On the Asymptotic Optimality of a Low-Complexity Coding Strategy for WSS, MA, and AR Vector Sources

Jesús Gutiérrez-Gutiérrez * [ID]**, Marta Zárraga-Rodríguez** [ID] **and Xabier Insausti** [ID]

Tecnun, University of Navarra, Paseo Manuel Lardizábal 13, 20018 San Sebastián, Spain;
mzarraga@tecnun.es (M.Z.-R.); xinsausti@tecnun.es (X.I.)
* Correspondence: jgutierrez@tecnun.es; Tel.: +34-943-219-877

Abstract: In this paper, we study the asymptotic optimality of a low-complexity coding strategy for Gaussian vector sources. Specifically, we study the convergence speed of the rate of such a coding strategy when it is used to encode the most relevant vector sources, namely wide sense stationary (WSS), moving average (MA), and autoregressive (AR) vector sources. We also study how the coding strategy considered performs when it is used to encode perturbed versions of those relevant sources. More precisely, we give a sufficient condition for such perturbed versions so that the convergence speed of the rate remains unaltered.

Keywords: source coding; low-complexity; wide sense stationary (WSS) vector source; moving average (MA) vector source; autoregressive (AR) vector source

1. Introduction

In [1], Kolmogorov gave a formula for the rate distortion function (RDF) of Gaussian vectors and for the RDF of Gaussian wide sense stationary (WSS) sources. In [2], Pearl presented an upper bound for the RDF of finite-length data blocks of any Gaussian WSS source and proved that such a bound tends to the RDF of the source when the length of the data block grows. However, he did not propose a coding strategy to achieve his bound for a given block length. In [3], we presented a tighter upper bound for the RDF of finite-length data blocks of any Gaussian WSS source, and we proposed a low-complexity coding strategy to achieve our bound. Obviously, since such a bound is tighter than the one given by Pearl, it also tends to the RDF of the source when the length of the data block grows. In [4], we generalized our low-complexity coding strategy to encode (compress) finite-length data blocks of any Gaussian vector source. Moreover, in [4], we also gave a sufficient condition for the vector source in order to make such a coding strategy asymptotically optimal. We recall that a coding strategy is asymptotically optimal if its rate tends to the RDF of the source as the length of the data block grows. Such a sufficient condition requires the Gaussian vector source to be asymptotically WSS (AWSS). The definition of the AWSS process was first introduced in [5], Section 6, and extended to vector processes in [6], Definition 7.1. However, the convergence speed of the rate of the coding strategy considered (i.e., how fast the rate of the coding strategy tends to the RDF of the AWSS vector source) was not studied in [4].

In this paper, we present a less restrictive sufficient condition for the vector source to make the coding strategy considered asymptotically optimal. Moreover, we study the convergence speed of the rate of such a coding strategy when it is used to encode the most relevant vector sources, namely, WSS, moving average (MA), and autoregressive (AR) vector sources. In this paper, we also study how the coding strategy considered performs when it is used to encode perturbed versions of those relevant sources. Specifically, we give a sufficient condition for such perturbed versions so that the convergence speed of the rate remains unaltered.

The study of the convergence speed in any information-theoretic problem is not an easy task. To study the aforementioned convergence speed, we first need to derive new mathematical results on block Toeplitz matrices and new mathematical results on the correlation matrices of the WSS, MA, and AR vector processes. These new mathematical results are useful not only to study the convergence speed in the information-theoretic problem considered, but also in other problems. In fact, as an example, in Appendix H, we use such mathematical results to study the convergence speed in a statistical signal processing problem on filtering WSS vector processes.

The paper is organized as follows. In Section 2, we give several new mathematical results on block Toeplitz matrices. In Section 3, using the results obtained in Section 2, we give several new mathematical results on the correlation matrices of WSS, MA, and AR vector processes. In Section 4, we recall the low-complexity coding strategy presented in [4], and using the results obtained in Section 3, we study the asymptotic optimality of such a coding strategy when it is used to encode WSS, MA, and AR vector sources. In Section 4, we also study how the coding strategy considered performs when it is used to encode perturbed versions of those sources. Finally, in Section 5, some conclusions are presented.

2. Several New Results on Block Toeplitz Matrices

In this section, we present new results on the product of block Toeplitz matrices, on the inverse of a block Toeplitz matrix, and on block circulant matrices. These results will be used in Section 3. We begin by introducing some notation.

2.1. Notation

In this paper, $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{R}$, and $\mathbb{C}$ denote the set of natural numbers (that is, the set of positive integers), the set of integer numbers, the set of real numbers, and the set of complex numbers, respectively. $\mathbb{C}^{M \times N}$ is the set of all $M \times N$ complex matrices. I_N stands for the $N \times N$ identity matrix. $0_{M \times N}$ denotes the $M \times N$ zero matrix. V_n is the $n \times n$ Fourier unitary matrix, i.e.,

$$[V_n]_{j,k} = \frac{1}{\sqrt{n}} e^{-\frac{2\pi(j-1)(k-1)}{n}i}, \qquad j,k \in \{1,\ldots,n\},$$

with i being the imaginary unit. We denote by $\lambda_1(A),\ldots,\lambda_n(A)$ the eigenvalues of an $n \times n$ Hermitian matrix A arranged in decreasing order. $*$ denotes the conjugate transpose. $\otimes$ is the Kronecker product. $\|\cdot\|_2$ and $\|\cdot\|_F$ are the spectral norm and the Frobenius norm, respectively.

If $n \in \mathbb{N}$ and $A_j \in \mathbb{C}^{M \times N}$ for all $j \in \{1,\ldots,n\}$, then $\mathrm{diag}(A_1,\ldots,A_n)$ is the $n \times n$ block diagonal matrix whose $M \times N$ blocks are given by:

$$[\mathrm{diag}(A_1,\ldots,A_n)]_{j,k} = \delta_{j,k} A_j, \qquad j,k \in \{1,\ldots,n\},$$

where δ is the Kronecker delta. We also denote by $\mathrm{diag}_{1 \le k \le n}(A_k)$ the matrix $\mathrm{diag}(A_1,\ldots,A_n)$.

If $n \in \mathbb{N}$ and $F \colon \mathbb{R} \to \mathbb{C}^{M \times N}$ is a continuous 2π-periodic function, $T_n(F)$ stands for the $n \times n$ block Toeplitz matrix generated by F whose $M \times N$ blocks are given by:

$$[T_n(F)]_{j,k} = \mathsf{F}_{j-k}, \qquad j,k \in \{1,\ldots,n\},$$

where $\{\mathsf{F}_k\}_{k \in \mathbb{Z}}$ is the sequence of Fourier coefficients of F, that is,

$$\mathsf{F}_k = \frac{1}{2\pi} \int_0^{2\pi} e^{-k\omega i} F(\omega) d\omega \qquad \forall k \in \mathbb{Z}.$$

We denote by $C_n(F)$ the $n \times n$ block circulant matrix with $M \times N$ blocks defined as:

$$C_n(F) = (V_n \otimes I_M) \text{diag}_{1 \le k \le n} \left(F \left(\frac{2\pi(k-1)}{n} \right) \right) (V_n \otimes I_N)^*.$$

If $A_n \in \mathbb{C}^{nM \times nN}$, then C_{A_n} is the $n \times n$ block circulant matrix with the $M \times N$ blocks given by:

$$C_{A_n} = (V_n \otimes I_M) \text{diag}_{1 \le k \le n} \left([(V_n \otimes I_M)^* A_n (V_n \otimes I_N)]_{k,k} \right) (V_n \otimes I_N)^*.$$

We denote by $\widehat{C}_n(F)$ the $n \times n$ block circulant matrix with the $M \times N$ blocks defined as $\widehat{C}_n(F) = C_{T_n(F)}$.

If $F(\omega)$ is Hermitian for all $\omega \in \mathbb{R}$ (or equivalently, $T_n(F)$ is Hermitian for all $n \in \mathbb{N}$ (see, e.g., [6], Theorem 4.4), then $\inf F$ denotes $\inf_{\omega \in [0, 2\pi]} \lambda_N(F(\omega))$. We recall that (see [7], Proposition 3):

$$\inf_{n \in \mathbb{N}} \lambda_{nN}(T_n(F)) = \inf F = \min_{\omega \in [0, 2\pi]} \lambda_N(F(\omega)). \tag{1}$$

2.2. Product of Block Toeplitz Matrices

We begin this subsection with a result on the entries of the block Toeplitz matrices generated by the product of two functions, which is a direct consequence of the Parseval theorem.

Lemma 1. *Consider two continuous 2π-periodic functions $F \colon \mathbb{R} \to \mathbb{C}^{M \times N}$ and $G \colon \mathbb{R} \to \mathbb{C}^{N \times K}$. Let $\{F_k\}_{k \in \mathbb{Z}}$ and $\{G_k\}_{k \in \mathbb{Z}}$ be the sequences of Fourier coefficients of F and G, respectively. Then:*

$$[T_n(FG)]_{j,k} = \sum_{h=-\infty}^{\infty} F_{j-h} G_{h-k}$$

for all $n \in \mathbb{N}$ and $j, k \in \{1, \ldots, n\}$.

Proof. See Appendix A. $\square$

We can now give a result on the product of two block Toeplitz matrices when one of them is generated by a trigonometric polynomial. We recall that an $M \times N$ trigonometric polynomial of degree $p \in \mathbb{N} \cup \{0\}$ is a function $F \colon \mathbb{R} \to \mathbb{C}^{M \times N}$ of the form:

$$F(\omega) = \sum_{k=-p}^{p} e^{k \omega i} A_k \qquad \forall \omega \in \mathbb{R}, \tag{2}$$

where $A_k \in \mathbb{C}^{M \times N}$ with $|k| \le p$. It can be easily proven (see, e.g., [6], Example 4.3) that the sequence of the Fourier coefficients $\{F_k\}_{k \in \mathbb{Z}}$ of the continuous 2π-periodic function F in Equation (2) is given by:

$$F_k = \begin{cases} A_k & \text{if } |k| \le p, \\ 0_{M \times N} & \text{if } |k| > p. \end{cases}$$

Lemma 2. *Let F, G, $\{F_k\}_{k \in \mathbb{Z}}$, and $\{G_k\}_{k \in \mathbb{Z}}$ be as in Lemma 1.*

1. *If F is a trigonometric polynomial of degree p, then:*

$$[T_n(F)T_n(G) - T_n(FG)]_{j,k} = \begin{cases} -\sum_{h=j-p}^{0} F_{j-h} G_{h-k} & \text{if } j \le p, \\ 0_{M \times K} & \text{if } p+1 \le j \le n-p, \\ -\sum_{h=n+1}^{j+p} F_{j-h} G_{h-k} & \text{if } j \ge n-p+1, \end{cases} \tag{3}$$

and:

$$\|T_n(F)T_n(G) - T_n(FG)\|_F \leq \sqrt{p(p+1)\left(\frac{1}{2\pi}\int_0^{2\pi}\|F(\omega)\|_F^2 d\omega\right)\left(\frac{1}{2\pi}\int_0^{2\pi}\|G(\omega)\|_F^2 d\omega\right)} \quad (4)$$

for all $n \in \mathbb{N}$ and $j, k \in \{1, \ldots, n\}$.

2. If G is a trigonometric polynomial of degree q, then:

$$[T_n(F)T_n(G) - T_n(FG)]_{j,k} = \begin{cases} -\sum_{h=k-q}^{0} \mathsf{F}_{j-h}\mathsf{G}_{h-k} & \text{if } k \leq q, \\ 0_{M\times K} & \text{if } q+1 \leq k \leq n-q, \\ -\sum_{h=n+1}^{k+q} \mathsf{F}_{j-h}\mathsf{G}_{h-k} & \text{if } k \geq n-q+1, \end{cases} \quad (5)$$

and:

$$\|T_n(F)T_n(G) - T_n(FG)\|_F \leq \sqrt{q(q+1)\left(\frac{1}{2\pi}\int_0^{2\pi}\|F(\omega)\|_F^2 d\omega\right)\left(\frac{1}{2\pi}\int_0^{2\pi}\|G(\omega)\|_F^2 d\omega\right)} \quad (6)$$

for all $n \in \mathbb{N}$ and $j, k \in \{1, \ldots, n\}$.

3. If F is a trigonometric polynomial of degree p and G is a trigonometric polynomial of degree q, then:

$$T_n(F)T_n(G) - T_n(FG) = \begin{pmatrix} \xi_1(F,G) & 0_{pM\times(n-2q)K} & 0_{pM\times qK} \\ 0_{(n-2p)M\times qK} & 0_{(n-2p)M\times(n-2q)K} & 0_{(n-2p)M\times qK} \\ 0_{pM\times qK} & 0_{pM\times(n-2q)K} & \xi_2(F,G) \end{pmatrix}$$

and:

$$\|T_n(F)T_n(G) - T_n(FG)\|_F = \sqrt{\|\xi_1(F,G)\|_F^2 + \|\xi_2(F,G)\|_F^2} \quad (7)$$

for all $n \geq \max\{2p, 2q\}$, where $\xi_1(F,G), \xi_2(F,G) \in \mathbb{C}^{pM\times qK}$ are given by:

$$[\xi_1(F,G)]_{j,k} = -\sum_{h=\max\{j-p,k-q\}}^{0} \mathsf{F}_{j-h}\mathsf{G}_{h-k}$$

and:

$$[\xi_2(F,G)]_{j,k} = -\sum_{h=1}^{\min\{j,k\}} \mathsf{F}_{j-p-h}\mathsf{G}_{h+q-k}$$

for all $j \in \{1, \ldots, p\}$ and $k \in \{1, \ldots, q\}$.

Proof. See Appendix B. □

2.3. Inverse of a Block Toeplitz Matrix

Lemma 3. *Let* $F\colon \mathbb{R} \to \mathbb{C}^{N\times N}$ *be a trigonometric polynomial of degree* p.

1. *If* $F(\omega)$ *is invertible for all* $\omega \in \mathbb{R}$ *and* $\{T_n(F)\}$ *is stable (i.e.,* $T_n(F)$ *is invertible for all* $n \in \mathbb{N}$ *and* $\{\|(T_n(F))^{-1}\|_2\}$ *is bounded), then:*

$$\|(T_n(F))^{-1} - T_n(F^{-1})\|_F \leq \sup_{m\in\mathbb{N}}\|(T_m(F))^{-1}\|_2\sqrt{p(p+1)\left(\frac{1}{2\pi}\int_0^{2\pi}\|F(\omega)\|_F^2 d\omega\right)\left(\frac{1}{2\pi}\int_0^{2\pi}\|(F(\omega))^{-1}\|_F^2 d\omega\right)}$$

for all $n \in \mathbb{N}$.

2. *If* $F(\omega)$ *is positive definite for all* $\omega \in \mathbb{R}$, *then:*

$$\|(T_n(F))^{-1} - T_n(F^{-1})\|_F \leq \frac{1}{\inf F}\sqrt{p(p+1)\left(\frac{1}{2\pi}\int_0^{2\pi}\|F(\omega)\|_F^2 d\omega\right)\left(\frac{1}{2\pi}\int_0^{2\pi}\|(F(\omega))^{-1}\|_F^2 d\omega\right)} \quad (8)$$

for all $n \in \mathbb{N}$.

Proof. See Appendix C. □

2.4. Block Circulant Matrices

Lemma 4. *Consider $A_n, B_n \in \mathbb{C}^{nM \times nN}$. Then:*

$$\|C_{A_n} - C_{B_n}\|_F \le \|A_n - B_n\|_F$$

and:

$$\|A_n - C_{A_n}\|_F \le 2\|A_n - B_n\|_F + \|B_n - C_{B_n}\|_F. \tag{9}$$

Moreover, if B_n is an $n \times n$ block circulant matrix with $M \times N$ blocks, then:

$$C_{B_n} = B_n \tag{10}$$

and:

$$\|A_n - C_{A_n}\|_F \le 2\|A_n - B_n\|_F. \tag{11}$$

Proof. See Appendix D. □

Lemma 5. *Let $F \colon \mathbb{R} \to \mathbb{C}^{M \times N}$ be a trigonometric polynomial of degree p. Then:*

$$\|T_n(F) - \widehat{C}_n(F)\|_F \le \|T_n(F) - C_n(F)\|_F = \sqrt{\sum_{k=1}^{p} k \left(\|\mathsf{F}_k\|_F^2 + \|\mathsf{F}_{-k}\|_F^2 \right)}$$

for all $n > 2p$. Furthermore,

$$\lim_{n \to \infty} \|T_n(F) - \widehat{C}_n(F)\|_F = \lim_{n \to \infty} \|T_n(F) - C_n(F)\|_F.$$

Proof. See Appendix E. □

3. Several New Results on the Correlation Matrices of Certain Random Vector Processes

Let $\{x_n\}$ be a (complex) random N-dimensional vector process, that is x_n is a (complex) random N-dimensional (column) vector for all $n \in \mathbb{N}$. In this section, we study the boundedness of the sequence $\left\{ \left\| E\left(x_{n:1} x_{n:1}^* \right) - C_{E\left(x_{n:1} x_{n:1}^* \right)} \right\|_F \right\}$ when $\{x_n\}$ is a WSS, MA, or AR vector process, where:

$$x_{n:1} = \begin{pmatrix} x_n \\ x_{n-1} \\ \vdots \\ x_1 \end{pmatrix}, \qquad n \in \mathbb{N},$$

and E denotes expectation.

3.1. WSS Vector Processes

In this subsection, we review the concept of the WSS vector process, and we prove that the sequence $\left\{ \left\| E\left(x_{n:1} x_{n:1}^* \right) - C_{E\left(x_{n:1} x_{n:1}^* \right)} \right\|_F \right\}$ is bounded when $\{x_n\}$ is a WSS vector process whose power spectral density (PSD) is a trigonometric polynomial.

Definition 1. *Let $X : \mathbb{R} \to \mathbb{C}^{N \times N}$ be continuous and 2π-periodic. A random N-dimensional vector process $\{x_n\}$ is said to be WSS with PSD X if it has constant mean (i.e., $E(x_{n_1}) = E(x_{n_2})$ for all $n_1, n_2 \in \mathbb{N}$) and $\{E(x_{n:1}x_{n:1}^*)\} = \{T_n(X)\}$.*

Lemma 6. *If $\{x_n\}$ is a WSS vector process whose PSD is a trigonometric polynomial, then $\left\{ \left\| E(x_{n:1}x_{n:1}^*) - C_{E(x_{n:1}x_{n:1}^*)} \right\|_F \right\}$ is bounded.*

Proof. This is a direct consequence of Lemma 5. □

3.2. VMA Processes

In this subsection, we review the concept of the MA vector (VMA) process, and we prove that the sequence $\left\{ \left\| E(x_{n:1}x_{n:1}^*) - C_{E(x_{n:1}x_{n:1}^*)} \right\|_F \right\}$ is bounded when $\{x_n\}$ is a VMA process of finite order.

Definition 2. *A zero-mean random N-dimensional vector process $\{x_n\}$ is said to be a VMA process if:*

$$x_n = w_n + \sum_{k=1}^{n-1} \mathsf{G}_{-k} w_{n-k} \qquad \forall n \in \mathbb{N}, \tag{12}$$

where $\mathsf{G}_{-k} \in \mathbb{C}^{N \times N}$ for all $k \in \mathbb{N}$ and $\{w_n\}$ is a zero-mean WSS N-dimensional vector process whose PSD is an $N \times N$ positive semidefinite matrix Λ. If there exists $q \in \mathbb{N}$ such that $\mathsf{G}_{-k} = 0_{N \times N}$ for all $k > q$, then $\{x_n\}$ is called a VMA process of (finite) order q or a VMA(q) process.

Lemma 7. *If $\{x_n\}$ is a VMA(q) process as in Definition 2, then $\left\{ \left\| E(x_{n:1}x_{n:1}^*) - C_{E(x_{n:1}x_{n:1}^*)} \right\|_F \right\}$ is bounded.*

Proof. See Appendix F. □

3.3. VAR Processes

In this subsection, we review the concept of the AR vector (VAR) process, and we study the boundedness of the sequence $\left\{ \left\| E(x_{n:1}x_{n:1}^*) - C_{E(x_{n:1}x_{n:1}^*)} \right\|_F \right\}$ when $\{x_n\}$ is a VAR process of finite order.

Definition 3. *A zero-mean random N-dimensional vector process $\{x_n\}$ is said to be a VAR process if:*

$$x_n = w_n - \sum_{k=1}^{n-1} \mathsf{F}_{-k} x_{n-k} \qquad \forall n \in \mathbb{N}, \tag{13}$$

where $\mathsf{F}_{-k} \in \mathbb{C}^{N \times N}$ for all $k \in \mathbb{N}$ and $\{w_n\}$ is a zero-mean WSS N-dimensional vector process whose PSD is an $N \times N$ positive definite matrix Λ. If there exists $p \in \mathbb{N}$ such that $\mathsf{F}_{-k} = 0_{N \times N}$ for all $k > p$, then $\{x_n\}$ is called a VAR process of (finite) order p or a VAR(p) process.

Lemma 8. *Let $\{x_n\}$ be a VAR(p) process as in Definition 3. Suppose that $F(\omega) = I_N + \sum_{k=1}^{p} e^{-k\omega i} \mathsf{F}_{-k}$ is invertible for all $\omega \in \mathbb{R}$ and $\{\|(T_n(F))^{-1}\|_2\}$ is bounded. Then, $\left\{ \left\| E(x_{n:1}x_{n:1}^*) - C_{E(x_{n:1}x_{n:1}^*)} \right\|_F \right\}$ is bounded.*

Proof. See Appendix G. □

4. On the Asymptotic Optimality of a Low-Complexity Coding Strategy for Gaussian Vector Sources

4.1. Low-Complexity Coding Strategy Considered

In [1], Kolmogorov gave a formula for the RDF of a real zero-mean Gaussian N-dimensional vector x with a positive definite correlation matrix $E\left(\mathrm{xx}^\top\right)$, namely,

$$R_{\mathrm{x}}(D) = \frac{1}{N} \sum_{k=1}^{N} \max\left\{0, \frac{1}{2}\ln\frac{\lambda_k\left(E\left(\mathrm{xx}^\top\right)\right)}{\theta}\right\} \qquad \forall D \in \left(0, \frac{\mathrm{tr}\left(E\left(\mathrm{xx}^\top\right)\right)}{N}\right],$$

where $\top$ stands for the transpose, tr denotes the trace, and θ is a real number satisfying:

$$D = \frac{1}{N} \sum_{k=1}^{N} \min\left\{\theta, \lambda_k\left(E\left(\mathrm{xx}^\top\right)\right)\right\}.$$

We recall that $R_{\mathrm{x}}(D)$ can be thought of as the minimum rate (measured in nats) at which x can be encoded (compressed) in order to be able to recover it with a mean squared error (MSE) per dimension no larger than a given distortion D, that is:

$$\frac{E\left(\|\mathrm{x} - \widetilde{\mathrm{x}}\|_2^2\right)}{N} \le D,$$

where $\widetilde{\mathrm{x}}$ denotes the estimation of x.

If $D \in \left(0, \lambda_N\left(E\left(\mathrm{xx}^\top\right)\right)\right]$, an optimal coding strategy to achieve $R_{\mathrm{x}}(D)$ is to encode $[z]_{1,1}, \ldots, [z]_{N,1}$ separately with $E(\|[z]_{k,1} - \widetilde{[z]_{k,1}}\|_2^2) \le D$ for all $k \in \{1, \ldots, N\}$, where $\mathrm{z} = U^\top \mathrm{x}$ with U being a real orthogonal eigenvector matrix of $E\left(\mathrm{xx}^\top\right)$ (see [8], Corollary 1). Observe that in order to obtain U, we need to know the correlation matrix $E\left(\mathrm{xx}^\top\right)$. This coding strategy also requires an optimal coding method for real Gaussian random variables.

In [4], Theorem 3, we gave a low-complexity coding strategy for any Gaussian N-dimensional vector source $\{x_n\}$. According to that strategy, to encode a finite-length data block $x_{n:1}$ of such a source, we first compute the block discrete Fourier transform (DFT) of $x_{n:1}$:

$$y_{n:1} = \left(V_n^* \otimes I_N\right) x_{n:1}, \tag{14}$$

and then, we encode $y_{\lceil \frac{n}{2} \rceil}, \ldots, y_n$ separately (i.e., if n is even, we encode $y_{\frac{n}{2}}, \widehat{y_{\frac{n}{2}+1}}, \ldots, \widehat{y_{n-1}}, y_n$ separately, and if n is odd, we encode $\widehat{y_{\frac{n+1}{2}}}, \ldots, \widehat{y_{n-1}}, y_n$ separately) with:

$$\frac{E\left(\left\|\widehat{y_k} - \widehat{\widetilde{y_k}}\right\|_2^2\right)}{2N} \le \frac{D}{2}, \qquad k \in \left\{\left\lceil\frac{n}{2}\right\rceil, \ldots, n-1\right\} \setminus \left\{\frac{n}{2}\right\},$$

and:

$$\frac{E\left(\|y_k - \widetilde{y_k}\|_2^2\right)}{N} \le D, \qquad k \in \left\{\frac{n}{2}, n\right\} \cap \mathbb{N},$$

where $\lceil x \rceil$ denotes the smallest integer higher than or equal to $x \in \mathbb{R}$ and:

$$\widehat{z} = \begin{pmatrix} \mathrm{Re}(z) \\ \mathrm{Im}(z) \end{pmatrix} = \begin{pmatrix} \mathrm{Re}([z]_{1,1}) \\ \vdots \\ \mathrm{Re}([z]_{N,1}) \\ \mathrm{Im}([z]_{1,1}) \\ \vdots \\ \mathrm{Im}([z]_{N,1}) \end{pmatrix} \qquad \forall z \in \mathbb{C}^{N \times 1}$$

with Re and Im being the real part and the imaginary part of a complex number, respectively.

As our coding strategy requires the computation of the block DFT, its computational complexity is $O(nN \log n)$ whenever the fast Fourier transform (FFT) algorithm is used. We recall that the computational complexity of the optimal coding strategy for $x_{n:1}$ is $O(n^2 N^2)$ since it requires the computation of $U_n^\top x_{n:1}$, where U_n is a real orthogonal eigenvector matrix of $E\left(x_{n:1}x_{n:1}^\top\right)$. Observe that such an eigenvector matrix U_n also needs to be computed, which further increases the complexity. Hence, the main advantage of our coding strategy is that it notably reduces the computational complexity of coding $x_{n:1}$. Moreover, our coding strategy does not require the knowledge of $E\left(x_{n:1}x_{n:1}^\top\right)$. It only requires the knowledge of $E\left(\widehat{y_k}\widehat{y_k}^\top\right)$, with $k \in \{\lceil \frac{n}{2} \rceil, \ldots, n\}$.

We finish this subsection by reviewing a result that provides an upper bound for the distance between $R_{x_{n:1}}(D)$ and the rate of our coding strategy $\widetilde{R}_{x_{n:1}}(D)$ (see [4], Theorem 3).

Theorem 1. *Consider $n, N \in \mathbb{N}$. Let x_k be a random N-dimensional vector for all $k \in \{1, \ldots, n\}$. Suppose that $x_{n:1}$ is a real zero-mean Gaussian vector with a positive definite correlation matrix (or equivalently, $\lambda_{nN}\left(E\left(x_{n:1}x_{n:1}^\top\right)\right) > 0$). Let $y_{n:1}$ be the random vector given by Equation (14). If $D \in \left(0, \lambda_{nN}\left(E\left(x_{n:1}x_{n:1}^\top\right)\right)\right]$, then:*

$$0 \le \widetilde{R}_{x_{n:1}}(D) - R_{x_{n:1}}(D) \le \frac{1}{2} \ln\left(1 + \frac{\left\| E\left(x_{n:1}x_{n:1}^\top\right) - C_{E\left(x_{n:1}x_{n:1}^\top\right)} \right\|_F}{\sqrt{nN}\lambda_{nN}\left(E\left(x_{n:1}x_{n:1}^\top\right)\right)} \right), \tag{15}$$

where:

$$\widetilde{R}_{x_{n:1}}(D) = \begin{cases} \dfrac{R_{y_{\frac{n}{2}}}(D) + 2\sum_{k=\frac{n}{2}+1}^{n-1} R_{\widehat{y_k}}\left(\frac{D}{2}\right) + R_{y_n}(D)}{n} & \text{if } n \text{ is even,} \\[2em] \dfrac{2\sum_{k=\frac{n+1}{2}}^{n-1} R_{\widehat{y_k}}\left(\frac{D}{2}\right) + R_{y_n}(D)}{n} & \text{if } n \text{ is odd.} \end{cases}$$

4.2. On the Asymptotic Optimality of the Low-Complexity Coding Strategy Considered

In this subsection, we study the asymptotic optimality of our coding strategy for Gaussian vector sources. We begin by presenting a new result that provides a sufficient condition for the source to make such a coding strategy asymptotically optimal.

Theorem 2. *Let $\{x_n\}$ be a real zero-mean Gaussian N-dimensional vector process. Suppose that $\inf_{n \in \mathbb{N}} \lambda_{nN}\left(E\left(x_{n:1}x_{n:1}^\top\right)\right) > 0$ and $\lim_{n \to \infty} \dfrac{\left\| E\left(x_{n:1}x_{n:1}^\top\right) - C_{E\left(x_{n:1}x_{n:1}^\top\right)} \right\|_F}{\sqrt{n}} = 0$. If $D \in \left(0, \inf_{n \in \mathbb{N}} \lambda_{nN}\left(E\left(x_{n:1}x_{n:1}^\top\right)\right)\right]$, then:*

$$\lim_{n \to \infty} \left(\widetilde{R}_{x_{n:1}}(D) - R_{x_{n:1}}(D) \right) = 0.$$

Hence, if $\{R_{x_{n:1}}(D)\}$ *is convergent, then:*

$$\lim_{n\to\infty} \widetilde{R}_{x_{n:1}}(D) = \lim_{n\to\infty} R_{x_{n:1}}(D). \tag{16}$$

Proof. From Equation (15), we have:

$$0 \le \widetilde{R}_{x_{n:1}}(D) - R_{x_{n:1}}(D) \le \frac{1}{2}\ln\left(1 + \frac{\left\|E\left(x_{n:1}x_{n:1}^\top\right) - C_{E\left(x_{n:1}x_{n:1}^\top\right)}\right\|_F}{\sqrt{nN}\inf_{m\in\mathbb{N}}\lambda_{mN}\left(E\left(x_{m:1}x_{m:1}^\top\right)\right)}\right) \qquad \forall n \in \mathbb{N},$$

and therefore, Theorem 2 is proven. $\square$

We recall that $\lim_{n\to\infty} R_{x_{n:1}}(D)$ is the RDF of the source $\{x_n\}$.

In [4], Theorem 4, we gave a more restrictive sufficient condition for the source to make the coding strategy considered asymptotically optimal. Specifically, in [4], Theorem 4, we proved that Equation (16) holds if $\{x_n\}$ is AWSS. However, the convergence speed of the rate of the coding strategy considered (i.e., how fast the rate of the coding strategy tends to the RDF of the AWSS vector source) was not studied in [4]. We now study the convergence speed of the rate of such a coding strategy when it is used to encode the most relevant vector sources, namely WSS vector sources, VMA sources, and VAR sources. It should be mentioned that this convergence speed depends on the sequence $\left\{\left\|E\left(x_{n:1}x_{n:1}^\top\right) - C_{E\left(x_{n:1}x_{n:1}^\top\right)}\right\|_F\right\}$ whose boundedness is studied in Section 3 for these three types of vector sources.

Theorem 3. *Let* $\{x_n\}$ *be a real zero-mean Gaussian WSS N-dimensional vector process whose PSD X is a trigonometric polynomial. Suppose that* $\inf X > 0$ *(or equivalently,* $\det(X(\omega)) \neq 0$ *for all* $\omega \in \mathbb{R}$*). If* $D \in (0, \inf X]$*, there exists* $K \in [0, \infty)$ *such that:*

$$0 \le \widetilde{R}_{x_{n:1}}(D) - R_{x_{n:1}}(D) \le \frac{1}{2}\ln\left(1 + \frac{K}{\sqrt{n}}\right) \qquad \forall n \in \mathbb{N}. \tag{17}$$

Proof. As $\{T_n(X)\} = \{E\left(x_{n:1}x_{n:1}^*\right)\}$, $T_n(X)$ is positive semidefinite for all $n \in \mathbb{N}$. Consequently, from [7], Proposition 3, $X(\omega)$ is positive semidefinite for all $\omega \in \mathbb{R}$. Therefore, applying Equation (1), $\inf X > 0$ if and only if $\det(X(\omega)) \neq 0$ for all $\omega \in \mathbb{R}$. Equation (17) is a direct consequence of Equation (1), Theorem 1, and Lemma 6. $\square$

Theorem 4. *Let* $\{x_n\}$ *be a VMA(q) process as in Definition 2. Suppose that* $\det(\Lambda) \neq 0$ *and* $\{\|(T_n(G))^{-1}\|_2\}$ *is bounded with* $G(\omega) = I_N + \sum_{k=1}^q e^{-k\omega i}G_{-k}$ *for all* $\omega \in \mathbb{R}$*. If* $\{x_n\}$ *is real and Gaussian, and* $D \in \left(0, \inf_{n\in\mathbb{N}} \lambda_{nN}\left(E\left(x_{n:1}x_{n:1}^\top\right)\right)\right]$*, there exists* $K \in [0, \infty)$ *such that:*

$$0 \le \widetilde{R}_{x_{n:1}}(D) - R_{x_{n:1}}(D) \le \frac{1}{2}\ln\left(1 + \frac{K}{\sqrt{n}}\right) \qquad \forall n \in \mathbb{N}.$$

Proof. Since $\det(T_n(G)) = 1$ for all $n \in \mathbb{N}$, from Equation (A3), we have:

$$\lambda_{nN}\left(E\left(x_{n:1}x_{n:1}^\top\right)\right) = \frac{1}{\left\|\left(E\left(x_{n:1}x_{n:1}^\top\right)\right)^{-1}\right\|_2} = \frac{1}{\left\|\left(T_n(G)T_n(\Lambda)\left(T_n(G)\right)^*\right)^{-1}\right\|_2}$$

$$= \frac{1}{\left\|\left((T_n(G))^{-1}\right)^* T_n(\Lambda^{-1})\left(T_n(G)\right)^{-1}\right\|_2} \ge \frac{1}{\left\|\left((T_n(G))^{-1}\right)^*\right\|_2 \|T_n(\Lambda^{-1})\|_2 \left\|(T_n(G))^{-1}\right\|_2}$$

$$= \frac{\lambda_N(\Lambda)}{\left\|(T_n(G))^{-1}\right\|_2^2} \ge \frac{\lambda_N(\Lambda)}{\left(\sup_{m\in\mathbb{N}}\left\|(T_m(G))^{-1}\right\|_2\right)^2} > 0 \qquad \forall n \in \mathbb{N}.$$

Hence, $\inf_{n \in \mathbb{N}} \lambda_{nN} \left(E \left(x_{n:1} x_{n:1}^\top \right) \right) > 0$. Theorem 1 and Lemma 7 prove Theorem 4. $\quad\square$

Theorem 5. *Let $\{x_n\}$ be a VAR(p) process as in Definition 3. Suppose that $F(\omega) = I_N + \sum_{k=1}^p e^{-k\omega i} F_{-k}$ is invertible for all $\omega \in \mathbb{R}$ and $\{\|(T_n(F))^{-1}\|_2\}$ is bounded. If $\{x_n\}$ is real and Gaussian and $D \in \left(0, \inf_{n \in \mathbb{N}} \lambda_{nN} \left(E \left(x_{n:1} x_{n:1}^\top \right) \right)\right]$, there exists $K \in [0, \infty)$ such that:*

$$0 \le \widetilde{R}_{x_{n:1}}(D) - R_{x_{n:1}}(D) \le \frac{1}{2} \ln \left(1 + \frac{K}{\sqrt{n}} \right) \qquad \forall n \in \mathbb{N}.$$

Proof. As $\det(T_n(F)) = 1$ for all $n \in \mathbb{N}$, applying Equation (A4) and [6], Theorem 4.3, yields:

$$\lambda_{nN} \left(E \left(x_{n:1} x_{n:1}^\top \right) \right) = \frac{1}{\left\| \left(E \left(x_{n:1} x_{n:1}^\top \right) \right)^{-1} \right\|_2} = \frac{1}{\left\| \left((T_n(F))^{-1} T_n(\Lambda) \left((T_n(F))^* \right)^{-1} \right)^{-1} \right\|_2}$$

$$= \frac{1}{\|(T_n(F))^* T_n(\Lambda^{-1}) T_n(F)\|_2} \ge \frac{1}{\|(T_n(F))^*\|_2 \|T_n(\Lambda^{-1})\|_2 \|T_n(F)\|_2}$$

$$= \frac{\lambda_N(\Lambda)}{\|T_n(F)\|_2^2} \ge \frac{\lambda_N(\Lambda)}{\left(\sup_{m \in \mathbb{N}} \|T_m(F)\|_2 \right)^2} > 0 \qquad \forall n \in \mathbb{N}.$$

Thus, $\inf_{n \in \mathbb{N}} \lambda_{nN} \left(E \left(x_{n:1} x_{n:1}^\top \right) \right) > 0$. Theorem 1 and Lemma 8 prove Theorem 5. $\quad\square$

4.3. On How the Low-Complexity Coding Strategy Considered Performs under Perturbations

In this subsection, we study how the low-complexity coding strategy considered performs when it is used to encode a perturbed version, $\{z_n\}$, of a WSS, MA, or AR vector source $\{x_n\}$. Observe that if $\left\{ \left\| E \left(z_{n:1} z_{n:1}^\top \right) - E \left(x_{n:1} x_{n:1}^\top \right) \right\|_F \right\}$ is bounded, from Equation (9), we conclude that our coding strategy can also be used to optimally encode $\{z_n\}$, and the convergence speed of the rate remains unaltered.

We now present three numerical examples that show how the coding strategy considered performs in the presence of a perturbation. In all of them, $N = 2$ and:

$$E \left(z_{n:1} z_{n:1}^\top \right) = E \left(x_{n:1} x_{n:1}^\top \right) + \begin{pmatrix} 0_{2n-2 \times 2n-2} & 0_{2n-2 \times 2} \\ 0_{2 \times 2n-2} & I_2 \end{pmatrix} \qquad \forall n \in \mathbb{N}.$$

Obviously, $\left\{ \left\| E \left(z_{n:1} z_{n:1}^\top \right) - E \left(x_{n:1} x_{n:1}^\top \right) \right\|_F \right\}$ is bounded since $\left\| E \left(z_{n:1} z_{n:1}^\top \right) - E \left(x_{n:1} x_{n:1}^\top \right) \right\|_F = \sqrt{2}$ for all $n \in \mathbb{N}$. The three vector sources $\{x_n\}$ considered in our numerical examples are the zero-mean WSS vector source in [9], Section 4, the VMA(1) source in [10], Example 2.1, and the VAR(1) source in [10], Example 2.3. In [9], Section 4, the Fourier coefficients of the PSD X are:

$$\mathsf{X}_0 = \begin{pmatrix} 2.0002 & 0.7058 \\ 0.7058 & 2.0000 \end{pmatrix}, \quad \mathsf{X}_{-1} = \mathsf{X}_1^* = \begin{pmatrix} -0.3542 & 0.1016 \\ 0.1839 & -0.2524 \end{pmatrix}, \quad \mathsf{X}_{-2} = \mathsf{X}_2^* = \begin{pmatrix} -0.0923 & 0.0153 \\ 0.1490 & 0.0696 \end{pmatrix},$$

$$\mathsf{X}_{-3} = \mathsf{X}_3^* = \begin{pmatrix} -0.1443 & -0.0904 \\ 0.0602 & 0.0704 \end{pmatrix}, \quad \mathsf{X}_{-4} = \mathsf{X}_4^* = \begin{pmatrix} -0.0516 & -0.0603 \\ 0 & 0 \end{pmatrix},$$

and $\mathsf{X}_k = 0_{2 \times 2}$ with $|k| > 4$. In [10], Example 2.1, G_{-1} and Λ are given by:

$$\begin{pmatrix} -0.8 & -0.7 \\ 0.4 & -0.6 \end{pmatrix} \tag{18}$$

and:

$$\begin{pmatrix} 4 & 1 \\ 1 & 2 \end{pmatrix}, \tag{19}$$

respectively. In [10], Example 2.3, F_{-1} and Λ are given by Equations (18) and (19), respectively.

Figures 1a, 2a and 3a show $R_{x_{n:1}}(D)$ and $\widetilde{R}_{x_{n:1}}(D)$ for the three vector sources $\{x_n\}$ considered by assuming that they are Gaussian. Figures 1b, 2b and 3b show $R_{z_{n:1}}(D)$ and $\widetilde{R}_{z_{n:1}}(D)$ for these three vector sources. In Figures 1–3, $n \leq 100$ and $D = 0.001$. The figures bear the evidence of the fact that the rate of the low-complexity coding strategy considered tends to the RDF of the source even in the presence of a perturbation.
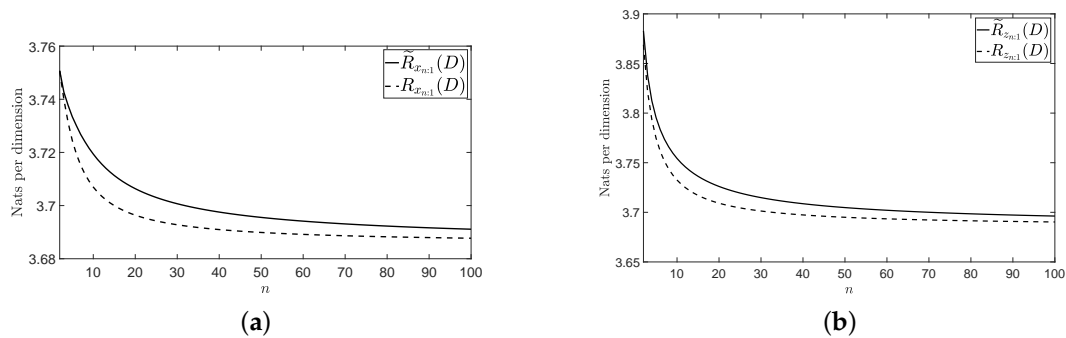
Figure 1. Rates for the considered wide sense stationary (WSS) vector source: (**a**) without perturbation and (**b**) with perturbation.
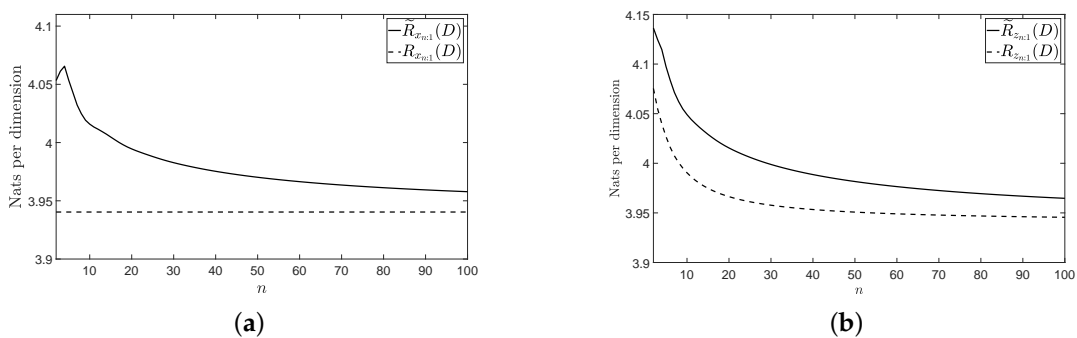
Figure 2. Rates for the considered VMA(1) source: (**a**) without perturbation and (**b**) with perturbation.
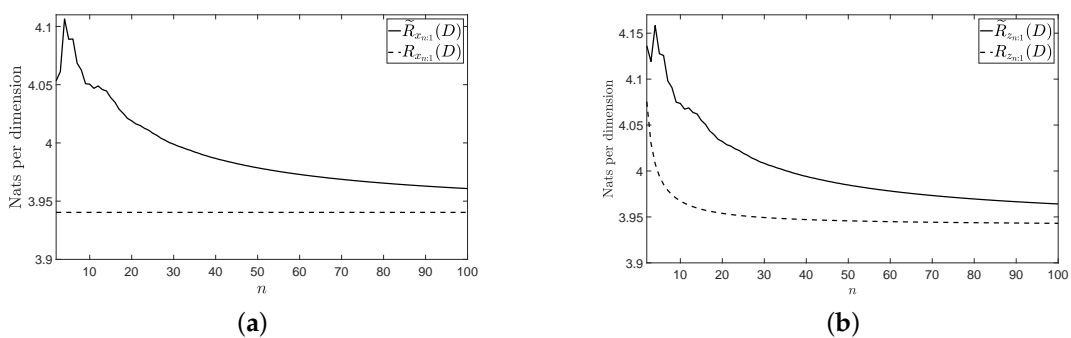
Figure 3. Rates for the considered VAR(1) source: (**a**) without perturbation and (**b**) with perturbation.

5. Conclusions

In [4], we proposed a low-complexity coding strategy to encode finite-length data blocks of any Gaussian vector source. In this paper, we proved that the convergence speed of the rate of our coding strategy is $O\left(\frac{1}{\sqrt{n}}\right)$ when it is used to encode the most relevant vector sources, namely WSS, MA, and AR vector sources. This means that the rate of our coding strategy will be close enough to the RDF of the source even if the length n of the data blocks is relatively small. Therefore, we conclude

that our coding strategy is not only low-complexity and asymptotically optimal, but also low-latency. These three features make our coding strategy very useful in practical coding applications.

Author Contributions: Authors are listed in order of their degree of involvement in the work, with the most active contributors listed first. J.G.-G. conceived the research question. All authors were involved in the research and wrote the paper. All authors have read and agreed to the published version of the manuscript.

Appendix A. Proof of Lemma 1

Proof. Fix $n \in \mathbb{N}$ and $j, k \in \{1, \ldots, n\}$. As $F \colon \mathbb{R} \to \mathbb{C}^{M \times N}$ and $G \colon \mathbb{R} \to \mathbb{C}^{N \times K}$ are continuous and 2π-periodic, $FG \colon \mathbb{R} \to \mathbb{C}^{M \times K}$ is also continuous and 2π-periodic. Applying the Parseval theorem (see, e.g., [11], p. 191) yields:

$$
\begin{aligned}
\left[[T_n(FG)]_{j,k} \right]_{r,s} &= \left[\frac{1}{2\pi} \int_0^{2\pi} e^{-(j-k)\omega \mathrm{i}} F(\omega) G(\omega) d\omega \right]_{r,s} \\
&= \frac{1}{2\pi} \int_0^{2\pi} e^{-(j-k)\omega \mathrm{i}} [F(\omega) G(\omega)]_{r,s} d\omega \\
&= \frac{1}{2\pi} \int_0^{2\pi} e^{(k-j)\omega \mathrm{i}} \sum_{t=1}^N [F(\omega)]_{r,t} [G(\omega)]_{t,s} d\omega \\
&= \sum_{t=1}^N \frac{1}{2\pi} \int_0^{2\pi} e^{(k-j)\omega \mathrm{i}} [F(\omega)]_{r,t} [G(\omega)]_{t,s} d\omega \\
&= \sum_{t=1}^N \frac{1}{2\pi} \int_0^{2\pi} [G(\omega)]_{t,s} e^{k\omega \mathrm{i}} [F(\omega)]_{r,t} e^{-j\omega \mathrm{i}} d\omega \\
&= \sum_{t=1}^N \sum_{h=-\infty}^\infty \left(\left(\frac{1}{2\pi} \int_0^{2\pi} e^{-h\omega \mathrm{i}} [G(\omega)]_{t,s} e^{k\omega \mathrm{i}} d\omega \right) \overline{\left(\frac{1}{2\pi} \int_0^{2\pi} e^{-h\omega \mathrm{i}} \overline{[F(\omega)]_{r,t} e^{-j\omega \mathrm{i}}} d\omega \right)} \right) \\
&= \sum_{t=1}^N \sum_{h=-\infty}^\infty \left(\left(\frac{1}{2\pi} \int_0^{2\pi} e^{-h\omega \mathrm{i}} [G(\omega)]_{t,s} e^{k\omega \mathrm{i}} d\omega \right) \left(\frac{1}{2\pi} \int_0^{2\pi} e^{h\omega \mathrm{i}} [F(\omega)]_{r,t} e^{-j\omega \mathrm{i}} d\omega \right) \right) \\
&= \sum_{t=1}^N \sum_{h=-\infty}^\infty \left(\left(\frac{1}{2\pi} \int_0^{2\pi} e^{-(j-h)\omega \mathrm{i}} [F(\omega)]_{r,t} d\omega \right) \left(\frac{1}{2\pi} \int_0^{2\pi} e^{-(h-k)\omega \mathrm{i}} [G(\omega)]_{t,s} d\omega \right) \right) \\
&= \sum_{t=1}^N \sum_{h=-\infty}^\infty \left(\left[\frac{1}{2\pi} \int_0^{2\pi} e^{-(j-h)\omega \mathrm{i}} F(\omega) d\omega \right]_{r,t} \left[\frac{1}{2\pi} \int_0^{2\pi} e^{-(h-k)\omega \mathrm{i}} G(\omega) d\omega \right]_{t,s} \right) \\
&= \sum_{t=1}^N \sum_{h=-\infty}^\infty \left(\left[\mathsf{F}_{j-h} \right]_{r,t} [\mathsf{G}_{h-k}]_{t,s} \right) \\
&= \sum_{h=-\infty}^\infty \sum_{t=1}^N \left(\left[\mathsf{F}_{j-h} \right]_{r,t} [\mathsf{G}_{h-k}]_{t,s} \right) \\
&= \sum_{h=-\infty}^\infty \left[\mathsf{F}_{j-h} \mathsf{G}_{h-k} \right]_{r,s} \\
&= \left[\sum_{h=-\infty}^\infty \mathsf{F}_{j-h} \mathsf{G}_{h-k} \right]_{r,s}
\end{aligned}
$$

for all $r \in \{1, \ldots, M\}$ and $s \in \{1, \ldots, K\}$. $\quad\square$

Appendix B. Proof of Lemma 2

Proof. (1) Fix $n \in \mathbb{N}$. As $\mathsf{F}_r = 0_{M \times N}$ with $|r| > p$, from Lemma 1, we have:

$$
\begin{aligned}
[T_n(F)T_n(G) - T_n(FG)]_{j,k} &= [T_n(F)T_n(G)]_{j,k} - [T_n(FG)]_{j,k} \\
&= \sum_{h=1}^{n} [T_n(F)]_{j,h}[T_n(G)]_{h,k} - \sum_{h=-\infty}^{\infty} \mathsf{F}_{j-h}\mathsf{G}_{h-k} \\
&= \sum_{h=1}^{n} \mathsf{F}_{j-h}\mathsf{G}_{h-k} - \lim_{H \to \infty}\left(\sum_{h=-H}^{H} \mathsf{F}_{j-h}\mathsf{G}_{h-k} \right) \\
&= -\lim_{H \to \infty}\left(\sum_{h=-H}^{0} \mathsf{F}_{j-h}\mathsf{G}_{h-k} + \sum_{h=n+1}^{H} \mathsf{F}_{j-h}\mathsf{G}_{h-k} \right) \\
&= -\sum_{h=j-p}^{0} \mathsf{F}_{j-h}\mathsf{G}_{h-k} - \sum_{h=n+1}^{j+p} \mathsf{F}_{j-h}\mathsf{G}_{h-k},
\end{aligned}
$$

and consequently, Equation (3) holds for all $j,k \in \{1,\ldots,n\}$. Applying Equation (3), the Schwarz inequality (see, e.g., [11], p. 15), the Parseval theorem for continuous matrix-valued functions (see, e.g., [6], p. 208), and the well-known formula for the partial sums of the arithmetic series yields:

$$
\begin{aligned}
&\|T_n(F)T_n(G) - T_n(FG)\|_F^2 \\
&= \sum_{j=1}^{n}\sum_{k=1}^{n} \left\| [T_n(F)T_n(G) - T_n(FG)]_{j,k} \right\|_F^2 \\
&= \sum_{j=1}^{p}\sum_{k=1}^{n} \left\| -\sum_{h=j-p}^{0} \mathsf{F}_{j-h}\mathsf{G}_{h-k} \right\|_F^2 + \sum_{j=n-p+1}^{n}\sum_{k=1}^{n} \left\| -\sum_{h=n+1}^{j+p} \mathsf{F}_{j-h}\mathsf{G}_{h-k} \right\|_F^2 \\
&= \sum_{j=1}^{p}\sum_{k=1}^{n} \left\| \sum_{h=j-p}^{0} \mathsf{F}_{j-h}\mathsf{G}_{h-k} \right\|_F^2 + \sum_{j=n-p+1}^{n}\sum_{k=1}^{n} \left\| \sum_{h=n+1}^{j+p} \mathsf{F}_{j-h}\mathsf{G}_{h-k} \right\|_F^2 \\
&\leq \sum_{j=1}^{p}\sum_{k=1}^{n} \left(\sum_{h=j-p}^{0} \|\mathsf{F}_{j-h}\mathsf{G}_{h-k}\|_F \right)^2 + \sum_{j=n-p+1}^{n}\sum_{k=1}^{n} \left(\sum_{h=n+1}^{j+p} \|\mathsf{F}_{j-h}\mathsf{G}_{h-k}\|_F \right)^2 \\
&\leq \sum_{j=1}^{p}\sum_{k=1}^{n} \left(\sum_{h=j-p}^{0} \|\mathsf{F}_{j-h}\|_F \|\mathsf{G}_{h-k}\|_F \right)^2 + \sum_{j=n-p+1}^{n}\sum_{k=1}^{n} \left(\sum_{h=n+1}^{j+p} \|\mathsf{F}_{j-h}\|_F \|\mathsf{G}_{h-k}\|_F \right)^2 \\
&\leq \sum_{j=1}^{p}\sum_{k=1}^{n}\sum_{h=j-p}^{0} \|\mathsf{F}_{j-h}\|_F^2 \sum_{l=j-p}^{0} \|\mathsf{G}_{l-k}\|_F^2 + \sum_{j=n-p+1}^{n}\sum_{k=1}^{n}\sum_{h=n+1}^{j+p} \|\mathsf{F}_{j-h}\|_F^2 \sum_{l=n+1}^{j+p} \|\mathsf{G}_{l-k}\|_F^2 \\
&= \sum_{j=1}^{p}\sum_{h=j-p}^{0} \|\mathsf{F}_{j-h}\|_F^2 \sum_{l=j-p}^{0}\sum_{k=1}^{n} \|\mathsf{G}_{l-k}\|_F^2 + \sum_{j=n-p+1}^{n}\sum_{h=n+1}^{j+p} \|\mathsf{F}_{j-h}\|_F^2 \sum_{l=n+1}^{j+p}\sum_{k=1}^{n} \|\mathsf{G}_{l-k}\|_F^2 \\
&\leq \sum_{j=1}^{p}\sum_{h=j-p}^{0} \|\mathsf{F}_{j-h}\|_F^2 \sum_{l=j-p}^{0} \frac{1}{2\pi}\int_0^{2\pi} \|G(\omega)\|_F^2 d\omega + \sum_{j=n-p+1}^{n}\sum_{h=n+1}^{j+p} \|\mathsf{F}_{j-h}\|_F^2 \sum_{l=n+1}^{j+p} \frac{1}{2\pi}\int_0^{2\pi} \|G(\omega)\|_F^2 d\omega \\
&= \frac{1}{2\pi}\int_0^{2\pi} \|G(\omega)\|_F^2 d\omega \left(\sum_{j=1}^{p}(p-j+1)\sum_{h=j-p}^{0} \|\mathsf{F}_{j-h}\|_F^2 + \sum_{j=n-p+1}^{n}(j+p-n)\sum_{h=n+1}^{j+p} \|\mathsf{F}_{j-h}\|_F^2 \right) \\
&\leq \frac{1}{2\pi}\int_0^{2\pi} \|G(\omega)\|_F^2 d\omega \left(\sum_{j=1}^{p}(p-j+1)\frac{1}{2\pi}\int_0^{2\pi} \|F(\omega)\|_F^2 d\omega + \sum_{j=n-p+1}^{n}(j+p-n)\frac{1}{2\pi}\int_0^{2\pi} \|F(\omega)\|_F^2 d\omega \right) \\
&= \left(\frac{1}{2\pi}\int_0^{2\pi} \|F(\omega)\|_F^2 d\omega \right)\left(\frac{1}{2\pi}\int_0^{2\pi} \|G(\omega)\|_F^2 d\omega \right)\left(\sum_{j=1}^{p}(p-j+1) + \sum_{j=n-p+1}^{n}(j+p-n) \right)
\end{aligned}
$$

$$= \left(\frac{1}{2\pi} \int_0^{2\pi} \|F(\omega)\|_F^2 d\omega \right) \left(\frac{1}{2\pi} \int_0^{2\pi} \|G(\omega)\|_F^2 d\omega \right) \left(\frac{p(p+1)}{2} + \frac{p(1+p)}{2} \right),$$

and therefore, Equation (4) is proven.

(2) Fix $n \in \mathbb{N}$. As $\mathsf{G}_r = 0_{N \times K}$ with $|r| > q$, from Lemma 1, we obtain:

$$[T_n(F)T_n(G) - T_n(FG)]_{j,k} = -\lim_{H \to \infty} \left(\sum_{h=-H}^{0} \mathsf{F}_{j-h}\mathsf{G}_{h-k} + \sum_{h=n+1}^{H} \mathsf{F}_{j-h}\mathsf{G}_{h-k} \right)$$

$$= -\sum_{h=k-q}^{0} \mathsf{F}_{j-h}\mathsf{G}_{h-k} - \sum_{h=n+1}^{k+q} \mathsf{F}_{j-h}\mathsf{G}_{h-k},$$

and hence, Equation (5) holds for all $j, k \in \{1, \ldots, n\}$. Since G is a trigonometric polynomial of degree q, G^* is also a trigonometric polynomial of degree q, where $G^*(\omega) := (G(\omega))^*$ for all $\omega \in \mathbb{R}$. Applying [6], Lemma 4.2, and Equation (4) yields:

$$\|T_n(F)T_n(G) - T_n(FG)\|_F = \left\| (T_n(F)T_n(G) - T_n(FG))^* \right\|_F$$

$$= \left\| (T_n(G))^* (T_n(F))^* - (T_n(FG))^* \right\|_F$$

$$= \left\| T_n(G^*) T_n(F^*) - T_n(G^*F^*) \right\|_F$$

$$\leq \sqrt{q(q+1) \left(\frac{1}{2\pi} \int_0^{2\pi} \|(G(\omega))^*\|_F^2 d\omega \right) \left(\frac{1}{2\pi} \int_0^{2\pi} \|(F(\omega))^*\|_F^2 d\omega \right)},$$

and thus, Equation (6) is proven.

(3) Fix $n \geq \max\{2p, 2q\}$. As $\mathsf{F}_r = 0_{M \times N}$ with $|r| > p$ and $\mathsf{G}_s = 0_{N \times K}$ with $|s| > q$, from Lemma 1, we obtain:

$$[T_n(F)T_n(G) - T_n(FG)]_{j,k} = -\lim_{H \to \infty} \left(\sum_{h=-H}^{0} \mathsf{F}_{j-h}\mathsf{G}_{h-k} + \sum_{h=n+1}^{H} \mathsf{F}_{j-h}\mathsf{G}_{h-k} \right)$$

$$= -\sum_{h=\max\{j-p,k-q\}}^{0} \mathsf{F}_{j-h}\mathsf{G}_{h-k} - \sum_{h=n+1}^{\min\{j+p,k+q\}} \mathsf{F}_{j-h}\mathsf{G}_{h-k}$$

$$= \begin{cases} -\sum_{h=\max\{j-p,k-q\}}^{0} \mathsf{F}_{j-h}\mathsf{G}_{h-k} & \text{if } j \leq p \text{ and } k \leq q, \\ 0_{M \times K} & \text{if } j \leq p \text{ and } k \geq q+1, \\ 0_{M \times K} & \text{if } p+1 \leq j \leq n-p, \\ 0_{M \times K} & \text{if } j \geq n-p+1 \text{ and } k \leq n-q, \\ -\sum_{h=n+1}^{\min\{j+p,k+q\}} \mathsf{F}_{j-h}\mathsf{G}_{h-k} & \text{if } j \geq n-p+1 \text{ and } k \geq n-q+1, \end{cases}$$

for all $j, k \in \{1, \ldots, n\}$. Observe that:

$$[T_n(F)T_n(G) - T_n(FG)]_{n-p+j,n-q+k} = -\sum_{h=n+1}^{\min\{n+j,n+k\}} \mathsf{F}_{n-p+j-h}\mathsf{G}_{h-n+q-k} = -\sum_{l=1}^{\min\{j,k\}} \mathsf{F}_{-p+j-l}\mathsf{G}_{l+q-k}$$

for all $j \in \{1, \ldots, p\}$ and $k \in \{1, \ldots, q\}$. $\square$

Appendix C. Proof of Lemma 3

Proof. (1) Since $F: \mathbb{R} \to \mathbb{C}^{N \times N}$ is continuous and 2π-periodic, $F^{-1}: \mathbb{R} \to \mathbb{C}^{N \times N}$ is also continuous and 2π-periodic, where $F^{-1}(\omega) := (F(\omega))^{-1}$ for all $\omega \in \mathbb{R}$. As:

$$\|(T_n(F))^{-1} - T_n(F^{-1})\|_F = \|(T_n(F))^{-1}(I_{nN} - T_n(F)T_n(F^{-1}))\|_F$$

$$\leq \|(T_n(F))^{-1}\|_2 \|I_{nN} - T_n(F)T_n(F^{-1})\|_F$$

$$=\|(T_n(F))^{-1}\|_2\|T_n(I_N) - T_n(F)T_n(F^{-1})\|_F$$
$$=\|(T_n(F))^{-1}\|_2\|T_n(F)T_n(F^{-1}) - T_n(FF^{-1})\|_F$$

for all $n \in \mathbb{N}$, Equation (4) proves Assertion 1 of Lemma 3.

(2) Since $F(\omega)$ is positive definite for all $\omega \in \mathbb{R}$ (or equivalently, $F(\omega)$ is Hermitian and $\lambda_N(F(\omega)) > 0$ for all $\omega \in \mathbb{R}$), $F(\omega)$ is invertible for all $\omega \in \mathbb{R}$ (or equivalently, $\det(F(\omega)) = \prod_{k=1}^{N} \lambda_k(F(\omega))$ is non-zero for all $\omega \in \mathbb{R}$), $T_n(F)$ is Hermitian, and $\lambda_{nN}(T_n(F)) \geq \inf F > 0$ for all $n \in \mathbb{N}$ (see Equation (1)). As $T_n(F)$ is positive definite for all $n \in \mathbb{N}$, $(T_n(F))^{-1}$ is also positive definite for all $n \in \mathbb{N}$. Therefore,

$$\|(T_n(F))^{-1}\|_2 = \lambda_1((T_n(F))^{-1}) = \frac{1}{\lambda_{nN}(T_n(F))} \leq \frac{1}{\inf F}$$

for all $n \in \mathbb{N}$. Assertion 2 of Lemma 3 can now be obtained from Assertion 1 of Lemma 3. $\square$

Appendix D. Proof of Lemma 4

Proof. Consider $A_n, B_n \in \mathbb{C}^{nM \times nN}$. As $V_n \otimes I_m$ is unitary, $(V_n \otimes I_m)^*$ is also unitary for all $m \in \mathbb{N}$. Consequently, since the Frobenius norm is unitarily invariant, we have

$$\|C_{A_n} - C_{B_n}\|_F = \left\|(V_n \otimes I_M)\text{diag}_{1 \leq k \leq n}\left([(V_n \otimes I_M)^*(A_n - B_n)(V_n \otimes I_N)]_{k,k}\right)(V_n \otimes I_N)^*\right\|_F$$
$$= \left\|\text{diag}_{1 \leq k \leq n}\left([(V_n \otimes I_M)^*(A_n - B_n)(V_n \otimes I_N)]_{k,k}\right)\right\|_F$$
$$\leq \|(V_n \otimes I_M)^*(A_n - B_n)(V_n \otimes I_N)\|_F = \|A_n - B_n\|_F$$

and:

$$\|A_n - C_{A_n}\|_F \leq \|A_n - C_{B_n}\|_F + \|C_{B_n} - C_{A_n}\|_F$$
$$\leq \|A_n - B_n\|_F + \|B_n - C_{B_n}\|_F + \|C_{A_n} - C_{B_n}\|_F \leq 2\|A_n - B_n\|_F + \|B_n - C_{B_n}\|_F.$$

If B_n is an $n \times n$ block circulant matrix with $M \times N$ blocks, then (see, e.g., [6], Lemma 5.1, or [12], Lemma 3) there exist $\Lambda_1, \ldots, \Lambda_n \in \mathbb{C}^{M \times N}$ such that:

$$B_n = (V_n \otimes I_M)\text{diag}(\Lambda_1, \ldots, \Lambda_n)(V_n \otimes I_N)^*.$$

Therefore,

$$C_{B_n} = (V_n \otimes I_M)\text{diag}_{1 \leq k \leq n}\left([(V_n \otimes I_M)^*B_n(V_n \otimes I_N)]_{k,k}\right)(V_n \otimes I_N)^*$$
$$= (V_n \otimes I_M)\text{diag}_{1 \leq k \leq n}\left([\text{diag}(\Lambda_1, \ldots, \Lambda_n)]_{k,k}\right)(V_n \otimes I_N)^*$$
$$= (V_n \otimes I_M)\text{diag}_{1 \leq k \leq n}(\Lambda_k)(V_n \otimes I_N)^* = B_n,$$

and combining Equations (9) and (10), we obtain Equation (11). $\square$

Appendix E. Proof of Lemma 5

Proof. Fix $n > 2p$. From [13], p. 5, we obtain:

$$[\widehat{C}_n(F)]_{j,1} = \begin{cases} F_0 & \text{if } j = 1, \\ \left(1 - \frac{j-1}{n}\right)F_{j-1} + \frac{j-1}{n}F_{j-1-n} & \text{if } j \in \{2, \ldots, n\}, \end{cases}$$

$$
= \begin{cases}
F_0 & \text{if } j = 1, \\
\left(1 - \frac{j-1}{n}\right) F_{j-1} & \text{if } j \in \{2, \ldots, p+1\}, \\
0_{M \times N} & \text{if } j \in \{p+2, \ldots, n-p\}, \\
\frac{j-1}{n} F_{j-1-n} & \text{if } j \in \{n-p+1, \ldots, n\}.
\end{cases}
$$

Hence, the Frobenius norm of the $n \times n$ block Toeplitz matrix with $M \times N$ blocks $T_n(F) - \widehat{C}_n(F)$ is given by:

$$
\begin{aligned}
\|T_n(F) - \widehat{C}_n(F)\|_F^2 &= \sum_{j,k=1}^{n} \|[T_n(F) - \widehat{C}_n(F)]_{j,k}\|_F^2 \\
&= \sum_{j=1}^{n}(n-j+1)\|[T_n(F) - \widehat{C}_n(F)]_{j,1}\|_F^2 + \sum_{k=2}^{n}(n-k+1)\|[T_n(F) - \widehat{C}_n(F)]_{1,k}\|_F^2 \\
&= \sum_{j=1}^{n}(n-j+1)\|[T_n(F)]_{j,1} - [\widehat{C}_n(F)]_{j,1}\|_F^2 + \sum_{k=2}^{n}(n-k+1)\|[T_n(F)]_{1,k} - [\widehat{C}_n(F)]_{1,k}\|_F^2 \\
&= \sum_{j=1}^{n}(n-j+1)\|F_{j-1} - [\widehat{C}_n(F)]_{j,1}\|_F^2 + \sum_{k=2}^{n}(n-k+1)\|F_{1-k} - [\widehat{C}_n(F)]_{n+2-k,1}\|_F^2 \\
&= \sum_{j=2}^{n}(n-j+1)\|F_{j-1} - [\widehat{C}_n(F)]_{j,1}\|_F^2 + \sum_{h=2}^{n}(h-1)\|F_{h-1-n} - [\widehat{C}_n(F)]_{h,1}\|_F^2 \\
&= \sum_{j=2}^{n}\left((n-j+1)\|F_{j-1} - [\widehat{C}_n(F)]_{j,1}\|_F^2 + (j-1)\|F_{j-1-n} - [\widehat{C}_n(F)]_{j,1}\|_F^2\right) \\
&= \sum_{j=2}^{p+1}\left((n-j+1)\|F_{j-1} - [\widehat{C}_n(F)]_{j,1}\|_F^2 + (j-1)\|F_{j-1-n} - [\widehat{C}_n(F)]_{j,1}\|_F^2\right) \\
&\quad + \sum_{j=n-p+1}^{n}\left((n-j+1)\|F_{j-1} - [\widehat{C}_n(F)]_{j,1}\|_F^2 + (j-1)\|F_{j-1-n} - [\widehat{C}_n(F)]_{j,1}\|_F^2\right) \quad \text{(A1)} \\
&= \sum_{j=2}^{p+1}\left((n-j+1)\left\|F_{j-1} - \left(1 - \frac{j-1}{n}\right)F_{j-1}\right\|_F^2 + (j-1)\left\|-\left(1-\frac{j-1}{n}\right)F_{j-1}\right\|_F^2\right) \\
&\quad + \sum_{j=n-p+1}^{n}\left((n-j+1)\left\|-\frac{j-1}{n}F_{j-1-n}\right\|_F^2 + (j-1)\left\|F_{j-1-n} - \frac{j-1}{n}F_{j-1-n}\right\|_F^2\right) \\
&= \sum_{j=2}^{p+1}\left((n-j+1)\left(\frac{j-1}{n}\right)^2\|F_{j-1}\|_F^2 + (j-1)\left(\frac{n-j+1}{n}\right)^2\|F_{j-1}\|_F^2\right) \\
&\quad + \sum_{j=n-p+1}^{n}\left((n-j+1)\left(\frac{j-1}{n}\right)^2\|F_{j-1-n}\|_F^2 + (j-1)\left(\frac{n-j+1}{n}\right)^2\|F_{j-1-n}\|_F^2\right) \\
&= \sum_{j=2}^{p+1}\frac{(n-j+1)(j-1)}{n}\|F_{j-1}\|_F^2 + \sum_{j=n-p+1}^{n}\frac{(n-j+1)(j-1)}{n}\|F_{j-1-n}\|_F^2 \\
&= \sum_{k=1}^{p}k\frac{n-k}{n}\|F_k\|_F^2 + \sum_{h=1}^{p}h\frac{n-h}{n}\|F_{-h}\|_F^2 \\
&= \sum_{k=1}^{p}k\frac{n-k}{n}\left(\|F_k\|_F^2 + \|F_{-k}\|_F^2\right).
\end{aligned}
$$

Applying [6], Lemma 5.4, we have:

$$
[C_n(F)]_{j,1} = \begin{cases}
F_0 & \text{if } j = 1, \\
F_{j-1} & \text{if } j \in \{2, \ldots, p+1\}, \\
0_{M \times N} & \text{if } j \in \{p+2, \ldots, n-p\}, \\
F_{j-1-n} & \text{if } j \in \{n-p+1, \ldots, n\}.
\end{cases}
$$

Consequently, the Frobenius norm of the $n \times n$ block Toeplitz matrix with $M \times N$ blocks $T_n(F) - C_n(F)$ is given by:

$$
\begin{aligned}
\|T_n(F) - C_n(F)\|_F^2 &= \sum_{j=2}^{p+1} \left((n-j+1)\|F_{j-1} - [C_n(F)]_{j,1}\|_F^2 + (j-1)\|F_{j-1-n} - [C_n(F)]_{j,1}\|_F^2 \right) \\
&\quad + \sum_{j=n-p+1}^{n} \left((n-j+1)\|F_{j-1} - [C_n(F)]_{j,1}\|_F^2 + (j-1)\|F_{j-1-n} - [C_n(F)]_{j,1}\|_F^2 \right) \\
&= \sum_{j=2}^{p+1} (j-1)\| - F_{j-1}\|_F^2 + \sum_{j=n-p+1}^{n} (n-j+1)\| - F_{j-1-n}\|_F^2 \\
&= \sum_{k=1}^{p} k\|F_k\|_F^2 + \sum_{h=1}^{p} h\|F_{-h}\|_F^2 \\
&= \sum_{k=1}^{p} k\left(\|F_k\|_F^2 + \|F_{-k}\|_F^2 \right).
\end{aligned}
\tag{A2}
$$

Equations (A1) and (A2) prove Lemma 5. $\quad\square$

Appendix F. Proof of Lemma 7

Proof. From Equation (12), we obtain:

$$
x_{n:1} = T_n(G)w_{n:1} \qquad \forall n \in \mathbb{N},
$$

with $G(\omega) = I_N + \sum_{k=1}^{q} e^{-k\omega i} G_{-k}$ for all $\omega \in \mathbb{R}$. Therefore, applying [6], Lemma 4.2, yields:

$$
\{E\left(x_{n:1}x_{n:1}^*\right)\} = \{T_n(G)E\left(w_{n:1}w_{n:1}^*\right)(T_n(G))^*\} = \{T_n(G)T_n(\Lambda)T_n(G^*)\} = \{T_n(G\Lambda)T_n(G^*)\}. \tag{A3}
$$

Hence, using Equation (9), we have:

$$
\begin{aligned}
\left\| E\left(x_{n:1}x_{n:1}^*\right) - C_{E\left(x_{n:1}x_{n:1}^*\right)} \right\|_F &\leq 2\|E\left(x_{n:1}x_{n:1}^*\right) - T_n(G\Lambda G^*)\|_F + \left\| T_n(G\Lambda G^*) - C_{T_n(G\Lambda G^*)} \right\|_F \\
&= 2\|T_n(G\Lambda)T_n(G^*) - T_n(G\Lambda G^*)\|_F + \left\| T_n(G\Lambda G^*) - \widehat{C}_n(G\Lambda G^*) \right\|_F
\end{aligned}
$$

for all $n \in \mathbb{N}$. Thus, to finish the proof, we only need to show that $\{\|T_n(G\Lambda)T_n(G^*) - T_n(G\Lambda G^*)\|_F\}$ and $\left\{ \left\| T_n(G\Lambda G^*) - \widehat{C}_n(G\Lambda G^*) \right\|_F \right\}$ are bounded. As $G\Lambda$ and G^* are trigonometric polynomials, from Equation (7), we obtain that $\{\|T_n(G\Lambda)T_n(G^*) - T_n(G\Lambda G^*)\|_F\}$ is bounded. Since $G\Lambda G^*$ is a trigonometric polynomial, applying Lemma 5, we conclude that $\left\{ \left\| T_n(G\Lambda G^*) - \widehat{C}_n(G\Lambda G^*) \right\|_F \right\}$ is bounded. $\quad\square$

Appendix G. Proof of Lemma 8

Proof. As Λ is positive definite, if $\omega \in \mathbb{R}$ and $v \in \mathbb{C}^{N \times 1}$, then:

$$
v^*(F(\omega))^{-1}\Lambda((F(\omega))^{-1})^*v = (((F(\omega))^{-1})^*v)^*\Lambda((F(\omega))^{-1})^*v = (((F(\omega))^*)^{-1}v)^*\Lambda((F(\omega))^*)^{-1}v > 0
$$

whenever $((F(\omega))^*)^{-1}v \neq 0_{N \times 1}$, or equivalently, $v \neq 0_{N \times 1}$. Since $(F(\omega))^{-1}\Lambda((F(\omega))^{-1})^*$ is positive definite for all $\omega \in \mathbb{R}$, we have:

$$
\|(T_n(F^{-1}\Lambda(F^{-1})^*))^{-1}\|_2 \leq \frac{1}{\inf(F^{-1}\Lambda(F^{-1})^*)} \qquad \forall n \in \mathbb{N}.
$$

From Equation (13), we obtain:

$$
w_{n:1} = T_n(F)x_{n:1} \qquad \forall n \in \mathbb{N}.
$$

Consequently,
$$\{T_n(\Lambda)\} = \{E\left(w_{n:1}w_{n:1}^*\right)\} = \{T_n(F)E\left(x_{n:1}x_{n:1}^*\right)(T_n(F))^*\}.$$

Therefore, as $\det(T_n(F)) = 1$ for all $n \in \mathbb{N}$, we have:

$$\{E\left(x_{n:1}x_{n:1}^*\right)\} = \{(T_n(F))^{-1}T_n(\Lambda)((T_n(F))^*)^{-1}\}. \tag{A4}$$

Hence, applying Equation (11) and [6], Lemma 4.2, yields:

$$\left\|E\left(x_{n:1}x_{n:1}^*\right) - C_{E\left(x_{n:1}x_{n:1}^*\right)}\right\|_F$$

$$\leq 2\|E\left(x_{n:1}x_{n:1}^*\right) - C_n(F^{-1}\Lambda(F^{-1})^*)\|_F$$

$$\leq 2\left(\|(T_n(F))^{-1}T_n(\Lambda)((T_n(F))^*)^{-1} - T_n(F^{-1}\Lambda(F^{-1})^*)\|_F + \|T_n(F^{-1}\Lambda(F^{-1})^*) - C_n(F^{-1}\Lambda(F^{-1})^*)\|_F\right)$$

$$\leq 2\left(\|(T_n(F))^{-1}T_n(\Lambda)((T_n(F))^*)^{-1}\|_2\|I_{nN} - (T_n(F))^*(T_n(\Lambda))^{-1}T_n(F)T_n(F^{-1}\Lambda(F^{-1})^*)\|_F\right.$$

$$\left. + \|T_n(F^{-1}\Lambda(F^{-1})^*)\|_2\|I_{nN} - (T_n(F^{-1}\Lambda(F^{-1})^*))^{-1}C_n(F^{-1}\Lambda(F^{-1})^*)\|_F\right)$$

$$\leq 2\left(\|(T_n(F))^{-1}\|_2\|T_n(\Lambda)\|_2\|((T_n(F))^*)^{-1}\|_2\|I_{nN} - T_n(F^*)T_n(\Lambda^{-1})T_n(F)T_n(F^{-1}\Lambda(F^{-1})^*)\|_F\right.$$

$$\left. + \|T_n(F^{-1}\Lambda(F^{-1})^*)\|_2\|(C_n(F^{-1}\Lambda(F^{-1})^*))^{-1} - (T_n(F^{-1}\Lambda(F^{-1})^*))^{-1}\|_F\|C_n(F^{-1}\Lambda(F^{-1})^*)\|_2\right)$$

$$= 2\left(\|(T_n(F))^{-1}\|_2\lambda_1(\Lambda)\|((T_n(F))^{-1})^*\|_2\|I_{nN} - T_n(F^*)T_n(\Lambda^{-1}F)T_n(F^{-1}\Lambda(F^{-1})^*)\|_F\right.$$

$$\left. + \|T_n(F^{-1}\Lambda(F^{-1})^*)\|_2\|C_n((F^{-1}\Lambda(F^{-1})^*)^{-1}) - (T_n(F^{-1}\Lambda(F^{-1})^*))^{-1}\|_F\|C_n(F^{-1}\Lambda(F^{-1})^*)\|_2\right)$$

$$\leq 2\|T_n(F^{-1}\Lambda(F^{-1})^*)\|_2\left(\|(T_n(F))^{-1}\|_2^2\lambda_1(\Lambda)\|(T_n(F^{-1}\Lambda(F^{-1})^*))^{-1} - T_n(F^*)T_n(\Lambda^{-1}F)\|_F\right.$$

$$\left. + \|C_n(F^{-1}\Lambda(F^{-1})^*)\|_2\|C_n((F^{-1}\Lambda(F^*)^{-1})^{-1}) - (T_n(F^{-1}\Lambda(F^{-1})^*))^{-1}\|_F\right)$$

$$\leq 2\|T_n(F^{-1}\Lambda(F^{-1})^*)\|_2\left(\|(T_n(F))^{-1}\|_2^2\lambda_1(\Lambda)\left(\|(T_n(F^{-1}\Lambda(F^{-1})^*))^{-1} - T_n(F^*\Lambda^{-1}F)\|_F\right.\right.$$

$$\left. + \|T_n(F^*\Lambda^{-1}F) - T_n(F^*)T_n(\Lambda^{-1}F)\|_F\right) + \|C_n(F^{-1}\Lambda(F^{-1})^*)\|_2\left(\|C_n(F^*\Lambda^{-1}F) - T_n(F^*\Lambda^{-1}F)\|_F\right.$$

$$\left.\left. + \|T_n(F^*\Lambda^{-1}F) - (T_n(F^{-1}\Lambda(F^{-1})^*))^{-1}\|_F\right)\right)$$

$$= 2\|T_n(F^{-1}\Lambda(F^{-1})^*)\|_2\left(\|(T_n(F))^{-1}\|_2^2\lambda_1(\Lambda)\|T_n(F^*)T_n(\Lambda^{-1}F) - T_n(F^*\Lambda^{-1}F)\|_F\right.$$

$$+ \left(\|(T_n(F))^{-1}\|_2^2\lambda_1(\Lambda) + \|C_n(F^{-1}\Lambda(F^{-1})^*)\|_2\right)\|(T_n(F^{-1}\Lambda(F^{-1})^*))^{-1} - T_n(F^*\Lambda^{-1}F)\|_F$$

$$\left. + \|C_n(F^{-1}\Lambda(F^{-1})^*)\|_2\|T_n(F^*\Lambda^{-1}F) - C_n(F^*\Lambda^{-1}F)\|_F\right)$$

$$\leq 2\|T_n(F^{-1}\Lambda(F^{-1})^*)\|_2\left(\|(T_n(F))^{-1}\|_2^2\lambda_1(\Lambda)\|T_n(F^*)T_n(\Lambda^{-1}F) - T_n(F^*\Lambda^{-1}F)\|_F\right.$$

$$+ \left(\|(T_n(F))^{-1}\|_2^2\lambda_1(\Lambda) + \|C_n(F^{-1}\Lambda(F^{-1})^*)\|_2\right)\|(T_n(F^{-1}\Lambda(F^{-1})^*))^{-1}\|_2\|T_n(F^*\Lambda^{-1}F)\|_2$$

$$\times \|(T_n(F^*\Lambda^{-1}F))^{-1} - T_n(F^{-1}\Lambda(F^{-1})^*)\|_F$$

$$\left. + \|C_n(F^{-1}\Lambda(F^{-1})^*)\|_2\|T_n(F^*\Lambda^{-1}F) - C_n(F^*\Lambda^{-1}F)\|_F\right)$$

$$= 2\|T_n(F^{-1}\Lambda(F^{-1})^*)\|_2\left(\|(T_n(F))^{-1}\|_2^2\lambda_1(\Lambda)\|T_n(F^*)T_n(\Lambda^{-1}F) - T_n(F^*\Lambda^{-1}F)\|_F\right.$$

$$+ \left(\|(T_n(F))^{-1}\|_2^2\lambda_1(\Lambda) + \|C_n(F^{-1}\Lambda(F^{-1})^*)\|_2\right)\|(T_n(F^{-1}\Lambda(F^{-1})^*))^{-1}\|_2\|T_n(F^*\Lambda^{-1}F)\|_2$$

$$\times \|(T_n(F^*\Lambda^{-1}F))^{-1} - T_n((F^*\Lambda^{-1}F)^{-1})\|_F$$

$$\left. + \|C_n(F^{-1}\Lambda(F^{-1})^*)\|_2\|T_n(F^*\Lambda^{-1}F) - C_n(F^*\Lambda^{-1}F)\|_F\right)$$

for all $n \in \mathbb{N}$. Thus, to finish the proof, we only need to show that $\{\|T_n(F^{-1}\Lambda(F^{-1})^*)\|_2\}$, $\{\|T_n(F^*)T_n(\Lambda^{-1}F) - T_n(F^*\Lambda^{-1}F)\|_F\}$, $\{\|C_n(F^{-1}\Lambda(F^{-1})^*)\|_2\}$, $\{\|T_n(F^*\Lambda^{-1}F)\|_2\}$,

$\{\|(T_n(F^*\Lambda^{-1}F))^{-1} - T_n((F^*\Lambda^{-1}F)^{-1})\|_F\}$, and $\{\|T_n(F^*\Lambda^{-1}F) - C_n(F^*\Lambda^{-1}F)\|_F\}$ are bounded. From [6], Theorem 4.3, we obtain that $\{\|T_n(F^{-1}\Lambda(F^{-1})^*)\|_2\}$ and $\{\|T_n(F^*\Lambda^{-1}F)\|_2\}$ are bounded. Applying [6], Lemma 5.2, we have that $\{\|C_n(F^{-1}\Lambda(F^{-1})^*)\|_2\}$ is bounded. Since F^* and $\Lambda^{-1}F$ are trigonometric polynomials, from Equation (7), we obtain that $\{\|T_n(F^*)T_n(\Lambda^{-1}F) - T_n(F^*\Lambda^{-1}F)\|_F\}$ is bounded. As $F^*\Lambda^{-1}F$ is a trigonometric polynomial, applying Lemma 5, we have that $\{\|T_n(F^*\Lambda^{-1}F) - C_n(F^*\Lambda^{-1}F)\|_F\}$ is bounded. Since $(F(\omega))^{-1}\Lambda((F(\omega))^{-1})^* = (F(\omega))^{-1}\Lambda((F(\omega))^*)^{-1}$ is positive definite for all $\omega \in \mathbb{R}$, $((F(\omega))^{-1}\Lambda((F(\omega))^*)^{-1})^{-1} = (F(\omega))^*\Lambda^{-1}F(\omega)$ is also positive definite for all $\omega \in \mathbb{R}$, and consequently, from Equation (8), we conclude that $\{\|(T_n(F^*\Lambda^{-1}F))^{-1} - T_n((F^*\Lambda^{-1}F)^{-1})\|_F\}$ is bounded. $\square$

Appendix H. A Statistical Signal Processing Application on Filtering WSS Vector Processes

Consider a zero-mean WSS M-dimensional vector process $\{x_n\}$. Let Y be the PSD of a zero-mean WSS N-dimensional vector process $\{y_n\}$ with $\inf Y > 0$. Assume that those two processes are jointly WSS with joint PSD Z, that is $Z \colon \mathbb{R} \to \mathbb{C}^{M \times N}$ is a continuous 2π-periodic function satisfying that $\{E(x_{n:1}y_{n:1}^*)\} = \{T_n(Z)\}$.

For every $n \in \mathbb{N}$, if $\widetilde{x}_{n:1}$ is an estimation of $x_{n:1}$ from $y_{n:1}$ of the form:

$$\widetilde{x}_{n:1} = Wy_{n:1} \tag{A5}$$

with $W \in \mathbb{C}^{nM \times nN}$, the MSE per sample is:

$$\frac{\text{MSE}(W)}{n} = \frac{E\left(\|x_{n:1} - \widetilde{x}_{n:1}\|_2^2\right)}{n},$$

and the minimum MSE (MMSE) is given by $\text{MMSE} = \text{MSE}(W_0)$, where W_0 is the Wiener filter, i.e.,

$$W_0 = E(x_{n:1}y_{n:1}^*)\left(E(y_{n:1}y_{n:1}^*)\right)^{-1} = T_n(Z)(T_n(Y))^{-1}.$$

In [13], Equation (6), it was shown that there is no difference in the MSE per sample for large enough n if we substitute the optimal filter W_0 by W_C, where $W_C = \widehat{C}_n(Z)(\widehat{C}_n(Y))^{-1}$, that is,

$$\lim_{n \to \infty} \left(\frac{\text{MSE}(W_C)}{n} - \frac{\text{MMSE}}{n}\right) = 0.$$

Obviously, the computational complexity of the operation (A5) is notably reduced when applying this substitution and the FFT algorithm is used. Specifically, the computational complexity is reduced from $O(n^2)$ to $O(n \log n)$.

We here study the convergence speed of the sequence $\left\{\frac{\text{MSE}(W_C)}{n} - \frac{\text{MMSE}}{n}\right\}$ (i.e., how fast this sequence tends to zero) by assuming that Y and Z are trigonometric polynomials. Applying [13], p. 11, and Lemma 5, we conclude that there exists $K \in [0, \infty)$ such that:

$$0 \le \frac{\text{MSE}(W_C)}{n} - \frac{\text{MMSE}}{n} \le \frac{\sqrt{M}\sigma_1(Z)}{\inf Y}\left(1 + \frac{\sup Y}{\inf Y}\right)\left(\frac{\|\widehat{C}_n(Z) - T_n(Z)\|_F}{\sqrt{n}} + \frac{\sigma_1(Z)}{\inf Y}\frac{\|\widehat{C}_n(Y) - T_n(Y)\|_F}{\sqrt{n}}\right)$$

$$\le \frac{K}{\sqrt{n}} \qquad \forall n \in \mathbb{N},$$

where $\sigma_1(Z) = \sup_{\omega \in [0,2\pi]} \|Z(\omega)\|_2$ and $\sup Y = \max_{\omega \in [0,2\pi]} \lambda_1(Y(\omega))$. Therefore,

$$\frac{\text{MSE}(W_C)}{n} - \frac{\text{MMSE}}{n} = O\left(\frac{1}{\sqrt{n}}\right). \tag{A6}$$

Equation (A6) was proven in [2] for the case $M = N = 1$.

References

1. Kolmogorov, A.N. On the Shannon theory of information transmission in the case of continuous signals. *IRE Trans. Inf. Theory* **1956**, *IT-2*, 102–108. [CrossRef]

2. Pearl, J. On coding and filtering stationary signals by discrete Fourier transforms. *IEEE Trans. Inf. Theory* **1973**, *19*, 229–232. [CrossRef]

3. Gutiérrez-Gutiérrez, J.; Zárraga-Rodríguez, M.; Insausti, X. Upper bounds for the rate distortion function of finite-length data blocks of Gaussian WSS sources. *Entropy* **2017**, *19*, 554. [CrossRef]

4. Zárraga-Rodríguez, M.; Gutiérrez-Gutiérrez, J.; Insausti, X. A low-complexity and asymptotically optimal coding strategy for Gaussian vectors sources. *Entropy* **2019**, *21*, 965. [CrossRef]

5. Gray, R.M. Toeplitz and circulant matrices: A review. *Found. Trends Commun. Inf. Theory* **2006**, *2*, 155–239. [CrossRef]

6. Gutiérrez-Gutiérrez, J.; Crespo, P.M. Block Toeplitz matrices: Asymptotic results and applications. *Found. Trends Commun. Inf. Theory* **2011**, *8*, 179–257. [CrossRef]

7. Gutiérrez-Gutiérrez, J. A modified version of the Pisarenko method to estimate the power spectral density of any asymptotically wide sense stationary vector process. *Appl. Math. Comput.* **2019**, *362*, 124526. [CrossRef]

8. Gutiérrez-Gutiérrez, J.; Zárraga-Rodríguez, M.; Villar-Rosety, F.M.; Insausti, X. Rate-distortion function upper bounds for Gaussian vectors and their applications in coding AR sources. *Entropy* **2018**, *20*, 399. [CrossRef] [PubMed]

9. Gutiérrez-Gutiérrez, J.; Iglesias, I.; Podhorski, A. Geometric MMSE for one-sided and two-sided vector linear predictors: From the finite-length case to the infinite-length case. *Signal Process.* **2011**, *91*, 2237–2245. [CrossRef]

10. Reinsel, G.C. *Elements of Multivariate Time Series Analysis*; Springer: Berlin/Heidelberg, Germany, 1993.

11. Rudin, W. *Principles of Mathematical Analysis*; McGraw-Hill: New York, NY, USA, 1976.

12. Gutiérrez-Gutiérrez, J.; Crespo, P.M. Asymptotically equivalent sequences of matrices and multivariate ARMA processes. *IEEE Trans. Inf. Theory* **2011**, *57*, 5444–5454. [CrossRef]

13. Gutiérrez-Gutiérrez, J.; Zárraga-Rodríguez, M.; Insausti, X.; Hogstad, B.O. On the complexity reduction of coding WSS vector processes by using a sequence of block circulant matrices. *Entropy* **2017**, *19*, 95. [CrossRef]