful in Sect. 6.

Lemma 2.1 *For $n \in \mathbb{N}$ and $s, t \in \{1, \ldots, n\}$, we have the following two equalities:*

$$\left(\tilde{A}_n^* \tilde{A}_n\right)^{s,t} = \sum_{\ell=1}^{s \wedge t} \tilde{a}_{s-\ell}^* \tilde{a}_{t-\ell},$$

$$\left(A_n^* A_n\right)^{s,t} = \sum_{\ell=s \vee t}^{n} a_{\ell-s}^* a_{\ell-t}.$$

The proof of Lemma 2.1 is straightforward and will be omitted.

3 Proof of Theorem 2.1

In this section, we prove Theorem 2.1. We assume (1.2) and (1.3). Let $\{X_k\} = \{X_k : k \in \mathbb{Z}\}$ be a $\mathbb{C}^d$-valued, centered, weakly stationary process, defined on a probability space $(\Omega, \mathcal{F}, P)$, that has spectral density w, hence autocovariance function γ. Thus we have $E[X_k X_0^*] = \gamma(k) = \int_{-\pi}^{\pi} e^{-ik\theta} w(e^{i\theta})(d\theta/(2\pi))$ for $k \in \mathbb{Z}$.

Write $X_k = (X_k^1, \ldots, X_k^d)^\top$, and let V be the complex Hilbert space spanned by all the entries $\{X_k^j : k \in \mathbb{Z},\ j = 1, \ldots, d\}$ in $L^2(\Omega, \mathcal{F}, P)$, which has inner product $(x, y)_V := E[x\overline{y}]$ and norm $\|x\|_V := (x, x)_V^{1/2}$. For $J \subset \mathbb{Z}$ such as $\{n\}$, $(-\infty, n] := \{n, n-1, \ldots\}$, $[n, \infty) := \{n, n+1, \ldots\}$, and $[m, n] := \{m, \ldots, n\}$ with $m \leq n$, we define the closed subspace V_J^X of V by

$$V_J^X := \overline{\mathrm{sp}}\{X_k^j : j = 1, \ldots, d,\ k \in J\}.$$

Let P_J and $P_J^\perp$ be the orthogonal projection operators of V onto V_J^X and $(V_J^X)^\perp$, respectively, where $(V_J^X)^\perp$ denotes the orthogonal complement of V_J^X in V.

By Theorem 3.1 in [11] for $d = 1$ and Corollary 3.6 in [15] for general $d \geq 1$, the conditions (1.2) and (1.3) imply the following *intersection of past and future* property:

$$V_{(-\infty,n]}^X \cap V_{[1,\infty)}^X = V_{[1,n]}^X, \quad n \in \mathbb{N}. \tag{3.1}$$

Let V^d be the space of $\mathbb{C}^d$-valued random variables on $(\Omega, \mathcal{F}, P)$ whose entries belong to V. The norm $\|x\|_{V^d}$ of $x = (x^1, \ldots, x^d)^\top \in V^d$ is given by $\|x\|_{V^d} := (\sum_{i=1}^d \|x^i\|_V^2)^{1/2}$. For $J \subset \mathbb{Z}$ and $x = (x^1, \ldots, x^d)^\top \in V^d$, we write $P_J x$ for $(P_J x^1, \ldots, P_J x^d)^\top$. We define $P_J^\perp x$ in a similar way. For $x = (x^1, \ldots, x^d)^\top$ and $y = (y^1, \ldots, y^d)^\top$ in V^d,

$$\langle x, y \rangle := E[xy^*] = ((x^k, y^\ell)_V)_{1 \leq k, \ell \leq d} \in \mathbb{C}^{d \times d}$$

stands for the Gram matrix of x and y.

Let

$$X_k = \int_{-\pi}^{\pi} e^{-ik\theta} \eta(d\theta), \quad k \in \mathbb{Z},$$

be the spectral representation of $\{X_k\}$, where η is a $\mathbb{C}^d$-valued random spectral measure. We define a d-variate stationary process $\{\varepsilon_k : k \in \mathbb{Z}\}$, called the *forward innovation process* of $\{X_k\}$, by

$$\varepsilon_k := \int_{-\pi}^{\pi} e^{-ik\theta} h(e^{i\theta})^{-1} \eta(d\theta), \quad k \in \mathbb{Z}.$$

Then, $\{\varepsilon_k\}$ satisfies $\langle \varepsilon_n, \varepsilon_m \rangle = \delta_{nm} I_d$ and $V^X_{(-\infty,n]} = V^{\varepsilon}_{(-\infty,n]}$ for $n \in \mathbb{Z}$, hence

$$(V^X_{(-\infty,n]})^{\perp} = V^{\varepsilon}_{[n+1,\infty)}, \quad n \in \mathbb{Z}.$$

Recall the outer function $h_\sharp$ in $H_2^{d\times d}(\mathbb{T})$ from (1.4). We define the *backward innovation process* $\{\tilde{\varepsilon}_k : k \in \mathbb{Z}\}$ of $\{X_k\}$ by

$$\tilde{\varepsilon}_k := \int_{-\pi}^{\pi} e^{ik\theta} \{h_\sharp(e^{i\theta})^*\}^{-1} \eta(d\theta), \quad k \in \mathbb{Z}.$$

Then, $\{\tilde{\varepsilon}_k\}$ satisfies $\langle \tilde{\varepsilon}_n, \tilde{\varepsilon}_m \rangle = \delta_{nm} I_d$ and $V^X_{[-n,\infty)} = V^{\tilde{\varepsilon}}_{(-\infty,n]}$ for $n \in \mathbb{Z}$, hence

$$(V^X_{[-n,\infty)})^{\perp} = V^{\tilde{\varepsilon}}_{[n+1,\infty)}, \quad n \in \mathbb{Z}$$

(see Sect. 2 in [16]). Moreover, by Lemma 4.1 in [16], we have

$$\langle \varepsilon_\ell, \tilde{\varepsilon}_m \rangle = -\beta_{\ell+m}, \quad \langle \tilde{\varepsilon}_m, \varepsilon_\ell \rangle = -\beta^*_{\ell+m}, \quad \ell, m \in \mathbb{Z}. \tag{3.2}$$

By (3.2), for $\{s_\ell\} \in \ell_{2+}^{d\times d}$ and $n \in \mathbb{N}$,

$$P^{\perp}_{[1,\infty)}\left(\sum_{\ell=0}^{\infty} s_\ell \varepsilon_{n+1+\ell}\right) = -\sum_{\ell=0}^{\infty}\left(\sum_{m=0}^{\infty} s_m \beta_{n+1+\ell+m}\right)\tilde{\varepsilon}_\ell, \tag{3.3}$$

$$P^{\perp}_{(-\infty,n]}\left(\sum_{\ell=0}^{\infty} s_\ell \tilde{\varepsilon}_\ell\right) = -\sum_{\ell=0}^{\infty}\left(\sum_{m=0}^{\infty} s_m \beta^*_{n+1+\ell+m}\right)\varepsilon_{n+1+\ell}. \tag{3.4}$$

Therefore,

$$\left\{\sum_{m=0}^{\infty} s_m \beta_{n+1+\ell+m}\right\}_{\ell=0}^{\infty}, \quad \left\{\sum_{m=0}^{\infty} s_m \beta^*_{n+1+\ell+m}\right\}_{\ell=0}^{\infty} \in \ell_{2+}^{d\times d}.$$

See Lemma 4.2 in [16]. In particular, for $n \in \mathbb{N}$, $u \in \{1, \ldots, n\}$ and $k \in \mathbb{N}$, we can define the sequences $\{b^k_{n,u,\ell}\}_{\ell=0}^{\infty} \in \ell_{2+}^{d\times d}$ and $\{\tilde{b}^k_{n,u,\ell}\}_{\ell=0}^{\infty} \in \ell_{2+}^{d\times d}$ by the recursions (2.7) and (2.8), respectively.

By (1.2) and (1.3), $\{X_k\}$ has the dual process $\{X'_k : k \in \mathbb{Z}\}$, which is a $\mathbb{C}^d$-valued, centered, weakly stationary process characterized by the biorthogonality relation

$$\langle X_s, X'_t \rangle = \delta_{st} I_d, \quad s, t \in \mathbb{Z}$$

(see [19]). Recall $\{a_k\} \in \ell_{2+}^{d\times d}$ and $\{\tilde{a}_k\} \in \ell_{2+}^{d\times d}$ from (2.3) and (2.5), respectively. The dual process $\{X'_k\}$ admits the following two MA representations (see Sect. 5 in [16]):

$$X'_n = -\sum_{\ell=n}^{\infty} a^*_{\ell-n}\varepsilon_\ell, \quad n \in \mathbb{Z}, \tag{3.5}$$

$$X'_n = -\sum_{\ell=-n}^{\infty} \tilde{a}^*_{\ell+n}\tilde{\varepsilon}_\ell, \quad n \in \mathbb{Z}. \tag{3.6}$$

The next theorem is the key to the proof of Theorem 2.1.

Theorem 3.1 *Assume* (1.2) *and* (1.3). *Then, for $n \in \mathbb{N}$ and $s, t \in \{1, \dots, n\}$, we have* (1.5).

Proof Fix $n \in \mathbb{N}$. For $s \in \{1, \dots, n\}$, we can write $P_{[1,n]}X'_s = \sum_{k=1}^n q_{s,k}X_k$ for some $q_{s,k} \in \mathbb{C}^{d\times d}$, $k \in \{1, \dots, n\}$. For $s, t \in \{1, \dots, n\}$, we have

$$\begin{aligned}\delta_{st}I_d &= \langle X'_s, X_t\rangle = \langle X'_s, P_{[1,n]}X_t\rangle = \langle P_{[1,n]}X'_s, X_t\rangle = \left\langle \sum_{k=1}^n q_{s,k}X_k, X_t \right\rangle \\ &= \sum_{k=1}^n q_{s,k}\langle X_k, X_t\rangle = \sum_{k=1}^n q_{s,k}\gamma(k-t),\end{aligned}$$

or $Q_nT_n(w) = I_{dn}$, where $Q_n := (q_{s,k})_{1\le s,k\le n} \in \mathbb{C}^{dn\times dn}$. Therefore, we have $Q_n = T_n(w)^{-1}$. However,

$$\langle X'_s, P_{[1,n]}X'_t\rangle = \langle P_{[1,n]}X'_s, X'_t\rangle = \left\langle \sum_{k=1}^n q_{s,k}X_k, X'_t \right\rangle = \sum_{k=1}^n q_{s,k}\langle X_k, X'_t\rangle = q_{s,t}.$$

Thus, the theorem follows. □

Lemma 3.1 *Assume* (1.2) *and* (1.3). *Then, for $n \in \mathbb{N}$ and $s, t \in \{1, \dots, n\}$, the following two equalities hold:*

$$\langle X'_s, P_{[1,n]}X'_t\rangle = \sum_{\ell=s\vee t}^{n} a^*_{\ell-s}a_{\ell-t} + \sum_{u=t}^{n}\langle X'_s, P^{\perp}_{[1,n]}\varepsilon_u\rangle a_{u-t}, \tag{3.7}$$

$$\langle X'_s, P_{[1,n]}X'_t\rangle = \sum_{\ell=1}^{s\wedge t} \tilde{a}^*_{s-\ell}\tilde{a}_{t-\ell} + \sum_{u=1}^{t}\langle X'_s, P^{\perp}_{[1,n]}\tilde{\varepsilon}_{-u}\rangle \tilde{a}_{t-u}. \tag{3.8}$$

Proof First, we prove (3.7). Since $V^X_{[1,n]} \subset V^X_{(-\infty,n]}$, we have

$$\begin{aligned}\langle X'_s, P_{[1,n]} X'_t\rangle &= \langle X'_s, P_{[1,n]} P_{(-\infty,n]} X'_t\rangle \\ &= \langle X'_s, P_{(-\infty,n]} X'_t\rangle - \langle X'_s, P^{\perp}_{[1,n]} P_{(-\infty,n]} X'_t\rangle.\end{aligned}$$

On the other hand, from (3.5), we have $P_{(-\infty,n]} X'_t = -\sum_{m=t}^{n} a^*_{m-t}\varepsilon_m$, hence

$$\langle X'_s, P_{(-\infty,n]} X'_t\rangle = \left\langle \sum_{\ell=s}^{\infty} a^*_{\ell-s}\varepsilon_\ell, \sum_{m=t}^{n} a^*_{m-t}\varepsilon_m \right\rangle = \sum_{\ell=s\vee t}^{n} a^*_{\ell-s} a_{\ell-t},$$

and $\langle X'_s, P^{\perp}_{[1,n]} P_{(-\infty,n]} X'_t\rangle$ is equal to

$$-\left\langle X'_s, P^{\perp}_{[1,n]}\left(\sum_{u=t}^{n} a^*_{u-t}\varepsilon_u\right)\right\rangle = -\sum_{u=t}^{n}\langle X'_s, P^{\perp}_{[1,n]}\varepsilon_u\rangle a_{u-t}.$$

Combining, we obtain (3.7).

Next, we prove (3.8). Since $V^X_{[1,n]} \subset V^X_{[1,\infty)}$, we have

$$\begin{aligned}\langle X'_s, P_{[1,n]} X'_t\rangle &= \langle X'_s, P_{[1,n]} P_{[1,\infty)} X'_t\rangle \\ &= \langle X'_s, P_{[1,\infty)} X'_t\rangle - \langle X'_s, P^{\perp}_{[1,n]} P_{[1,\infty)} X'_t\rangle.\end{aligned}$$

On the other hand, from (3.6), we have $P_{[1,\infty)} X'_t = -\sum_{m=1}^{t} \tilde{a}^*_{t-m}\tilde{\varepsilon}_{-m}$, hence

$$\langle X'_s, P_{[1,\infty)} X'_t\rangle = \left\langle \sum_{\ell=-\infty}^{s} \tilde{a}^*_{s-\ell}\tilde{\varepsilon}_{-\ell}, \sum_{m=1}^{t} \tilde{a}^*_{t-m}\tilde{\varepsilon}_{-m} \right\rangle = \sum_{\ell=1}^{s\wedge t} \tilde{a}^*_{s-\ell}\tilde{a}_{t-\ell},$$

and $\langle X'_s, P^{\perp}_{[1,n]} P_{[1,\infty)} X'_t\rangle$ is equal to

$$-\left\langle X'_s, P^{\perp}_{[1,n]}\left(\sum_{u=1}^{t} \tilde{a}^*_{t-u}\tilde{\varepsilon}_{-u}\right)\right\rangle = -\sum_{u=1}^{t}\langle X'_s, P^{\perp}_{[1,n]}\tilde{\varepsilon}_{-u}\rangle \tilde{a}_{t-u}.$$

Combining, we obtain (3.8). □

For $n \in \mathbb{N}$ and $u \in \{1, \ldots, n\}$, we define the sequence $\{W^k_{n,u}\}_{k=1}^{\infty}$ in V^d by

$$\begin{aligned}W^{2k-1}_{n,u} &= -P^{\perp}_{[1,\infty)}(P^{\perp}_{(-\infty,n]} P^{\perp}_{[1,\infty)})^{k-1}\varepsilon_u, \quad k \in \mathbb{N},\\ W^{2k}_{n,u} &= (P^{\perp}_{(-\infty,n]} P^{\perp}_{[1,\infty)})^{k}\varepsilon_u, \quad k \in \mathbb{N}.\end{aligned}$$

Lemma 3.2 *We assume (1.2) and (1.3). Then, for $n \in \mathbb{N}$ and $u \in \{1, \ldots, n\}$, we have*

$$P^{\perp}_{[1,n]}\varepsilon_u = -\sum_{k=1}^{\infty} W^k_{n,u}, \tag{3.9}$$

the sum converging strongly in V^d.

Proof Since ε_u is in $V^X_{(-\infty,n]}$, (3.9) follows from (3.1) and Theorem 3.2 in [16]. □

Proposition 3.1 *We assume* (1.2) *and* (1.3)*. Then, for* $n \in \mathbb{N}$*,* $u \in \{1, \ldots, n\}$ *and* $k \in \mathbb{N}$*, we have*

$$W^{2k-1}_{n,u} = \sum_{\ell=0}^{\infty} b^{2k-1}_{n,u,\ell} \tilde{\varepsilon}_\ell, \tag{3.10}$$

$$W^{2k}_{n,u} = \sum_{\ell=0}^{\infty} b^{2k}_{n,u,\ell} \varepsilon_{n+1+\ell}. \tag{3.11}$$

Proof Note that, from the definition of $W^k_{n,u}$,

$$W^{2k+1}_{n,u} = -P^{\perp}_{[1,\infty)} W^{2k}_{n,u}, \quad W^{2k+2}_{n,u} = -P^{\perp}_{(-\infty,n]} W^{2k+1}_{n,u}.$$

We prove (3.10) and (3.11) by induction. First, by (3.2), we have

$$W^1_{n,u} = -P^{\perp}_{[1,\infty)} \varepsilon_u = -\sum_{\ell=0}^{\infty} \langle \varepsilon_u, \tilde{\varepsilon}_\ell \rangle \tilde{\varepsilon}_\ell = \sum_{\ell=0}^{\infty} \beta_{u+\ell} \tilde{\varepsilon}_\ell = \sum_{\ell=0}^{\infty} b^1_{n,u,\ell} \tilde{\varepsilon}_\ell.$$

For $k \in \mathbb{N}$, assume that $W^{2k-1}_{n,u} = \sum_{\ell=0}^{\infty} b^{2k-1}_{n,u,\ell} \tilde{\varepsilon}_\ell$. Then, by (3.4),

$$\begin{aligned} W^{2k}_{n,u} &= -P^{\perp}_{(-\infty,n]} \left(\sum_{\ell=0}^{\infty} b^{2k-1}_{n,u,\ell} \tilde{\varepsilon}_\ell \right) = \sum_{\ell=0}^{\infty} \left(\sum_{m=0}^{\infty} b^{2k-1}_{n,u,m} \beta^*_{n+1+m+\ell} \right) \varepsilon_{n+1+\ell} \\ &= \sum_{\ell=0}^{\infty} b^{2k}_{n,u,\ell} \varepsilon_{n+1+\ell}, \end{aligned}$$

and, by (3.3),

$$\begin{aligned} W^{2k+1}_{n,u} &= -P^{\perp}_{[1,\infty)} \left(\sum_{\ell=0}^{\infty} b^{2k}_{n,u,\ell} \varepsilon_{n+1+\ell} \right) = \sum_{\ell=0}^{\infty} \left(\sum_{m=0}^{\infty} b^{2k}_{n,u,m} \beta_{n+1+m+\ell} \right) \tilde{\varepsilon}_\ell \\ &= \sum_{\ell=0}^{\infty} b^{2k+1}_{n,u,\ell} \tilde{\varepsilon}_\ell. \end{aligned}$$

Thus (3.10) and (3.11) follow. □

For $n \in \mathbb{N}$ and $u \in \{1, \ldots, n\}$, we define the sequence $\{\tilde{W}^k_{n,u}\}_{k=1}^{\infty}$ in V^d by

$$\begin{aligned} \tilde{W}^{2k-1}_{n,u} &= -P^{\perp}_{(-\infty,n]} (P^{\perp}_{[1,\infty)} P^{\perp}_{(-\infty,n]})^{k-1} \tilde{\varepsilon}_{-u}, \quad k \in \mathbb{N}, \\ \tilde{W}^{2k}_{n,u} &= (P^{\perp}_{[1,\infty)} P^{\perp}_{(-\infty,n]})^{k} \tilde{\varepsilon}_{-u}, \quad k \in \mathbb{N}. \end{aligned}$$

Lemma 3.3 *We assume* (1.2) *and* (1.3). *Then, for* $n \in \mathbb{N}$ *and* $u \in \{1, \ldots, n\}$, *we have*

$$P_{[1,n]}^{\perp}\tilde{\varepsilon}_{-u} = -\sum_{k=1}^{\infty} \tilde{W}_{n,u}^{k}, \tag{3.12}$$

the sum converging strongly in V^d.

Proof Since $\tilde{\varepsilon}_{-u}$ is in $V_{[1,\infty)}^{X}$, (3.12) follows from (3.1) and Theorem 3.2 in [16]. □

Proposition 3.2 *We assume* (1.2) *and* (1.3). *Then, for* $n \in \mathbb{N}$, $u \in \{1, \ldots, n\}$ *and* $k \in \mathbb{N}$, *we have*

$$\tilde{W}_{n,u}^{2k-1} = \sum_{\ell=0}^{\infty} \tilde{b}_{n,u,\ell}^{2k-1} \varepsilon_{n+1+\ell}, \tag{3.13}$$

$$\tilde{W}_{n,u}^{2k} = \sum_{\ell=0}^{\infty} \tilde{b}_{n,u,\ell}^{2k} \tilde{\varepsilon}_{\ell}. \tag{3.14}$$

Proof Note that, from the definition of $\tilde{W}_{n,u}^{k}$,

$$\tilde{W}_{n,u}^{2k+1} = -P_{(-\infty,n]}^{\perp}\tilde{W}_{n,u}^{2k}, \quad \tilde{W}_{n,u}^{2k+2} = -P_{[1,\infty)}^{\perp}\tilde{W}_{n,u}^{2k+1}.$$

We prove (3.13) and (3.14) by induction. First, by (3.2), we have

$$\begin{aligned}\tilde{W}_{n,u}^{1} &= -P_{(-\infty,n]}^{\perp}\tilde{\varepsilon}_{-u} = -\sum_{\ell=0}^{\infty} \langle \tilde{\varepsilon}_{-u}, \varepsilon_{n+1+\ell}\rangle \varepsilon_{n+1+\ell} \\ &= \sum_{\ell=0}^{\infty} \beta_{n+1-u+\ell}^{*} \varepsilon_{n+1+\ell} = \sum_{\ell=0}^{\infty} \tilde{b}_{n,u,\ell}^{1} \varepsilon_{n+1+\ell}.\end{aligned}$$

For $k \in \mathbb{N}$, assume that $\tilde{W}_{n,u}^{2k-1} = \sum_{\ell=0}^{\infty} \tilde{b}_{n,u,\ell}^{2k-1} \varepsilon_{n+1+\ell}$. Then, by (3.3),

$$\begin{aligned}\tilde{W}_{n,u}^{2k} &= -P_{[1,\infty)}^{\perp}\left(\sum_{\ell=0}^{\infty} \tilde{b}_{n,u,\ell}^{2k-1} \varepsilon_{n+1+\ell}\right) = \sum_{\ell=0}^{\infty}\left(\sum_{m=0}^{\infty} \tilde{b}_{n,u,m}^{2k-1} \beta_{n+1+m+\ell}\right)\tilde{\varepsilon}_{l} \\ &= \sum_{\ell=0}^{\infty} \tilde{b}_{n,u,\ell}^{2k} \tilde{\varepsilon}_{\ell},\end{aligned}$$

and, by (3.4),

$$\begin{aligned}\tilde{W}_{n,u}^{2k+1} &= -P_{(\infty,n]}^{\perp}\left(\sum_{\ell=0}^{\infty} \tilde{b}_{n,u,\ell}^{2k} \tilde{\varepsilon}_{\ell}\right) = \sum_{\ell=0}^{\infty}\left(\sum_{m=0}^{\infty} \tilde{b}_{n,u,m}^{2k} \beta_{n+1+m+\ell}^{*}\right)\varepsilon_{n+1+\ell} \\ &= \sum_{\ell=0}^{\infty} \tilde{b}_{n,u,\ell}^{2k+1} \varepsilon_{n+1+\ell}.\end{aligned}$$

Thus (3.13) and (3.14) follow. □

We are ready to prove Theorem 2.1.

Proof (i) For $n \in \mathbb{N}$, $s, u \in \{1, \dots, n\}$ and $k \in \mathbb{N}$, we see from (3.5) and (3.13) that

$$\langle X_s', \tilde{W}_{n,u}^{2k-1} \rangle = -\sum_{\ell=0}^{\infty} a_{n+1-s+\ell}^* (\tilde{b}_{n,u,\ell}^{2k-1})^*,$$

and from (3.6) and (3.14) that

$$\langle X_s', \tilde{W}_{n,u}^{2k} \rangle = -\sum_{\ell=0}^{\infty} \tilde{a}_{s+\ell}^* (\tilde{b}_{n,u,\ell}^{2k})^*.$$

Therefore, by Lemma 3.3, $\langle X_s', P_{[1,n]}^{\perp} \tilde{\varepsilon}_{-u} \rangle$ is equal to

$$-\sum_{k=1}^{\infty} \langle X_s', \tilde{W}_{n,u}^{k} \rangle = \sum_{k=1}^{\infty} \left\{ \sum_{\ell=0}^{\infty} \tilde{b}_{n,u,\ell}^{2k-1} a_{n+1-s+\ell} + \sum_{\ell=0}^{\infty} \tilde{b}_{n,u,\ell}^{2k} \tilde{a}_{s+\ell} \right\}^* .$$

The assertion (i) follows from this, Theorem 3.1 and Lemma 3.1.

(ii) For $n \in \mathbb{N}$, $s, u \in \{1, \dots, n\}$ and $k \in \mathbb{N}$, we see from (3.6) and (3.10) that

$$\langle X_s', W_{n,u}^{2k-1} \rangle = -\sum_{\ell=0}^{\infty} \tilde{a}_{s+\ell}^* (b_{n,u,\ell}^{2k-1})^*,$$

and from (3.5) and (3.11) that

$$\langle X_s', W_{n,u}^{2k} \rangle = -\sum_{\ell=0}^{\infty} a_{n+1-s+\ell}^* (b_{n,u,\ell}^{2k})^*.$$

Therefore, by Lemma 3.2, $\langle X_s', P_{[1,n]}^{\perp} \varepsilon_u \rangle$ is equal to

$$-\sum_{k=1}^{\infty} \langle X_s', W_{n,u}^{k} \rangle = \sum_{k=1}^{\infty} \left\{ \sum_{\ell=0}^{\infty} b_{n,u,\ell}^{2k-1} \tilde{a}_{s+\ell} + \sum_{\ell=0}^{\infty} b_{n,u,\ell}^{2k} a_{n+1-s+\ell} \right\}^* .$$

The assertion (ii) follows from this, Theorem 3.1 and Lemma 3.1. □

4 Strong convergence result for Toeplitz systems

In this section, we use Theorem 2.1 to show a strong convergence result for solutions of block Toeplitz systems. We assume (1.2) and (1.6). Then w is continuous on $\mathbb{T}$ since $w(e^{i\theta}) = (2\pi)^{-1} \sum_{k \in \mathbb{Z}} e^{ik\theta} \gamma(k)$. In particular, (1.3) is also satisfied. The conditions

(1.2) and (1.6) also imply that all of $\{a_k\}$, $\{c_k\}$, $\{\tilde{a}_k\}$ and $\{\tilde{c}_k\}$ belong to $\ell_{1+}^{d\times d}$. See Theorem 3.3 and (3.3) in [17]; see also Theorem 4.1 in [12]. In particular, we have $h(e^{i\theta})^{-1} = -\sum_{k=0}^{\infty} e^{ik\theta} a_k$ and $h_\sharp(e^{i\theta}) = \tilde{h}(e^{-i\theta})^* = \sum_{k=0}^{\infty} e^{ik\theta} \tilde{c}_k^*$, hence, by (2.6),

$$\beta_k = \sum_{j=0}^{\infty} a_{j+k}\tilde{c}_j, \quad k \in \mathbb{N} \cup \{0\}. \tag{4.1}$$

Under (1.2) and (1.6), we define

$$F(n) := \left(\sum_{j=0}^{\infty} \|\tilde{c}_j\|\right) \sum_{\ell=n}^{\infty} \|a_\ell\|, \quad n \in \mathbb{N} \cup \{0\}.$$

Then $F(n)$ decreases to zero as $n \to \infty$.

We need the next lemma in the proof of Theorem 4.1 below.

Lemma 4.1 *Assume* (1.2) *and* (1.6)*. Then, for* $n, k \in \mathbb{N}$ *and* $u \in \{1, \ldots, n\}$*, we have*

$$\sum_{\ell=0}^{\infty} \|\tilde{b}_{n,u,\ell}^k\| \le F(n+1)^{k-1} F(n+1-u). \tag{4.2}$$

Proof For $m \in \mathbb{N}$, we see from (4.1) that

$$\sum_{\ell=0}^{\infty} \|\beta_{m+\ell}\| \le \sum_{j=0}^{\infty} \|\tilde{c}_j\| \sum_{\ell=0}^{\infty} \|a_{m+j+\ell}\| \le \sum_{j=0}^{\infty} \|\tilde{c}_j\| \sum_{\ell=m}^{\infty} \|a_\ell\|,$$

hence

$$\sum_{\ell=0}^{\infty} \|\beta_{m+\ell}\| \le F(m). \tag{4.3}$$

Let $n \in \mathbb{N}$ and $u \in \{1, \ldots, n\}$. We use induction on k to prove (4.2). Since $\tilde{b}_{n,u,\ell}^1 = \beta_{n+1-u+\ell}^*$, we see from (4.3) that

$$\sum_{\ell=0}^{\infty} \|\tilde{b}_{n,u,\ell}^1\| = \sum_{\ell=0}^{\infty} \|\beta_{n+1-u+\ell}\| \le F(n+1-u).$$

We assume (4.2) for $k \in \mathbb{N}$. Then, again by (4.3),

$$\begin{aligned}\sum_{\ell=0}^{\infty} \|\tilde{b}_{n,u,\ell}^{k+1}\| &\le \sum_{m=0}^{\infty} \|\tilde{b}_{n,u,m}^k\| \sum_{\ell=0}^{\infty} \|\beta_{n+1+m+\ell}\| \\ &\le F(n+1) \sum_{m=0}^{\infty} \|\tilde{b}_{n,u,m}^k\| \le F(n+1)^k F(n+1-u).\end{aligned}$$

Thus (4.2) with k replaced by $k+1$ also holds. □

For $\{y_k\}_{k=1}^{\infty} \in \ell_1^{d\times d}(\mathbb{N})$, the solution Z_∞ to (1.11) with (1.12) and (1.13) is given by (1.10) with

$$z_s = \sum_{t=1}^{\infty} \sum_{\ell=1}^{s\wedge t} \tilde{a}_{s-\ell}^* \tilde{a}_{t-\ell} y_t \in \mathbb{C}^{d\times d}, \quad s \in \mathbb{N} \tag{4.4}$$

(see Remark 2.1 in Sect. 2). Notice that the sum in (4.4) converges absolutely.

Theorem 4.1 *We assume* (1.2) *and* (1.6). *Let* $\{y_k\}_{k=1}^{\infty} \in \ell_1^{d\times d}(\mathbb{N})$. *Then, for* Z_n *in* (1.7)–(1.9) *and* Z_∞ *in* (1.10)–(1.13), *we have* (1.14).

Proof By Theorem 2.1 (i), we have

$$\begin{aligned} z_{n,s} = &\sum_{t=1}^{n} \sum_{\ell=1}^{s\wedge t} \tilde{a}_{s-\ell}^* \tilde{a}_{t-\ell} y_t + \sum_{t=1}^{n} \sum_{u=1}^{t} \sum_{\ell=0}^{\infty} a_{n+1-s+\ell}^* \beta_{n+1-u+\ell} \tilde{a}_{t-u} y_t \\ &+ \sum_{t=1}^{n} \sum_{u=1}^{t} \sum_{k=1}^{\infty} \left\{ \sum_{\ell=0}^{\infty} \tilde{b}_{n,u,\ell}^{2k+1} a_{n+1-s+\ell} + \sum_{\ell=0}^{\infty} \tilde{b}_{n,u,\ell}^{2k} \tilde{a}_{s+\ell} \right\}^* \tilde{a}_{t-u} y_t, \end{aligned}$$

hence, by (4.4), $\sum_{s=1}^{n} \|z_{n,s} - z_s\| \le S_1(n) + S_2(n) + S_3(n) + S_4(n)$, where

$$S_1(n) := \sum_{t=n+1}^{\infty} \sum_{s=1}^{n} \sum_{\ell=1}^{s} \|\tilde{a}_{s-\ell}\| \|\tilde{a}_{t-\ell}\| \|y_t\|,$$

$$S_2(n) := \sum_{s=1}^{n} \sum_{t=1}^{n} \sum_{u=1}^{t} \sum_{\ell=0}^{\infty} \|a_{n+1-s+\ell}\| \|\beta_{n+1-u+\ell}\| \|\tilde{a}_{t-u}\| \|y_t\|,$$

$$S_3(n) := \sum_{s=1}^{n} \sum_{t=1}^{n} \sum_{u=1}^{t} \sum_{k=1}^{\infty} \sum_{\ell=0}^{\infty} \|\tilde{b}_{n,u,\ell}^{2k+1}\| \|a_{n+1-s+\ell}\| \|\tilde{a}_{t-u}\| \|y_t\|$$

and

$$S_4(n) = \sum_{s=1}^{n} \sum_{t=1}^{n} \sum_{u=1}^{t} \sum_{k=1}^{\infty} \sum_{\ell=0}^{\infty} \|\tilde{b}_{n,u,\ell}^{2k}\| \|\tilde{a}_{s+\ell}\| \|\tilde{a}_{t-u}\| \|y_t\|.$$

By the change of variables $m = s - \ell + 1$, we have

$$\begin{aligned} S_1(n) &= \sum_{t=n+1}^{\infty} \sum_{s=1}^{n} \sum_{m=1}^{s} \|\tilde{a}_{m-1}\| \|\tilde{a}_{t+m-s-1}\| \|y_t\| \\ &= \sum_{t=n+1}^{\infty} \|y_t\| \sum_{m=1}^{n} \|\tilde{a}_{m-1}\| \sum_{s=m}^{n} \|\tilde{a}_{t+m-s-1}\| \\ &\le \left(\sum_{k=0}^{\infty} \|\tilde{a}_k\| \right)^2 \sum_{t=n+1}^{\infty} \|y_t\| \quad \to \quad 0, \quad n \to \infty. \end{aligned}$$

By (4.2) with $k = 1$ or (4.3), we have

$$\begin{aligned} S_2(n) &= \sum_{t=1}^{n} \sum_{u=1}^{t} \|\tilde{a}_{t-u}\| \|y_t\| \sum_{\ell=0}^{\infty} \|\beta_{n+1-u+\ell}\| \sum_{s=1}^{n} \|a_{n+1-s+\ell}\| \\ &\le \left(\sum_{s=1}^{\infty} \|a_s\| \right) \sum_{t=1}^{n} \sum_{u=1}^{t} \|\tilde{a}_{t-u}\| \|y_t\| F(n+1-u). \end{aligned}$$

Furthermore, by the change of variables $v = t - u + 1$, we obtain

$$\begin{aligned} \sum_{t=1}^{n} \sum_{u=1}^{t} \|\tilde{a}_{t-u}\| \|y_t\| F(n+1-u) &= \sum_{t=1}^{\infty} \sum_{u=1}^{t} \|\tilde{a}_{t-u}\| \|y_t\| 1_{[0,n]}(t) F(n+1-u) \\ &= \sum_{t=1}^{\infty} \sum_{v=1}^{t} \|\tilde{a}_{v-1}\| \|y_t\| 1_{[0,n]}(t) F(n-t+v) \\ &\le \sum_{t=1}^{\infty} \sum_{v=1}^{\infty} \|\tilde{a}_{v-1}\| \|y_t\| 1_{[0,n]}(t) F(n-t+v). \end{aligned}$$

Since

$$\begin{aligned} &\lim_{n\to\infty} \|\tilde{a}_{v-1}\| \|y_t\| 1_{[0,n]}(t) F(n-t+v) = 0, \quad t, v \in \mathbb{N}, \\ &\|\tilde{a}_{v-1}\| \|y_t\| 1_{[0,n]}(t) F(n-t+v) \le F(1) \|\tilde{a}_{v-1}\| \|y_t\|, \quad t, v \in \mathbb{N}, \\ &\sum_{t=1}^{\infty} \sum_{v=1}^{\infty} \|\tilde{a}_{v-1}\| \|y_t\| < \infty, \end{aligned}$$

the dominated convergence theorem yields

$$\lim_{n\to\infty} \sum_{t=1}^{\infty} \sum_{v=1}^{\infty} \|\tilde{a}_{v-1}\| \|y_t\| 1_{[0,n]}(t) F(n-t+v) = 0,$$

hence $\lim_{n\to\infty} S_2(n) = 0$.

Choose $N \in \mathbb{N}$ such that $F(N+1) < 1$. Then, by Lemma 4.1, we have, for $n \geq N$,

$$\begin{aligned} S_3(n) &= \sum_{t=1}^{n} \sum_{u=1}^{t} \|\tilde{a}_{t-u}\| \|y_t\| \sum_{k=1}^{\infty} \sum_{\ell=0}^{\infty} \|\tilde{b}_{n,u,\ell}^{2k+1}\| \sum_{s=1}^{n} \|a_{n+1-s+\ell}\| \\ &\leq F(1) \left(\sum_{s=1}^{\infty} \|a_s\| \right) \sum_{t=1}^{n} \|y_t\| \sum_{u=1}^{t} \|\tilde{a}_{t-u}\| \sum_{k=1}^{\infty} F(n+1)^{2k} \\ &\leq F(1) \left(\sum_{s=1}^{\infty} \|a_s\| \right) \left(\sum_{u=0}^{\infty} \|\tilde{a}_u\| \right) \left(\sum_{t=1}^{\infty} \|y_t\| \right) \frac{F(n+1)^2}{1 - F(n+1)^2}. \end{aligned}$$

Thus $\lim_{n\to\infty} S_3(n) = 0$. Similarly, we have, for $n \geq N$,

$$S_4(n) \leq F(1) \left(\sum_{s=0}^{\infty} \|\tilde{a}_s\| \right)^2 \left(\sum_{t=1}^{\infty} \|y_t\| \right) \frac{F(n+1)}{1 - F(n+1)^2},$$

hence $\lim_{n\to\infty} S_4(n) = 0$.

Combining, we obtain (1.14). □

5 Closed-form formulas

In this section, we use Theorem 2.1 to derive closed-form formulas for $T_n(w)^{-1}$ with rational symbol w that corresponds to a d-variate ARMA process. We assume that the symbol w of $T_n(w)$ is of the form (1.17) with $h : \mathbb{T} \to \mathbb{C}^{d\times d}$ satisfying (1.18). Then h is an outer function in $H_2^{d\times d}(\mathbb{T})$, and another outer function $h_\sharp \in H_2^{d\times d}(\mathbb{T})$ that appears in (1.4) also satisfies (1.18); see Sect. 6.2 in [16]. Notice that (1.17) with (1.18) implies (1.2) and (1.3).

We can write $h(z)^{-1}$ in the form

$$h(z)^{-1} = -\rho_{0,0} - \sum_{\mu=1}^{K} \sum_{j=1}^{m_\mu} \frac{1}{(1 - \overline{p}_\mu z)^j} \rho_{\mu,j} - \sum_{j=1}^{m_0} z^j \rho_{0,j}, \tag{5.1}$$

where

$$\begin{cases} K \in \mathbb{N} \cup \{0\}, \quad m_\mu \in \mathbb{N}, \quad \mu \in \{1, \ldots, K\}, \quad m_0 \in \mathbb{N} \cup \{0\}, \\ p_\mu \in \mathbb{D} \setminus \{0\}, \quad \mu \in \{1, \ldots, K\}, \quad p_\mu \neq p_\nu, \quad \mu \neq \nu, \\ \rho_{\mu,j} \in \mathbb{C}^{d\times d}, \quad \mu \in \{0, \ldots, K\}, \ j \in \{1, \ldots, m_\mu\}, \quad \rho_{0,0} \in \mathbb{C}^{d\times d}, \\ \rho_{\mu,m_\mu} \neq 0, \quad \mu \in \{1, \ldots, K\}, \\ \rho_{0,m_0} \neq 0 \quad \text{if} \ m_0 \geq 1. \end{cases} \tag{5.2}$$

Here the convention $\sum_{k=1}^{0} = 0$ is adopted in the sums on the right-hand side of (5.1). For example, if $m_0 = 0$, then

$$h(z)^{-1} = -\rho_{0,0} - \sum_{\mu=1}^{K} \sum_{j=1}^{m_\mu} \frac{1}{(1 - \overline{p}_\mu z)^j} \rho_{\mu,j},$$

while, if $K = 0$, then

$$h(z)^{-1} = -\rho_{0,0} - \sum_{j=1}^{m_0} z^j \rho_{0,j} \tag{5.3}$$

and the corresponding stationary process $\{X_k\}$ is a d-variate AR(m_0) process.

Remark 5.1 It should be noticed that the expression (5.1) with (5.2) is uniquely determined, up to a constant unitary factor, from $\{X_k\}$ satisfying (1.17) with (1.18) since so is h in the factorization (1.17) with (1.18) (see Sect. 2). Suppose that we start with a d-variate, causal and invertible ARMA process $\{X_k\}$ in the sense of [4], that is, a $\mathbb{C}^d$-valued, centered, weakly stationary process described by the ARMA equation

$$\Phi(B)X_n = \Psi(B)\xi_n, \quad n \in \mathbb{Z},$$

where, for $r, s \in \mathbb{N} \cup \{0\}$ and $\Phi_i, \Psi_j \in \mathbb{C}^{d \times d}$ $(i = 1, \ldots, r,\ j = 1, \ldots, s)$,

$$\Phi(z) = I_d - z\Phi_1 - \cdots - z^r \Phi_r \quad \text{and} \quad \Psi(z) = I_d - z\Psi_1 - \cdots - z^s \Psi_s$$

are $\mathbb{C}^{d \times d}$-valued polynomials satisfying $\det \Phi(z) \neq 0$ and $\det \Psi(z) \neq 0$ on $\overline{\mathbb{D}}$, B is the backward shift operator defined by $BX_m = X_{m-1}$, and $\{\xi_k : k \in \mathbb{Z}\}$ is a d-variate white noise, that is, a d-variate, centered process such that $E[\xi_n \xi_m^*] = \delta_{nm} V$ for some positive-definite $V \in \mathbb{C}^{d \times d}$. Notice that the pair $(\Phi(z), \Psi(z))$ is not uniquely determined from $\{X_k\}$; for example, we can replace $(\Phi(z), \Psi(z))$ by $((2 - z)\Phi(z), (2 - z)\Psi(z))$. However, if we put $h(z) = \Phi(z)^{-1}\Psi(z)V^{1/2}$, then h is an outer function belonging to $H_2^{d \times d}(\mathbb{T})$ and satisfies (1.17) for the spectral density w of $\{X_k\}$. Therefore, h is uniquely determined, up to a constant unitary factor, from $\{X_k\}$. In particular, the expression (5.1) with (5.2) for h is also uniquely determined, up to a constant unitary factor, from $\{X_k\}$. From these observations and the results in [13] and this paper, we are led to the idea of parameterizing the ARMA processes by the expression (5.1) with (5.2) (see Remark 8 in [13]). This point will be discussed in future work.

By Theorem 2 in [13], $h_\sharp^{-1}$ has the same m_0 and the same poles with the same multiplicities as h^{-1}, that is, for m_0, K and $(p_1, m_1), \ldots, (p_K, m_K)$ in (5.1) with (5.2), $h_\sharp^{-1}$ has the form

$$h_\sharp(z)^{-1} = -\rho_{0,0}^\sharp - \sum_{\mu=1}^{K} \sum_{j=1}^{m_\mu} \frac{1}{(1 - \overline{p}_\mu z)^j} \rho_{\mu,j}^\sharp - \sum_{j=1}^{m_0} z^j \rho_{0,j}^\sharp, \tag{5.4}$$

where

$$\begin{cases} \rho^{\sharp}_{\mu,j} \in \mathbb{C}^{d\times d}, \quad \mu \in \{0,\dots,K\},\ j \in \{1,\dots,m_\mu\}, \quad \rho^{\sharp}_{0,0} \in \mathbb{C}^{d\times d}, \\ \rho^{\sharp}_{\mu,m_\mu} \neq 0, \quad \mu \in \{1,\dots,K\}, \\ \rho^{\sharp}_{0,m_0} \neq 0 \quad \text{if} \ \ m_0 \geq 1. \end{cases}$$

Notice that if $d = 1$, then we can take $h_\sharp = h$, hence $\rho_{0,0} = \rho^{\sharp}_{0,0}$ and $\rho_{\mu,j} = \rho^{\sharp}_{\mu,j}$ for $\mu \in \{1,\dots,K\}$ and $j \in \{1,\dots,m_\mu\}$.

Recall $\tilde{h}$ from (2.1). From (5.4), we have

$$\tilde{h}(z)^{-1} = -\tilde{\rho}_{0,0} - \sum_{\mu=1}^{K}\sum_{j=1}^{m_\mu} \frac{1}{(1-p_\mu z)^j}\tilde{\rho}_{\mu,j} - \sum_{j=1}^{m_0} z^j \tilde{\rho}_{0,j},$$

where

$$\tilde{\rho}_{0,0} := (\rho^{\sharp}_{0,0})^*, \quad \tilde{\rho}_{\mu,j} := (\rho^{\sharp}_{\mu,j})^*, \quad \mu \in \{0,\dots,K\},\ j \in \{1,\dots,m_\mu\}.$$

Recall the sequences $\{a_k\}$ and $\{\tilde{a}_k\}$ from (2.3) and (2.5), respectively. We have

$$a_n = \sum_{\mu=1}^{K}\sum_{j=1}^{m_\mu} \binom{n+j-1}{j-1} \overline{p}_\mu^n \rho_{\mu,j}, \quad n \geq m_0 + 1, \tag{5.5}$$

$$\tilde{a}_n = \sum_{\mu=1}^{K}\sum_{j=1}^{m_\mu} \binom{n+j-1}{j-1} p_\mu^n \tilde{\rho}_{\mu,j}, \quad n \geq m_0 + 1 \tag{5.6}$$

and

$$a_n = \rho_{0,n} + \sum_{\mu=1}^{K}\sum_{j=1}^{m_\mu} \binom{n+j-1}{j-1} \overline{p}_\mu^n \rho_{\mu,j}, \quad n \in \{0,\dots,m_0\}, \tag{5.7}$$

$$\tilde{a}_n = \tilde{\rho}_{0,n} + \sum_{\mu=1}^{K}\sum_{j=1}^{m_\mu} \binom{n+j-1}{j-1} p_\mu^n \tilde{\rho}_{\mu,j}, \quad n \in \{0,\dots,m_0\}, \tag{5.8}$$

where the convention $\binom{0}{0} = 1$ is adopted; see Proposition 4 in [13].

We first consider the case of $K = 0$ that corresponds to a d-variate AR(m_0) process. As can be seen from the following theorem, in this case, we have simple closed-form formulas for $T_n(w)^{-1}$.

Theorem 5.1 *We assume* (1.17), (1.18) *and* $K = 0$ *for* K *in* (5.1). *Thus we assume* (5.3). *Then the following four assertions hold.*

(i) For $n \ge m_0 + 1$, $s \in \{1, \dots, n\}$ and $t \in \{1, \dots, n - m_0\}$, we have

$$\left(T_n(w)^{-1}\right)^{s,t} = \sum_{\lambda=1}^{s\wedge t} \tilde{a}^*_{s-\lambda}\tilde{a}_{t-\lambda}. \tag{5.9}$$

(ii) For $n \ge m_0 + 1$, $s \in \{1, \dots, n - m_0\}$ and $t \in \{1, \dots, n\}$, we have

$$\left(T_n(w)^{-1}\right)^{s,t} = \sum_{\lambda=1}^{s\wedge t} \tilde{a}^*_{s-\lambda}\tilde{a}_{t-\lambda}. \tag{5.10}$$

(iii) For $n \ge m_0 + 1$, $s \in \{1, \dots, n\}$ and $t \in \{m_0 + 1, \dots, n\}$, we have

$$\left(T_n(w)^{-1}\right)^{s,t} = \sum_{\lambda=s\vee t}^{n} a^*_{\lambda-s}a_{\lambda-t}. \tag{5.11}$$

(iv) For $n \ge m_0 + 1$, $s \in \{m_0 + 1, \dots, n\}$ and $t \in \{1, \dots, n\}$, we have

$$\left(T_n(w)^{-1}\right)^{s,t} = \sum_{\lambda=s\vee t}^{n} a^*_{\lambda-s}a_{\lambda-t}. \tag{5.12}$$

Proof For w satisfying (1.17), (1.18) and $K = 0$, let $\{X_k\}$, $\{X'_k\}$, $\{\varepsilon_k\}$ and $\{\tilde{\varepsilon}_k\}$ be as in Sect. 3.

(i) By (5.4) with $K = 0$, we have $\tilde{a}_0 = \tilde{\rho}_0$, $\tilde{a}_k = \tilde{\rho}_{0,k}$ for $k \in \{1, \dots, m_0\}$ and $\tilde{a}_k = 0$ for $k \ge m_0 + 1$. In particular, we have $\sum_{k=0}^{m_0} \tilde{a}_k X_{u+k} + \tilde{\varepsilon}_{-u} = 0$ for $u \in \mathbb{Z}$; see (2.15) in [16]. This implies $\tilde{\varepsilon}_{-u} \in V^X_{[1,n]}$, or $P^{\perp}_{[1,n]}\tilde{\varepsilon}_{-u} = 0$, for $u \in \{1, \dots, n - m_0\}$. Therefore, (5.9) follows from Theorem 3.1 and (3.8).

(iii) By (5.3), we have $a_0 = \rho_{0,0}$, $a_k = \rho_{0,k}$ for $k \in \{1, \dots, m_0\}$ and $a_k = 0$ for $k \ge m_0 + 1$. In particular, $\sum_{k=0}^{m_0} a_k X_{u-k} + \varepsilon_u = 0$ for $u \in \mathbb{Z}$; see (2.15) in [16]. This implies $\varepsilon_u \in V^X_{[1,n]}$, or $P^{\perp}_{[1,n]}\varepsilon_u = 0$, for $u \in \{m_0 + 1, \dots, n\}$. Therefore, (5.11) follows from Theorem 3.1 and (3.7).

(ii), (iv) By (2.9), (ii) and (iv) follow from (i) and (iii), respectively. □

We turn to the case of $K \ge 1$. In what follows in this section, for K in (5.1), we assume

$$K \ge 1.$$

For $m_1, \dots, m_K$ in (5.1), we define $M \in \mathbb{N}$ by

$$M := \sum_{\mu=1}^{K} m_\mu. \tag{5.13}$$

For $\mu \in \{1, \dots, K\}$, p_μ in (5.1) and $i \in \mathbb{N}$, we define $p_{\mu,i} : \mathbb{Z} \to \mathbb{C}^{d\times d}$ by

$$p_{\mu,i}(k) := \binom{k}{i-1} p_\mu^{k-i+1} I_d, \quad k \in \mathbb{Z}. \tag{5.14}$$

Notice that

$$p_{\mu,i}(0) = \binom{0}{i-1} p_\mu^{-i+1} I_d = \delta_{i,1} I_d.$$

For $n \in \mathbb{Z}$, we also define $\mathbf{p}_n \in \mathbb{C}^{dM\times d}$ by the following block representation:

$$\begin{aligned}\mathbf{p}_n := (&p_{1,1}(n), \dots, p_{1,m_1}(n) \mid p_{2,1}(n), \dots, p_{2,m_2}(n) \mid \\ &\cdots \mid p_{K,1}(n), \dots, p_{K,m_K}(n))^\top.\end{aligned}$$

Notice that

$$\mathbf{p}_0 = (I_d, 0, \dots, 0 \mid I_d, 0, \dots, 0 | \cdots \mid I_d, 0, \dots, 0)^\top \in \mathbb{C}^{dM\times d}.$$

We define $\Lambda \in \mathbb{C}^{dM\times dM}$ by

$$\Lambda := \sum_{\ell=0}^{\infty} \mathbf{p}_\ell \mathbf{p}_\ell^*.$$

For $\mu, \nu \in \{1, 2, \dots, K\}$, we define $\Lambda^{\mu,\nu} \in \mathbb{C}^{dm_\mu \times dm_\nu}$ by the block representation

$$\Lambda^{\mu,\nu} := \begin{pmatrix} \lambda^{\mu,\nu}(1,1) & \lambda^{\mu,\nu}(1,2) & \cdots & \lambda^{\mu,\nu}(1,m_\nu) \\ \lambda^{\mu,\nu}(2,1) & \lambda^{\mu,\nu}(2,2) & \cdots & \lambda^{\mu,\nu}(2,m_\nu) \\ \vdots & \vdots & & \vdots \\ \lambda^{\mu,\nu}(m_\mu,1) & \lambda^{\mu,\nu}(m_\mu,2) & \cdots & \lambda^{\mu,\nu}(m_\mu,m_\nu) \end{pmatrix},$$

where, for $i \in \{1, \dots, m_\mu\}$ and $j \in \{1, \dots, m_\nu\}$,

$$\lambda^{\mu,\nu}(i,j) := \sum_{r=0}^{j-1} \binom{i-1}{r} \binom{i+j-r-2}{i-1} \frac{p_\mu^{j-r-1} \overline{p}_\nu^{i-r-1}}{(1-p_\mu \overline{p}_\nu)^{i+j-r-1}} I_d \in \mathbb{C}^{d\times d}.$$

Then, by Lemma 3 in [13], the matrix Λ has the following block representation:

$$\Lambda = \begin{pmatrix} \Lambda^{1,1} & \Lambda^{1,2} & \dots & \Lambda^{1,K} \\ \Lambda^{2,1} & \Lambda^{2,2} & \dots & \Lambda^{2,K} \\ \vdots & \vdots & \ddots & \vdots \\ \Lambda^{K,1} & \Lambda^{K,2} & \dots & \Lambda^{K,K} \end{pmatrix}.$$

We define, for $\mu \in \{1, \dots, K\}$ and $j \in \{1, \dots, m_\mu\}$,

$$\theta_{\mu,j} := -\lim_{z \to p_\mu} \frac{1}{(m_\mu - j)!} \frac{d^{m_\mu - j}}{dz^{m_\mu - j}} \left\{ (z - p_\mu)^{m_\mu} h_\sharp(z) h^\dagger(z)^{-1} \right\} \in \mathbb{C}^{d \times d}, \tag{5.15}$$

where

$$h^\dagger(z) := h(1/\overline{z})^*. \tag{5.16}$$

We define $\Theta \in \mathbb{C}^{dM \times dM}$ by the block representation

$$\Theta := \begin{pmatrix} \Theta_1 & 0 & \cdots & 0 \\ 0 & \Theta_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \Theta_K \end{pmatrix},$$

where, for $\mu \in \{1, \dots, K\}$, $\Theta_\mu \in \mathbb{C}^{dm_\mu \times dm_\mu}$ is defined by

$$\Theta_\mu := \begin{pmatrix} \theta_{\mu,1} & \theta_{\mu,2} & \cdots & \theta_{\mu,m_\mu - 1} & \theta_{\mu,m_\mu} \\ \theta_{\mu,2} & \theta_{\mu,3} & \cdots & \theta_{\mu,m_\mu} & \\ \vdots & \vdots & & & \\ \theta_{\mu,m_\mu - 1} & \theta_{\mu,m_\mu} & & & \\ \theta_{\mu,m_\mu} & & & & 0 \end{pmatrix}$$

using $\theta_{\mu,j}$ in (5.15) with (5.16).

For $n \in \mathbb{Z}$, we define $\Pi_n \in \mathbb{C}^{dM \times dM}$ by the block representation

$$\Pi_n := \begin{pmatrix} \Pi_{1,n} & 0 & \cdots & 0 \\ 0 & \Pi_{2,n} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \Pi_{K,n} \end{pmatrix},$$

where, for $\mu \in \{1, \dots, K\}$ and $n \in \mathbb{Z}$, $\Pi_{\mu,n} \in \mathbb{C}^{dm_\mu \times dm_\mu}$ is defined by

$$\Pi_{\mu,n} := \begin{pmatrix} p_{\mu,1}(n) & p_{\mu,2}(n) & p_{\mu,3}(n) & \cdots & p_{\mu,m_\mu}(n) \\ & p_{\mu,1}(n) & p_{\mu,2}(n) & \cdots & p_{\mu,m_\mu - 1}(n) \\ & & \ddots & \ddots & \vdots \\ & & & \ddots & p_{\mu,2}(n) \\ 0 & & & & p_{\mu,1}(n) \end{pmatrix}$$

using $p_{\mu,i}(n)$ in (5.14).

The next lemma slightly extends Lemma 17 in [13].

Lemma 5.1 *We assume* (1.17), (1.18) *and* $K \geq 1$ *for* K *in* (5.1). *Then, for* $n, k, \ell \in \mathbb{Z}$ *such that* $n + k + \ell \geq m_0$, *we have*

$$\beta^*_{n+k+\ell+1} = \mathbf{p}_\ell^\top \Pi_n \Theta \mathbf{p}_k,$$

hence

$$\beta_{n+k+\ell+1} = \mathbf{p}_k^* (\Pi_n \Theta)^* \overline{\mathbf{p}}_\ell.$$

The proof of Lemma 5.1 is almost the same as that of Lemma 17 in [13], hence we omit it.

For $n \in \mathbb{Z}$, we define $G_n, \tilde{G}_n \in \mathbb{C}^{dM \times dM}$ by

$$G_n := \Pi_n \Theta \Lambda, \quad \tilde{G}_n := (\Pi_n \Theta)^* \Lambda^\top.$$

Lemma 5.2 *We assume* (1.17), (1.18) *and* $K \geq 1$ *for* K *in* (5.1). *Then the following two assertions hold.*

(i) We assume $n \geq u \geq m_0 + 1$. *Then, for* $k \in \mathbb{N}$ *and* $\ell \in \mathbb{N} \cup \{0\}$, *we have*

$$b^{2k-1}_{n,u,\ell} = \mathbf{p}^*_{u-n-1} (\tilde{G}_n G_n)^{k-1} (\Pi_n \Theta)^* \overline{\mathbf{p}}_\ell, \tag{5.17}$$

$$b^{2k}_{n,u,\ell} = \mathbf{p}^*_{u-n-1} (\tilde{G}_n G_n)^{k-1} \tilde{G}_n \Pi_n \Theta \mathbf{p}_\ell. \tag{5.18}$$

(ii) We assume $1 \leq u \leq n - m_0$. *Then, for* $k \in \mathbb{N}$ *and* $\ell \in \mathbb{N} \cup \{0\}$, *we have*

$$\tilde{b}^{2k-1}_{n,u,\ell} = \mathbf{p}^\top_{-u} (G_n \tilde{G}_n)^{k-1} \Pi_n \Theta \mathbf{p}_\ell, \tag{5.19}$$

$$\tilde{b}^{2k}_{n,u,\ell} = \mathbf{p}^\top_{-u} (G_n \tilde{G}_n)^{k-1} G_n (\Pi_n \Theta)^* \overline{\mathbf{p}}_\ell. \tag{5.20}$$

The proof of Lemma 5.2 will be given in the Appendix.

For $n \in \mathbb{N}$ and $\mu, \nu \in \{1, 2, \ldots, K\}$, we define $\Xi_n^{\mu,\nu} \in \mathbb{C}^{dm_\mu \times dm_\nu}$ by the block representation

$$\Xi_n^{\mu,\nu} := \begin{pmatrix} \xi_n^{\mu,\nu}(1,1) & \xi_n^{\mu,\nu}(1,2) & \cdots & \xi_n^{\mu,\nu}(1,m_\nu) \\ \xi_n^{\mu,\nu}(2,1) & \xi_n^{\mu,\nu}(2,2) & \cdots & \xi_n^{\mu,\nu}(2,m_\nu) \\ \vdots & \vdots & & \vdots \\ \xi_n^{\mu,\nu}(m_\mu,1) & \xi_n^{\mu,\nu}(m_\mu,2) & \cdots & \xi_n^{\mu,\nu}(m_\mu,m_\nu) \end{pmatrix},$$

where, for $n \in \mathbb{N}$, $i \in \{1, \ldots, m_\mu\}$ and $j \in \{1, \ldots, m_\nu\}$, $\xi_n^{\mu,\nu}(i,j) \in \mathbb{C}^{d \times d}$ is defined by

$$\xi_n^{\mu,\nu}(i,j) := \sum_{r=0}^{j-1} \binom{n+i+j-2}{r} \binom{i+j-r-2}{i-1} \frac{p_\mu^{j-r-1} \overline{p}_\nu^{n+i+j-r-2}}{(1 - p_\mu \overline{p}_\nu)^{i+j-r-1}} I_d.$$

For $n \in \mathbb{N}$, we define $\Xi_n \in \mathbb{C}^{dM \times dM}$ by

$$\Xi_n := \begin{pmatrix} \Xi_n^{1,1} & \Xi_n^{1,2} & \cdots & \Xi_n^{1,K} \\ \Xi_n^{2,1} & \Xi_n^{2,2} & \cdots & \Xi_n^{2,K} \\ \vdots & \vdots & \ddots & \vdots \\ \Xi_n^{K,1} & \Xi_n^{K,2} & \cdots & \Xi_n^{K,K} \end{pmatrix}.$$

We also define $\rho \in \mathbb{C}^{dM \times d}$ and $\tilde{\rho} \in \mathbb{C}^{dM \times d}$ by the block representations

$$\rho := (\rho_{1,1}^\top, \ldots, \rho_{1,m_1}^\top \mid \rho_{2,1}^\top, \ldots, \rho_{2,m_2}^\top \mid \cdots \mid \rho_{K,1}^\top, \ldots, \rho_{K,m_K}^\top)^\top$$

and

$$\begin{aligned}\tilde{\rho} &:= (\tilde{\rho}_{1,1}^\top, \ldots, \tilde{\rho}_{1,m_1}^\top \mid \tilde{\rho}_{2,1}^\top, \ldots, \tilde{\rho}_{2,m_2}^\top \mid \cdots \mid \tilde{\rho}_{K,1}^\top, \ldots, \tilde{\rho}_{K,m_K}^\top)^\top \\ &= \left(\overline{\rho_{1,1}^\sharp}, \ldots, \overline{\rho_{1,m_1}^\sharp} \mid \overline{\rho_{2,1}^\sharp}, \ldots, \overline{\rho_{2,m_2}^\sharp} \mid \cdots \mid \overline{\rho_{K,1}^\sharp}, \ldots, \overline{\rho_{K,m_K}^\sharp}\right)^\top,\end{aligned}$$

respectively. For $n \in \mathbb{N}$, we define $v_n, \tilde{v}_n \in \mathbb{C}^{dM \times d}$ by

$$v_n := \sum_{\ell=0}^{\infty} \mathbf{p}_\ell a_{n+\ell}, \quad \tilde{v}_n := \sum_{\ell=0}^{\infty} \overline{\mathbf{p}}_\ell \tilde{a}_{n+\ell}.$$

Then, by Lemma 5 in [13], we have

$$v_n = \Xi_n \rho, \quad \tilde{v}_n = \overline{\Xi}_n \tilde{\rho}, \quad n \geq m_0 + 1.$$

Moreover, if $m_0 \geq 1$, then we have

$$v_n = \Xi_n \rho + \sum_{\ell=0}^{m_0-n} \mathbf{p}_\ell \rho_{0,n+\ell}, \quad \tilde{v}_n = \overline{\Xi}_n \tilde{\rho} + \sum_{\ell=0}^{m_0-n} \overline{\mathbf{p}}_\ell \tilde{\rho}_{0,n+\ell}, \quad n \in \{1, \ldots, m_0\}.$$

For $n \in \mathbb{Z}$, we define $w_n, \tilde{w}_n \in \mathbb{C}^{dM \times d}$ by

$$w_n := \sum_{\ell=0}^{\infty} \mathbf{p}_{\ell-n} a_\ell, \quad \tilde{w}_n := \sum_{\ell=0}^{\infty} \overline{\mathbf{p}}_{\ell-n} \tilde{a}_\ell.$$

To give closed-form expressions for w_n and $\tilde{w}_n$, we introduce some matrices. For $n \in \mathbb{Z}$ and $\mu, \nu \in \{1, 2, \ldots, K\}$, we define $\Phi_n^{\mu,\nu} \in \mathbb{C}^{dm_\mu \times dm_\nu}$ by the block representation

$$\Phi_n^{\mu,\nu} := \begin{pmatrix} \varphi_n^{\mu,\nu}(1,1) & \varphi_n^{\mu,\nu}(1,2) & \cdots & \varphi_n^{\mu,\nu}(1,m_\nu) \\ \varphi_n^{\mu,\nu}(2,1) & \varphi_n^{\mu,\nu}(2,2) & \cdots & \varphi_n^{\mu,\nu}(2,m_\nu) \\ \vdots & \vdots & & \vdots \\ \varphi_n^{\mu,\nu}(m_\mu,1) & \varphi_n^{\mu,\nu}(m_\mu,2) & \cdots & \varphi_n^{\mu,\nu}(m_\mu,m_\nu) \end{pmatrix},$$

where, for $n \in \mathbb{Z}$, $i = 1, \ldots, m_\mu$ and $j = 1, \ldots, m_\nu$, $\varphi_n^{\mu,\nu}(i,j) \in \mathbb{C}^{d\times d}$ is defined by

$$\varphi_n^{\mu,\nu}(i,j) := \sum_{q=0}^{i-1}\sum_{r=0}^{j-1}\binom{j-1}{r}\binom{r+q}{q}\binom{r-n}{i-q-1}\frac{p_\mu^{r+q+1-i-n}\overline{p}_\nu^{r+q}}{(1-p_\mu\overline{p}_\nu)^{r+q+1}}I_d.$$

For $n \in \mathbb{Z}$, we define $\Phi_n \in \mathbb{C}^{dM\times dM}$ by

$$\Phi_n := \begin{pmatrix} \Phi_n^{1,1} & \Phi_n^{1,2} & \cdots & \Phi_n^{1,K} \\ \Phi_n^{2,1} & \Phi_n^{2,2} & \cdots & \Phi_n^{2,K} \\ \vdots & \vdots & \ddots & \vdots \\ \Phi_n^{K,1} & \Phi_n^{K,2} & \cdots & \Phi_n^{K,K} \end{pmatrix}.$$

Here are closed-form expressions for w_n and $\tilde{w}_n$.

Lemma 5.3 *We have*

$$w_n = \Phi_n\rho + \sum_{\ell=0}^{m_0}\mathbf{p}_{\ell-n}\rho_{0,\ell}, \quad n \in \mathbb{Z},$$
$$\tilde{w}_n = \overline{\Phi}_n\tilde{\rho} + \sum_{\ell=0}^{m_0}\overline{\mathbf{p}}_{\ell-n}\tilde{\rho}_{0,\ell}, \quad n \in \mathbb{Z}.$$

The proof of Lemma 5.3 will be given in the Appendix.

Recall M from (5.13). For $n \in \mathbb{N}$ and $s \in \{1, \ldots, n\}$, we define

$$\ell_{n,s} := \{w_{n+1-s} - v_{n+1-s}\}^*(I_{dM} - \tilde{G}_n G_n)^{-1} \in \mathbb{C}^{d\times dM},$$
$$\tilde{\ell}_{n,s} := \{\tilde{w}_s - \tilde{v}_s\}^*(I_{dM} - G_n\tilde{G}_n)^{-1} \in \mathbb{C}^{d\times dM},$$
$$r_{n,s} := (\Pi_n\Theta)^*\tilde{v}_s + \tilde{G}_n\Pi_n\Theta v_{n+1-s} \in \mathbb{C}^{dM\times d}$$

and

$$\tilde{r}_{n,s} := \Pi_n\Theta v_{n+1-s} + G_n(\Pi_n\Theta)^*\tilde{v}_s \in \mathbb{C}^{dM\times d}.$$

Here are closed-form formulas for $(T_n(w))^{-1}$ with w satisfying (1.18) and $K \geq 1$.

Theorem 5.2 *We assume* (1.17)*,* (1.18) *and* $K \geq 1$ *for* K *in* (5.1)*. Then the following four assertions hold.*

(i) For $n \geq m_0 + 1$*,* $s \in \{1, \ldots, n\}$ *and* $t \in \{1, \ldots, n - m_0\}$*, we have*

$$\left(T_n(w)^{-1}\right)^{s,t} = \tilde{r}_{n,s}^*\tilde{\ell}_{n,t}^* + \sum_{\lambda=1}^{s\wedge t}\tilde{a}_{s-\lambda}^*\tilde{a}_{t-\lambda}.$$

(ii) For $n \geq m_0 + 1$, $s \in \{1, \ldots, n - m_0\}$ and $t \in \{1, \ldots, n\}$, we have

$$\left(T_n(w)^{-1}\right)^{s,t} = \tilde{\ell}_{n,s}\tilde{r}_{n,t} + \sum_{\lambda=1}^{s \wedge t} \tilde{a}_{s-\lambda}^{*}\tilde{a}_{t-\lambda}.$$

(iii) For $n \geq m_0 + 1$, $s \in \{1, \ldots, n\}$ and $t \in \{m_0 + 1, \ldots, n\}$, we have

$$\left(T_n(w)^{-1}\right)^{s,t} = r_{n,s}^{*}\ell_{n,t}^{*} + \sum_{\lambda=s \vee t}^{n} a_{\lambda-s}^{*}a_{\lambda-t}.$$

(iv) For $n \geq m_0 + 1$, $s \in \{m_0 + 1, \ldots, n\}$ and $t \in \{1, \ldots, n\}$, we have

$$\left(T_n(w)^{-1}\right)^{s,t} = \ell_{n,s}r_{n,t} + \sum_{\lambda=s \vee t}^{n} a_{\lambda-s}^{*}a_{\lambda-t}.$$

Proof (i) We assume $n \geq m_0 + 1$, $s \in \{1, \ldots, n\}$ and $t \in \{1, \ldots, n - m_0\}$. Then, by Lemma 5.2 (ii) above and Lemma 19 in [13], we have

$$\begin{aligned}
&\sum_{u=1}^{t}\sum_{k=1}^{\infty}\left\{\sum_{\lambda=0}^{\infty}\tilde{b}_{n,u,\lambda}^{2k-1}a_{n+1-s+\lambda}\right\}^{*}\tilde{a}_{t-u} \\
&\quad= \sum_{u=1}^{t}\sum_{k=1}^{\infty}\left\{\sum_{\lambda=0}^{\infty}\mathbf{p}_{-u}^{\top}(G_n\tilde{G}_n)^{k-1}\Pi_n\Theta\mathbf{p}_{\lambda}a_{n+1-s+\lambda}\right\}^{*}\tilde{a}_{t-u} \\
&\quad= \sum_{u=1}^{t}\sum_{k=1}^{\infty}\left\{\mathbf{p}_{-u}^{\top}(G_n\tilde{G}_n)^{k-1}\Pi_n\Theta v_{n+1-s}\right\}^{*}\tilde{a}_{t-u} \\
&\quad= \sum_{u=1}^{t}\left\{\mathbf{p}_{-u}^{\top}(I_{dM} - G_n\tilde{G}_n)^{-1}\Pi_n\Theta v_{n+1-s}\right\}^{*}\tilde{a}_{t-u} \\
&\quad= v_{n+1-s}^{*}(\Pi_n\Theta)^{*}(I_{dM} - \tilde{G}_n^{*}G_n^{*})^{-1}\sum_{u=1}^{t}\overline{\mathbf{p}}_{-u}\tilde{a}_{t-u}.
\end{aligned}$$

Similarly, by Lemma 5.2 (ii) above and Lemma 19 in [13],

$$\sum_{u=1}^{t}\sum_{k=1}^{\infty}\left\{\sum_{\lambda=0}^{\infty}\tilde{b}_{n,u,\lambda}^{2k}\tilde{a}_{s+\lambda}\right\}^{*}\tilde{a}_{t-u} = \tilde{v}_s^{*}\Pi_n\Theta G_n^{*}(I_{dM} - \tilde{G}_n^{*}G_n^{*})^{-1}\sum_{u=1}^{t}\overline{\mathbf{p}}_{-u}\tilde{a}_{t-u}.$$

However, $\sum_{u=1}^{t}\overline{\mathbf{p}}_{-u}\tilde{a}_{t-u} = \sum_{\lambda=0}^{\infty}\overline{\mathbf{p}}_{\lambda-t}\tilde{a}_{\lambda} - \sum_{\lambda=0}^{\infty}\overline{\mathbf{p}}_{\lambda}\tilde{a}_{t+\lambda} = \tilde{w}_t - \tilde{v}_t$. Therefore, the assertion (i) follows from Theorem 2.1 (i).

(iii) We assume $n \geq m_0 + 1$, $s \in \{1, \ldots, n\}$ and $t \in \{m_0 + 1, \ldots, n\}$. Then, by Lemma 5.2 (i) above and Lemma 19 in [13], we have

$$\begin{aligned}
&\sum_{u=t}^{n} \sum_{k=1}^{\infty} \left\{ \sum_{\lambda=0}^{\infty} b_{n,u,\lambda}^{2k-1} \tilde{a}_{s+\lambda} \right\}^* a_{u-t} \\
&= \sum_{u=t}^{n} \sum_{k=1}^{\infty} \left\{ \sum_{\lambda=0}^{\infty} \mathbf{p}_{u-n-1}^* (\tilde{G}_n G_n)^{k-1} (\Pi_n \Theta)^* \overline{\mathbf{p}}_\lambda \tilde{a}_{s+\lambda} \right\}^* a_{u-t} \\
&= \sum_{u=t}^{n} \sum_{k=1}^{\infty} \left\{ \mathbf{p}_{u-n-1}^* (\tilde{G}_n G_n)^{k-1} (\Pi_n \Theta)^* \tilde{v}_s \right\}^* a_{u-t} \\
&= \sum_{u=t}^{n} \left\{ \mathbf{p}_{u-n-1}^* (I_{dM} - \tilde{G}_n G_n)^{-1} (\Pi_n \Theta)^* \tilde{v}_s \right\}^* a_{u-t} \\
&= \tilde{v}_s^* \Pi_n \Theta (I_{dM} - G_n^* \tilde{G}_n^*)^{-1} \sum_{u=t}^{n} \mathbf{p}_{u-n-1} a_{u-t}.
\end{aligned}$$

Similarly, by Lemma 5.2 (i) above and Lemma 19 in [13], we have

$$\begin{aligned}
&\sum_{u=t}^{n} \sum_{k=1}^{\infty} \left\{ \sum_{\lambda=0}^{\infty} b_{n,u,\lambda}^{2k} a_{n+1-s+\lambda} \right\}^* a_{u-t} \\
&\quad = v_{n+1-s}^* (\Pi_n \Theta)^* \tilde{G}_n^* (I_{dM} - G_n^* \tilde{G}_n^*)^{-1} \sum_{u=t}^{n} \mathbf{p}_{u-n-1} a_{u-t}.
\end{aligned}$$

However, $\sum_{u=t}^{n} \mathbf{p}_{u-n-1} a_{u-t} = w_{n+1-t} - v_{n+1-t}$. Therefore, the assertion (ii) follows from Theorem 2.1 (ii).

(ii), (iv) By (2.9), (ii) and (iv) follow from (i) and (iii), respectively. □

Example 5.1 Suppose that $K \geq 1$, $m_\mu = 1$ for $\mu \in \{1, \ldots, K\}$ and $m_0 = 0$. Then,

$$h(z)^{-1} = -\rho_{0,0} - \sum_{\mu=1}^{K} \frac{1}{1 - \overline{p}_\mu z} \rho_{\mu,1}, \quad h_\sharp(z)^{-1} = -\rho_{0,0}^\sharp - \sum_{\mu=1}^{K} \frac{1}{1 - \overline{p}_\mu z} \rho_{\mu,1}^\sharp.$$

We have

$$\begin{aligned}
\mathbf{p}_n^\top &= (p_1^n I_d, \ldots, p_K^n I_d) \in \mathbb{C}^{d \times dK}, \quad n \in \mathbb{Z}, \\
\rho^\top &= (\rho_{1,1}^\top, \rho_{2,1}^\top, \ldots, \rho_{K,1}^\top) \in \mathbb{C}^{dK \times d}, \quad \tilde{\rho}^\top = \left(\overline{\rho_{1,1}^\sharp}, \overline{\rho_{2,1}^\sharp}, \ldots, \overline{\rho_{K,1}^\sharp} \right) \in \mathbb{C}^{dK \times d}.
\end{aligned}$$

We also have

$$\Theta = \begin{pmatrix} p_1 h_\sharp(p_1)\rho_{1,1}^* & 0 & \cdots & 0 \\ 0 & p_2 h_\sharp(p_2)\rho_{2,1}^* & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & p_K h_\sharp(p_K)\rho_{K,1}^* \end{pmatrix} \in \mathbb{C}^{dK\times dK},$$

$$\Lambda = \begin{pmatrix} \frac{1}{1-p_1\overline{p}_1} I_d & \frac{1}{1-p_1\overline{p}_2} I_d & \cdots & \frac{1}{1-p_1\overline{p}_K} I_d \\ \frac{1}{1-p_2\overline{p}_1} I_d & \frac{1}{1-p_2\overline{p}_2} I_d & \cdots & \frac{1}{1-p_2\overline{p}_K} I_d \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{1-p_K\overline{p}_1} I_d & \frac{1}{1-p_K\overline{p}_2} I_d & \cdots & \frac{1}{1-p_K\overline{p}_K} I_d \end{pmatrix} \in \mathbb{C}^{dK\times dK},$$

$$\Pi_n = \begin{pmatrix} p_1^n I_d & 0 & \cdots & 0 \\ 0 & p_2^n I_d & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & p_K^n I_d \end{pmatrix} \in \mathbb{C}^{dK\times dK}, \quad n \in \mathbb{Z},$$

$$\Xi_n = \begin{pmatrix} \frac{\overline{p}_1^n}{1-p_1\overline{p}_1} I_d & \frac{\overline{p}_2^n}{1-p_1\overline{p}_2} I_d & \cdots & \frac{\overline{p}_K^n}{1-p_1\overline{p}_K} I_d \\ \frac{\overline{p}_1^n}{1-p_2\overline{p}_1} I_d & \frac{\overline{p}_2^n}{1-p_2\overline{p}_2} I_d & \cdots & \frac{\overline{p}_K^n}{1-p_2\overline{p}_K} I_d \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\overline{p}_1^n}{1-p_K\overline{p}_1} I_d & \frac{\overline{p}_2^n}{1-p_K\overline{p}_2} I_d & \cdots & \frac{\overline{p}_K^n}{1-p_K\overline{p}_K} I_d \end{pmatrix} \in \mathbb{C}^{dK\times dK}, \quad n \in \mathbb{N},$$

$$\Phi_n = \begin{pmatrix} \frac{p_1^{-n}}{1-p_1\overline{p}_1} I_d & \frac{p_1^{-n}}{1-p_1\overline{p}_2} I_d & \cdots & \frac{p_1^{-n}}{1-p_1\overline{p}_K} I_d \\ \frac{p_2^{-n}}{1-p_2\overline{p}_1} I_d & \frac{p_2^{-n}}{1-p_2\overline{p}_2} I_d & \cdots & \frac{p_2^{-n}}{1-p_2\overline{p}_K} I_d \\ \vdots & \vdots & \ddots & \vdots \\ \frac{p_K^{-n}}{1-p_K\overline{p}_1} I_d & \frac{p_K^{-n}}{1-p_K\overline{p}_2} I_d & \cdots & \frac{p_K^{-n}}{1-p_K\overline{p}_K} I_d \end{pmatrix} \in \mathbb{C}^{dK\times dK}, \quad n \in \mathbb{Z},$$

$$\begin{aligned} G_n &= \Pi_n\Theta\Lambda \in \mathbb{C}^{dK\times dK}, \quad \tilde{G}_n = (\Pi_n\Theta)^*\Lambda^\top \in \mathbb{C}^{dK\times dK}, \quad n \in \mathbb{Z}, \\ v_n &= \Xi_n\rho \in \mathbb{C}^{dK\times d}, \quad \tilde{v}_n = \overline{\Xi}_n\tilde{\rho} \in \mathbb{C}^{dK\times d}, \quad n \in \mathbb{N}, \\ w_n &= \Phi_n\rho + \mathbf{p}_{-n}\rho_{0,0} \in \mathbb{C}^{dK\times d}, \quad \tilde{w}_n = \overline{\Phi}_n\tilde{\rho} + \overline{\mathbf{p}}_{-n}\tilde{\rho}_{0,0} \in \mathbb{C}^{dK\times d}, \quad n \in \mathbb{Z}. \end{aligned}$$

Example 5.2 In Example 5.1, we further assume $d = K = 1$. Then, we can write $h(z) = h_\sharp(z) = -(1 - \overline{p}z)/\rho$, where $\rho \in \mathbb{C}\backslash\{0\}$ and $p \in \mathbb{D}\backslash\{0\}$. It follows that

$$\begin{aligned} &c_0 = -1/\rho, \quad c_1 = \overline{p}/\rho, \quad c_k = 0 \quad (k \geq 2), \\ &a_k = \rho(\overline{p})^k, \quad \tilde{a}_k = \overline{a}_k, \quad k \in \mathbb{N} \cup \{0\}. \end{aligned}$$

Since $\gamma(k)=\sum_{\ell=0}^{\infty} c_{k+\ell}\overline{c}_{\ell}$ and $\gamma(-k)=\overline{\gamma(k)}$ for $k\in\mathbb{N}\cup\{0\}$, we have

$$T_2(w)=\frac{1}{|\rho|^2}\begin{pmatrix}1+|p|^2 & -p\\ -\overline{p} & 1+|p|^2\end{pmatrix},$$

hence

$$T_2(w)^{-1}=\frac{|\rho|^2}{1+|p|^2+|p|^4}\begin{pmatrix}1+|p|^2 & p\\ \overline{p} & 1+|p|^2\end{pmatrix}.$$

We also have

$$\tilde{A}_2=\overline{\rho}\begin{pmatrix}1 & p\\ 0 & 1\end{pmatrix}\quad\text{and}\quad A_2=\rho\begin{pmatrix}1 & 0\\ \overline{p} & 1\end{pmatrix}$$

for $\tilde{A}_2$ and A_2 in (2.14) and (2.15), respectively. By simple calculations, we have

$$(\ell_{2,1},\ell_{2,2})=\frac{\overline{\rho}}{(\overline{p})^2(1-|p|^6)}(1+|p|^2,\overline{p}),\quad \left(\tilde{\ell}_{2,1},\tilde{\ell}_{2,2}\right)=\left(\overline{\ell_{2,2}},\overline{\ell_{2,1}}\right),$$
$$(r_{2,1},r_{2,2})=-\rho\overline{p}|p|^2(1-|p|^2)(\overline{p}(1+|p|^2),|p|^2),\quad \left(\tilde{r}_{2,1},\tilde{r}_{2,2}\right)=\left(\overline{r_{2,2}},\overline{r_{2,1}}\right)$$

hence

$$T_2(w)^{-1}=\tilde{A}_2^*\tilde{A}_2+\begin{pmatrix}\tilde{\ell}_{2,1}\\ \tilde{\ell}_{2,2}\end{pmatrix}\left(\tilde{r}_{2,1},\tilde{r}_{2,2}\right)=A_2^*A_2+\begin{pmatrix}\ell_{2,1}\\ \ell_{2,2}\end{pmatrix}\left(r_{2,1},r_{2,2}\right)$$

which agrees with equalities in Theorem 5.2.

6 Linear-time algorithm

As in Sect. 5, we assume (1.17) and (1.18). Let K be as in (5.1) with (5.2). In this section, we explain how Theorems 5.1 and 5.2 above provide us with a linear-time algorithm to compute the solution Z to the block Toeplitz system (1.19).

For

$$Y=(y_1^\top,\dots,y_n^\top)^\top\in\mathbb{C}^{dn\times d}\quad\text{with}\quad y_s\in\mathbb{C}^{d\times d},\quad s\in\{1,\dots,n\},\tag{6.1}$$

let

$$Z=(z_1^\top,\dots,z_n^\top)^\top\in\mathbb{C}^{dn\times d}\quad\text{with}\quad z_s\in\mathbb{C}^{d\times d},\quad s\in\{1,\dots,n\},$$

be the solution to (1.19), that is, $Z=T_n(w)^{-1}Y$. For m_0 in (5.1), let $n\geq 2m_0+1$ so that $n-m_0\geq m_0+1$ holds.

Recall $\tilde{A}_n$ and A_n from (2.14) and (2.15), respectively. If $K = 0$, then it follows from Lemma 2.1 and Theorem 5.1 (ii), (iv) that

$$z_s = \tilde{\alpha}_{n,s}, \quad s \in \{1, \ldots, n - m_0\},$$
$$z_s = \alpha_{n,s}, \quad s \in \{m_0 + 1, \ldots, n\},$$

where

$$(\tilde{\alpha}_{n,1}^\top, \ldots, \tilde{\alpha}_{n,n}^\top)^\top := \tilde{A}_n^* \tilde{A}_n Y \quad \text{with} \quad \tilde{\alpha}_{n,s} \in \mathbb{C}^{d\times d}, \quad s \in \{1, \ldots, n\},$$
$$(\alpha_{n,1}^\top, \ldots, \alpha_{n,n}^\top)^\top := A_n^* A_n Y \quad \text{with} \quad \alpha_{n,s} \in \mathbb{C}^{d\times d}, \quad s \in \{1, \ldots, n\}.$$

On the other hand, if $K \geq 1$, then we see from Lemma 2.1 and Theorem 5.2 (ii), (iv) that

$$z_s = \tilde{\ell}_{n,s} \tilde{R}_n + \tilde{\alpha}_{n,s}, \quad s \in \{1, \ldots, n - m_0\},$$
$$z_s = \ell_{n,s} R_n + \alpha_{n,s}, \quad s \in \{m_0 + 1, \ldots, n\},$$

where

$$\tilde{R}_n := \sum_{t=1}^{n} \tilde{r}_{n,t} y_t \in \mathbb{C}^{d\times d}, \quad R_n := \sum_{t=1}^{n} r_{n,t} y_t \in \mathbb{C}^{d\times d}.$$

Therefore, algorithms to compute $\tilde{A}_n^* \tilde{A}_n Y$ and $A_n^* A_n Y$ in $O(n)$ operations imply that of Z. We present the former ones below.

For $n \in \mathbb{N} \cup \{0\}$, $\mu \in \{1, \ldots, K\}$ and $j \in \{1, \ldots, m_\mu\}$, we define $q_{\mu,j}(n) \in \mathbb{C}^{d\times d}$ by $q_{\mu,j}(n) := p_{\mu,j}(n + j - 1)$, that is,

$$q_{\mu,j}(n) = \binom{n + j - 1}{j - 1} p_\mu^n I_d. \tag{6.2}$$

For $n \in \mathbb{N}$, $\mu \in \{1, \ldots, K\}$ and $j \in \{1, \ldots, m_\mu\}$, we define the upper trianglular block Toeplitz matrix $Q_{\mu,j,n} \in \mathbb{C}^{dn\times dn}$ by

$$Q_{\mu,j,n} := \begin{pmatrix} q_{\mu,j}(0) & q_{\mu,j}(1) & q_{\mu,j}(2) & \cdots & q_{\mu,j}(n-1) \\ & q_{\mu,j}(0) & q_{\mu,j}(1) & \cdots & q_{\mu,j}(n-2) \\ & & \ddots & \ddots & \vdots \\ & & & \ddots & q_{\mu,j}(1) \\ 0 & & & & q_{\mu,j}(0) \end{pmatrix}.$$

Notice that

$$Q^*_{\mu,j,n} := \begin{pmatrix} q^*_{\mu,j}(0) & & & & 0 \\ q^*_{\mu,j}(1) & q^*_{\mu,j}(0) & & & \\ q^*_{\mu,j}(2) & q^*_{\mu,j}(1) & \ddots & & \\ \vdots & \vdots & \ddots & \ddots & \\ q^*_{\mu,j}(n-1) & q^*_{\mu,j}(n-2) & \cdots & q^*_{\mu,j}(1) & q^*_{\mu,j}(0) \end{pmatrix}$$

with $q^*_{\mu,j}(n) = \binom{n+j-1}{j-1}\overline{p}^n_\mu I_d$. For $n \in \mathbb{N}$, $\mu \in \{1, \dots, K\}$ and $j \in \{1, \dots, m_\mu\}$, we define the block diagonal matrices $\tilde{D}_{\mu,j,n} \in \mathbb{C}^{dn \times dn}$ and $D_{\mu,j,n} \in \mathbb{C}^{dn \times dn}$ by

$$\tilde{D}_{\mu,j,n} := \begin{pmatrix} \tilde{\rho}_{\mu,j} & 0 & \cdots & 0 \\ 0 & \tilde{\rho}_{\mu,j} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \tilde{\rho}_{\mu,j} \end{pmatrix} \quad \text{and} \quad D_{\mu,j,n} := \begin{pmatrix} \rho_{\mu,j} & 0 & \cdots & 0 \\ 0 & \rho_{\mu,j} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \rho_{\mu,j} \end{pmatrix},$$

respectively. Moreover, for $n \geq m_0 + 1$, we define the upper and lower triangular block Toeplitz matrices $\tilde{\Delta}_n \in \mathbb{C}^{dn \times dn}$ and $\Delta_n \in \mathbb{C}^{dn \times dn}$ by

$$\tilde{\Delta}_n := \begin{pmatrix} \tilde{\rho}_{0,0} & \tilde{\rho}_{0,1} & \cdots & \tilde{\rho}_{0,m_0} & & 0 \\ & \tilde{\rho}_{0,0} & \tilde{\rho}_{0,1} & & \ddots & \\ & & \ddots & \ddots & & \tilde{\rho}_{0,m_0} \\ & & & \ddots & \ddots & \vdots \\ & & & & \tilde{\rho}_{0,0} & \tilde{\rho}_{0,1} \\ 0 & & & & & \tilde{\rho}_{0,0} \end{pmatrix}$$

and

$$\Delta_n := \begin{pmatrix} \rho_{0,0} & & & & & 0 \\ \rho_{0,1} & \rho_{0,0} & & & & \\ \vdots & \rho_{0,1} & \ddots & & & \\ \rho_{0,m_0} & & \ddots & \ddots & & \\ & \ddots & & \ddots & \rho_{0,0} & \\ 0 & & \rho_{0,m_0} & \cdots & \rho_{0,1} & \rho_{0,0} \end{pmatrix},$$

respectively. Note that both $\tilde{\Delta}_n$ and Δ_n are sparse matrices in the sense that they have only $O(n)$ nonzero elements.

By (5.5)–(5.8), we have

$$\tilde{A}_n = \tilde{\Delta}_n + \sum_{\mu=1}^{K}\sum_{j=1}^{m_\mu} Q_{\mu,j,n}\tilde{D}_{\mu,j,n}, \quad n \ge m_0 + 1,$$

$$A_n = \Delta_n + \sum_{\mu=1}^{K}\sum_{j=1}^{m_\mu} Q^*_{\mu,j,n}D_{\mu,j,n}, \quad n \ge m_0 + 1.$$

Therefore, it is enough to give linear-time algorithms to compute $Q_{\mu,i,n}Y$ and $Q^*_{\mu,i,n}Y$ for $Y \in \mathbb{C}^{dn\times d}$ in $O(n)$ operations. The following two propositions provide such linear-time algorithms.

Proposition 6.1 *Let $n \in \mathbb{N}$, $\mu \in \{1,\dots,K\}$ and Y be as in (6.1). We put $Z_{\mu,i} = Q_{\mu,i,n}Y$ for $i \in \{1,\dots,m_\mu\}$. Then the component blocks $z_{\mu,i}(s)$ of $Z_{\mu,i} = (z^\top_{\mu,i}(1),\dots,z^\top_{\mu,i}(n))^\top$ satisfy the following equalities:*

$$z_{\mu,i}(n) = q_{\mu,i}(0)y_n, \quad i \in \{1,\dots,m_\mu\}, \tag{6.3}$$

$$z_{\mu,1}(s) = p_\mu z_{\mu,1}(s+1) + q_{\mu,1}(0)y_s, \quad s \in \{1,\dots,n-1\} \tag{6.4}$$

$$\begin{aligned} z_{\mu,i}(s) &= p_\mu z_{\mu,i}(s+1) + z_{\mu,i-1}(s) + \{q_{\mu,i}(0) - q_{\mu,i-1}(0)\}y_s, \\ &\qquad i \in \{2,\dots,m_\mu\},\ s \in \{1,\dots,n-1\}. \end{aligned} \tag{6.5}$$

Proof From the definition of $Q_{\mu,i,n}$, (6.3) is trivial. For $q_{\mu,i}(k)$ in (6.2), Pascal's rule yields the following recursions:

$$q_{\mu,1}(k+1) = p_\mu q_{\mu,1}(k), \quad k \in \mathbb{N}\cup\{0\}, \tag{6.6}$$

$$q_{\mu,i}(k+1) = p_\mu q_{\mu,i}(k) + q_{\mu,i-1}(k+1), \quad i \in \{2,\dots,j\},\ k \in \mathbb{N}\cup\{0\}. \tag{6.7}$$

For $s \in \{1,\dots,n-1\}$, we see, from (6.6),

$$\begin{aligned} z_{\mu,1}(s) &= q_{\mu,1}(0)y_s + \sum_{t=0}^{n-s-1} q_{\mu,1}(t+1)y_{s+t+1} \\ &= q_{\mu,1}(0)y_s + p_\mu \sum_{t=0}^{n-s-1} q_{\mu,1}(t)y_{s+t+1} = q_{\mu,1}(0)y_s + p_\mu z_{\mu,1}(s+1), \end{aligned}$$

and, from (6.7),

$$\begin{aligned} z_{\mu,i}(s) &= q_{\mu,i}(0)y_s + \sum_{t=0}^{n-s-1} q_{\mu,i}(t+1)y_{s+t+1} \\ &= \{q_{\mu,i}(0) - q_{\mu,i-1}(0)\}y_s + p_\mu \sum_{t=0}^{n-s-1} q_{\mu,1}(t)y_{s+t+1} + \sum_{t=0}^{n-s} q_{\mu,i-1}(t)y_{s+t} \\ &= \{q_{\mu,i}(0) - q_{\mu,i-1}(0)\}y_s + p_\mu z_{\mu,i}(s+1) + z_{\mu,i-1}(s) \end{aligned}$$

for $i \in \{2, \ldots, j\}$. Thus, (6.4) and (6.5) follow. □

By Proposition 6.1, we can compute $z_{\mu,i}(s)$ in the following order in $O(n)$ operations:

$$\begin{aligned} &z_{\mu,1}(n) \to \cdots \to z_{\mu,1}(1) \to z_{\mu,2}(n) \to \cdots \to z_{\mu,2}(1) \\ &\to \cdots \to z_{\mu,m_\mu}(n) \to \cdots \to z_{\mu,m_\mu}(1). \end{aligned}$$

Proposition 6.2 *Let $n \in \mathbb{N}$, $\mu \in \{1, \ldots, K\}$ and Y be as in (6.1). We put $W_{\mu,i} = Q^*_{\mu,i,n} Y$ for $i \in \{1, \ldots, m_\mu\}$. Then the component blocks $w_{\mu,i}(s)$ of $W_{\mu,i} = (w^\top_{\mu,i}(1), \ldots, w^\top_{\mu,i}(n))^\top$ satisfy the following equalities:*

$$\begin{aligned} &w_{\mu,i}(1) = q^*_{\mu,i}(0) y_1, \quad i \in \{1, \ldots, m_\mu\}, \\ &w_{\mu,1}(s+1) = \overline{p}_\mu w_{\mu,1}(s) + q^*_{\mu,1}(0) y_{s+1}, \quad s \in \{1, \ldots, n-1\} \\ &w_{\mu,i}(s+1) = \overline{p}_\mu w_{\mu,i}(s) + w_{\mu,i-1}(s+1) + \{q^*_{\mu,i}(0) - q^*_{\mu,i-1}(0)\} y_{s+1}, \\ &\qquad i \in \{2, \ldots, m_\mu\}, \ s \in \{1, \ldots, n-1\}. \end{aligned}$$

The proof of Proposition 6.2 is similar to that of Proposition 6.1; we omit it.

By Proposition 6.2, we can compute $w_{\mu,i}(s)$ in the following order in $O(n)$ operations:

$$\begin{aligned} &w_{\mu,1}(1) \to \cdots \to w_{\mu,1}(n) \to w_{\mu,2}(1) \to \cdots \to w_{\mu,2}(n) \\ &\to \cdots \to w_{\mu,m_\mu}(1) \to \cdots \to w_{\mu,m_\mu}(n). \end{aligned}$$

Acknowledgements The author is grateful to two referees for their valuable comments and helpful suggestions.

A Proofs of Lemmas 5.2 and 5.3

As in Sect. 5, we assume (1.17) and (1.18). We use the same notation as in Sect. 5. For K in (5.1) with (5.2), we assume $K \geq 1$.

We prove Lemma 5.2.

Proof (i) We assume $n \geq u \geq m_0 + 1$, and prove (5.17) and (5.18) by induction. First, from Lemma 5.1,

$$b^1_{n,u,\ell} = \beta_{u+l} = \mathbf{p}^*_{u-n-1} (\Pi_n \Theta)^* \overline{\mathbf{p}}_\ell.$$

Next, for $k = 1, 2, \ldots$, we assume (5.17). Then, by Lemma 5.1,

$$\begin{aligned} b^{2k}_{n,u,\ell} &= \sum_{m=0}^{\infty} b^{2k-1}_{n,u,m} \beta^*_{n+1+m+\ell} \\ &= \mathbf{p}^*_{u-n-1} (\tilde{G}_n G_n)^{k-1} (\Pi_n \Theta)^* \left(\sum_{m=0}^{\infty} \overline{\mathbf{p}}_m \mathbf{p}^{\top}_m \right) \Pi_n \Theta \mathbf{p}_\ell \\ &= \mathbf{p}^*_{u-n-1} (\tilde{G}_n G_n)^{k-1} (\Pi_n \Theta)^* \Lambda^{\top} \Pi_n \Theta \mathbf{p}_\ell = \mathbf{p}^*_{u-n-1} (\tilde{G}_n G_n)^{k-1} \tilde{G}_n \Pi_n \Theta \mathbf{p}_\ell \end{aligned}$$

or (5.18). From this as well as Lemma 5.1,

$$\begin{aligned} b^{2k+1}_{n,u,\ell} &= \sum_{m=0}^{\infty} b^{2k}_{n,u,m} \beta_{n+1+m+\ell} \\ &= \mathbf{p}^*_{u-n-1} (\tilde{G}_n G_n)^{k-1} \tilde{G}_n \Pi_n \Theta \left(\sum_{m=0}^{\infty} \mathbf{p}_m \mathbf{p}^*_m \right) (\Pi_n \Theta)^* \overline{\mathbf{p}}_\ell \\ &= \mathbf{p}^*_{u-n-1} (\tilde{G}_n G_n)^{k-1} \tilde{G}_n \Pi_n \Theta \Lambda (\Pi_n \Theta)^* \overline{\mathbf{p}}_\ell = \mathbf{p}^*_{u-n-1} (\tilde{G}_n G_n)^{k} (\Pi_n \Theta)^* \overline{\mathbf{p}}_\ell \end{aligned}$$

or (5.17) with k replaced by $k+1$. Thus (5.17) and (5.18) follow.

(ii) We assume $1 \le u \le n - m_0$, and prove (5.19) and (5.20) by induction. First, from Lemma 5.1,

$$\tilde{b}^1_{n,u,\ell} = \beta^*_{n+1-u+\ell} = \mathbf{p}^{\top}_{-u} \Pi_n \Theta \mathbf{p}_\ell.$$

Next, for $k = 1, 2, \ldots$, we assume (5.19). Then, by Lemma 5.1,

$$\begin{aligned} \tilde{b}^{2k}_{n,u,\ell} &= \sum_{m=0}^{\infty} \tilde{b}^{2k-1}_{n,u,m} \beta_{n+1+m+\ell} \\ &= \mathbf{p}^{\top}_{-u} (G_n \tilde{G}_n)^{k-1} \Pi_n \Theta \left(\sum_{m=0}^{\infty} \mathbf{p}_m \mathbf{p}^*_m \right) (\Pi_n \Theta)^* \overline{\mathbf{p}}_\ell \\ &= \mathbf{p}^{\top}_{-u} (G_n \tilde{G}_n)^{k-1} \Pi_n \Theta \Lambda (\Pi_n \Theta)^* \overline{\mathbf{p}}_\ell = \mathbf{p}^{\top}_{-u} (G_n \tilde{G}_n)^{k-1} G_n (\Pi_n \Theta)^* \overline{\mathbf{p}}_\ell \end{aligned}$$

or (5.20). From this as well as Lemma 5.1,

$$\begin{aligned} \tilde{b}^{2k+1}_{n,u,\ell} &= \sum_{m=0}^{\infty} \tilde{b}^{2k}_{n,u,m} \beta^*_{n+1+m+\ell} = \mathbf{p}^{\top}_{-u} (G_n \tilde{G}_n)^{k-1} G_n (\Pi_n \Theta)^* \left(\sum_{m=0}^{\infty} \overline{\mathbf{p}}_m \mathbf{p}^{\top}_m \right) \Pi_n \Theta \mathbf{p}_\ell \\ &= \mathbf{p}^{\top}_{-u} (G_n \tilde{G}_n)^{k-1} G_n (\Pi_n \Theta)^* \Lambda^{\top} \Pi_n \Theta \mathbf{p}_\ell = \mathbf{p}^{\top}_{-u} (G_n \tilde{G}_n)^{k} \Pi_n \Theta \mathbf{p}_\ell \end{aligned}$$

or (5.19) with k replaced by $k+1$. Thus (5.19) and (5.20) follow. □

To prove Lemma 5.3, we need some propositions.

Proposition A.1 *For $m, n \in \mathbb{Z}$, $i, j \in \mathbb{N} \cup \{0\}$ and $x, y \in \mathbb{D}$, we have*

$$\begin{aligned}&\frac{1}{i!j!}\left(\frac{\partial}{\partial x}\right)^i \left(\frac{\partial}{\partial y}\right)^j \frac{x^m y^n}{1-xy}\\&\quad = \sum_{q=0}^{i}\sum_{r=0}^{j}\binom{n}{j-r}\binom{q+r}{q}\binom{m+r}{i-q}\frac{x^{m+q+r-i}y^{n+q+r-j}}{(1-xy)^{q+r+1}}.\end{aligned}$$

Proof Let $m, n \in \mathbb{Z}$, $i, j \in \mathbb{N} \cup \{0\}$ and $x, y \in \mathbb{D}$. Then, we have

$$\begin{aligned}\frac{1}{j!}\left(\frac{\partial}{\partial y}\right)^j \frac{y^n}{1-xy} &= \sum_{r=0}^{j}\left\{\frac{1}{r!}\left(\frac{\partial}{\partial y}\right)^r \frac{1}{1-xy}\right\}\left\{\frac{1}{(j-r)!}\left(\frac{\partial}{\partial y}\right)^{j-r} y^n\right\}\\&= \sum_{r=0}^{j}\binom{n}{j-r}\frac{x^r y^{n+r-j}}{(1-xy)^{r+1}},\end{aligned}$$

hence

$$\begin{aligned}&\frac{1}{i!j!}\left(\frac{\partial}{\partial x}\right)^i \left(\frac{\partial}{\partial y}\right)^j \frac{x^m y^n}{1-xy}\\&= \sum_{r=0}^{j}\binom{n}{j-r}y^{n+r-j}\sum_{q=0}^{i}\left\{\frac{1}{q!}\left(\frac{\partial}{\partial x}\right)^q \frac{1}{(1-xy)^{r+1}}\right\}\left\{\frac{1}{(i-q)!}\left(\frac{\partial}{\partial x}\right)^{i-q} x^{m+r}\right\}\\&= \sum_{r=0}^{j}\binom{n}{j-r}y^{n+r-j}\sum_{q=0}^{i}\left\{\binom{q+r}{q}\frac{y^q}{(1-xy)^{q+r+1}}\right\}\left\{\binom{m+r}{i-q}x^{m+q+r-i}\right\}\\&= \sum_{q=0}^{i}\sum_{r=0}^{j}\binom{n}{j-r}\binom{q+r}{q}\binom{m+r}{i-q}\frac{x^{m+q+r-i}y^{n+q+r-j}}{(1-xy)^{q+r+1}}.\end{aligned}$$

Thus, the proposition follows. □

Proposition A.2 *For $n \in \mathbb{Z}$, $i, j \in \mathbb{N} \cup \{0\}$ and $x, y \in \mathbb{D}$, we have*

$$\sum_{\ell=0}^{\infty}\binom{n+\ell}{i}\binom{j+\ell}{j}x^{n+\ell-i}y^{\ell} = \sum_{q=0}^{i}\sum_{r=0}^{j}\binom{j}{r}\binom{r+q}{q}\binom{n+r}{i-q}\frac{x^{n+r+q-i}y^{r+q}}{(1-xy)^{r+q+1}}.$$

Proof Let $n \in \mathbb{Z}$, $i, j \in \mathbb{N} \cup \{0\}$ and $x, y \in \mathbb{D}$. Since $x^n y^j/(1-xy) = \sum_{\ell=0}^{\infty} x^{n+\ell}y^{j+\ell}$, we have

$$\frac{1}{i!j!}\left(\frac{\partial}{\partial x}\right)^i \left(\frac{\partial}{\partial y}\right)^j \frac{x^n y^j}{1-xy} = \sum_{\ell=0}^{\infty}\binom{n+\ell}{i}\binom{j+\ell}{j}x^{n+\ell-i}y^{\ell}.$$

On the other hand, by Proposition A.1, we have

$$\frac{1}{i!j!}\left(\frac{\partial}{\partial x}\right)^i\left(\frac{\partial}{\partial y}\right)^j\frac{x^n y^j}{1-xy}=\sum_{q=0}^{i}\sum_{r=0}^{j}\binom{j}{r}\binom{r+q}{q}\binom{n+r}{i-q}\frac{x^{n+r+q-i}y^{r+q}}{(1-xy)^{r+q+1}}.$$

Comparing, we obtain the proposition. □

We are ready to prove Lemma 5.3.

Proof By (5.5)–(5.8) and Proposition A.2, we have, for $n \in \mathbb{Z}$, $\mu \in \{1, \dots, K\}$ and $i \in \{1, \dots, m_\mu\}$,

$$\begin{aligned}\sum_{\ell=0}^{\infty} p_{\mu,i}(\ell-n)a_\ell &= \sum_{\ell=0}^{m_0} p_{\mu,i}(\ell-n)\rho_{0,\ell}\\ &\quad+\sum_{\nu=1}^{K}\sum_{j=1}^{m_\nu}\left\{\sum_{\ell=0}^{\infty}\binom{\ell-n}{i-1}\binom{\ell+j-1}{j-1}p_\mu^{\ell-i+1-n}\overline{p}_\nu^{\ell}\right\}\rho_{\nu,j}\\ &= \sum_{\ell=0}^{m_0} p_{\mu,i}(\ell-n)\rho_{0,\ell}+\sum_{\nu=1}^{K}\sum_{j=1}^{m_\nu}\varphi_n^{\mu,\nu}(i,j)\rho_{\nu,j}\end{aligned}$$

and

$$\begin{aligned}\sum_{\ell=0}^{\infty} \overline{p}_{\mu,i}(\ell-n)\tilde{a}_\ell &= \sum_{\ell=0}^{m_0} \overline{p}_{\mu,i}(\ell-n)\tilde{\rho}_{0,\ell}\\ &\quad+\sum_{\nu=1}^{K}\sum_{j=1}^{m_\nu}\left\{\sum_{\ell=0}^{\infty}\binom{\ell-n}{i-1}\binom{\ell+j-1}{j-1}\overline{p}_\mu^{\ell-i+1-n}p_\nu^{\ell}\right\}\tilde{\rho}_{\nu,j}\\ &= \sum_{\ell=0}^{m_0} \overline{p}_{\mu,i}(\ell-n)\tilde{\rho}_{0,\ell}+\sum_{\nu=1}^{K}\sum_{j=1}^{m_\nu}\overline{\varphi}_n^{\mu,\nu}(i,j)\tilde{\rho}_{\nu,j}.\end{aligned}$$

Thus, the lemma follows. □

References

1. Baxter, G.: An asymptotic result for the finite predictor. Math. Scand. **10**, 137–144 (1962)
2. Böttcher, A., Silbermann, B.: Introduction to Large Truncated Toeplitz Matrices. Springer, New York (1999)
3. Böttcher, A., Silbermann, B.: Analysis of Toeplitz Operators, 2nd edn. Springer, Berlin (2006)
4. Brockwell, P.J., Davis, R.A.: Time Series: Theory and Methods, 2nd edn. Springer, New York (1991)
5. Cheng, R., Pourahmadi, M.: Baxter's inequality and convergence of finite predictors of multivariate stochastic processes. Probab. Theory Relat. Fields **95**(1), 115–124 (1993)
6. Gohberg, I.C., Fel'dman, I.A.: Convolution Equations and Projection Methods for Their Solution, volume 41 of Translations of Mathematical Monographs. American Mathematical Society, Providence (1974)

7. Gohberg, I.C., Semencul, A.A.: The inversion of finite Toeplitz matrices and their continual analogues. (Russian) Mat. Issled. **7**(2), 201–223 (1972)
8. Grenander, U., Szegö, G.: Toeplitz Forms and Their Applications, 2nd edn. Chelsea Publishing Co., New York (1984)
9. Hannan, E.J., Deistler, M.: The Statistical Theory of Linear Systems. Wiley, New York (1988)
10. Helson, H., Lowdenslager, D.: Prediction theory and Fourier series in several variables II. Acta Math. **106**, 175–213 (1961)
11. Inoue, A.: Asymptotics for the partial autocorrelation function of a stationary process. J. Anal. Math. **81**, 65–109 (2000)
12. Inoue, A.: AR and MA representation of partial autocorrelation functions, with applications. Probab. Theory Relat. Fields **140**(3–4), 523–551 (2008)
13. Inoue, A.: Closed-form expression for finite predictor coefficients of multivariate ARMA processes. J. Multivariate Anal. **176**, 104578 (2020)
14. Inoue, A., Kasahara, Y.: Explicit representation of finite predictor coefficients and its applications. Ann. Stat. **34**(2), 973–993 (2006)
15. Inoue, A., Kasahara, Y., Pourahmadi, M.: The intersection of past and future for multivariate stationary processes. Proc. Amer. Math. Soc. **144**(4), 1779–1786 (2016)
16. Inoue, A., Kasahara, Y., Pourahmadi, M.: Baxter's inequality for finite predictor coefficients of multivariate long-memory stationary processes. Bernoulli **24**(2), 1202–1232 (2018)
17. Kasahara, Y., Bingham, N.H.: Matricial Baxter's theorem with a Nehari sequence. Math. Nachr. **291**(17–18), 2590–2598 (2018)
18. Katsnelson, V.E., Kirstein, B.: On the theory of matrix-valued functions belonging to the Smirnov class. In: Topics in Interpolation Theory. (Leipzig, 1994). Oper. Theory Adv. Appl. vol. 95, pp. 299–350, Birkhäuser (1997)
19. Masani, P.: The prediction theory of multivariate stochastic processes. III. Unbounded spectral densities. Acta Math. **104**, 141–162 (1960)
20. Peller, V.V.: Hankel Operators and Their Applications. Springer, New York (2003)
21. Rozanov, Yu.A.: Stationary Random Processes. Holden-Day, San Francisco (1967)
22. Simon, B.: Orthogonal Polynomials on the Unit Circle. Part 1. Classical Theory. American Mathematical Society, Providence (2005)
23. Simon, B.: Orthogonal Polynomials on the Unit Circle. Part 2. Spectral Theory. American Mathematical Society, Providence (2005)
24. Simon, B.: Szegö's Theorem and Its Descendants. Spectral Theory for L^2 Perturbations of Orthogonal Polynomials. Princeton University Press, Princeton (2011)
25. Subba Rao, S., Yang, J.: Reconciling the Gaussian and Whittle Likelihood with an application to estimation in the frequency domain. Ann. Stat. **49**(5), 2774–2802 (2021)