Applied Mathematics and Computation 445 (2023) 127824

Contents lists available at ScienceDirect

Applied Mathematics and Computation

journal homepage: www.elsevier.com/locate/amc

Necessary and sufficient conditions for AR vector processes to be stationary: Applications in information theory and in statistical signal processing

Jesús Gutiérrez-Gutiérrez*, Íñigo Barasoain-Echepare, Marta Zárraga-Rodríguez, Xabier Insausti

University of Navarra, Tecnun School of Engineering, Manuel de Lardizábal 13, 20018 San Sebastián, Spain

A R T I C L E I N F O

Article history:
Received 5 July 2022
Revised 21 September 2022
Accepted 26 December 2022
Available online 14 January 2023

2020 MSC:
62M10
15B05
94A17
62M15

Keywords:
Autoregressive (AR) stationary vector process
Block Toeplitz matrix
Differential entropy rate
The Pisarenko spectral estimation method

A B S T R A C T

As the correlation matrices of stationary vector processes are block Toeplitz, autoregressive (AR) vector processes are non-stationary. However, in the literature, an AR vector process of finite order is said to be "*stationary*" if it satisfies the so-called stationarity condition (i.e., if the spectral radius of the associated companion matrix is less than one). Since the term "*stationary*" is used for such an AR vector process, its correlation matrices should "*somehow approach*" the correlation matrices of a stationary vector process, but the meaning of "*somehow approach*" has not been mathematically established in the literature. In the present paper we give necessary and sufficient conditions for AR vector processes to be "*stationary*". These conditions show in which sense the correlation matrices of an AR "*stationary*" vector process asymptotically behave like block Toeplitz matrices. Applications in information theory and in statistical signal processing of these necessary and sufficient conditions are also given.

1. Introduction

As the correlation matrices of wide sense stationary (WSS) vector processes are block Toeplitz, autoregressive (AR) vector processes are not WSS (and thus not stationary). However, in the literature, an AR vector process of finite order is said to be "*stationary*" if it satisfies the so-called stationarity condition (i.e., if the spectral radius of the associated companion matrix is less than one (see, e.g., [1, Section 2.2.1])). Since the term "*stationary*" is used for such an AR vector process, its correlation matrices should "*somehow approach*" the correlation matrices of a WSS vector process, but the meaning of "*somehow approach*" has not been mathematically established in the literature. In this paper we formally establish the meaning of "*somehow approach*" by using the definition of asymptotically WSS (AWSS) process (which was given in [2, p. 225] for 1-dimensional (or scalar) processes and in [3, Definition 7.1] for vector processes). In other words, in this paper we prove that a necessary and sufficient condition for an AR vector process to be "*stationary*" is to be AWSS.

* Corresponding author.
E-mail addresses: jgutierrez@tecnun.es (J. Gutiérrez-Gutiérrez), ibarasoaine@tecnun.es (Í. Barasoain-Echepare), mzarraga@tecnun.es (M. Zárraga-Rodríguez), xinsausti@tecnun.es (X. Insausti).

https://doi.org/10.1016/j.amc.2022.127824

Unlike the stationarity condition, the definition of AWSS process provides information about the asymptotic behaviour of the correlation matrices of the process. Another concept regarding the asymptotic behaviour of the correlation matrices of a stochastic (or random) process was given in [4, p. 223], namely, the concept of asymptotically stationary correlation structure. In the present paper, we also obtain a sufficient condition for stochastic vector processes to be AWSS and to have an asymptotically stationary correlation structure. Moreover, we show that this sufficient condition is also a necessary condition to be AWSS if the stochastic vector processes are AR. Consequently, in this paper we prove that a necessary condition for an AR vector process to be "*stationary*" is to have an asymptotically stationary correlation structure.

The necessary and sufficient conditions obtained here find application in practical situations involving computations with large correlation matrices of AR "*stationary*" vector processes. As examples of such practical applications we obtain a novel result in information theory and another one in statistical signal processing. Specifically:

1. We compute the differential entropy rate of any proper Gaussian AR "*stationary*" vector process.
2. We extend to AR "*stationary*" vector processes the Pisarenko spectral estimation method given in [5, Theorem 2] for WSS scalar processes.

The rest of the paper is organized as follows. In Section 2 some preliminary results on block Toeplitz matrices are given. In Section 3 we present a sufficient condition for stochastic vector processes to be AWSS and to have an asymptotically stationary correlation structure. In Section 4 we give necessary and sufficient conditions for AR vector processes to be "stationary". In Section 5 applications in information theory and in statistical signal processing of these necessary and sufficient conditions are presented. Finally, Section 6 is the conclusions section.

2. Preliminary results

2.1. The Barnett factorization

We first review a result on block Hankel matrices called the Barnett factorization (see [6, Theorem 4.27]).

Theorem 1. *Let*

$$\{\mathsf{H}_k\}_{k\in\mathbb{Z}} = \left\{\frac{1}{2\pi}\int_0^{2\pi} \mathrm{e}^{-k\omega\mathrm{i}} H(\omega)d\omega\right\}_{k\in\mathbb{Z}}$$

be the sequence of Fourier coefficients of a continuous 2π-periodic function $H:\mathbb{R}\to\mathbb{C}^{N\times N}$. Assume that H_0 is the $N\times N$ identity matrix $I_N=(\delta_{j,k})_{j,k=1}^N$. Suppose that H_k is the $N\times N$ zero matrix $0_{N\times N}$ for all $k\in\mathbb{N}$. If $n\in\mathbb{N}$ then

$$Z_n(H) = K_n(H)(\Phi_n(H))^n,$$

where

$$\Phi_n(H) = \left(\begin{array}{c|c} 0_{(n-1)N\times N} & I_{(n-1)N} \\ \hline -\mathsf{H}_{-n} & -\mathsf{H}_{-(n-1)} \ \cdots \ -\mathsf{H}_{-1}\end{array}\right),$$

and $Z_n(H)$ and $K_n(H)$ are the block Hankel matrices

$$Z_n(H) = \begin{pmatrix} 0_{N\times N} & \cdots & 0_{N\times N} & 0_{N\times N} & -H_{-n} \\ 0_{N\times N} & \cdots & 0_{N\times N} & -H_{-n} & -H_{-(n-1)} \\ 0_{N\times N} & \cdots & -H_{-n} & -H_{-(n-1)} & -H_{-(n-2)} \\ \vdots & \iddots & \vdots & \vdots & \vdots \\ -H_{-n} & \cdots & -H_{-3} & -H_{-2} & -H_{-1}\end{pmatrix}$$

and

$$K_n(H) = \begin{pmatrix} H_{-(n-1)} & \cdots & H_{-2} & H_{-1} & I_N \\ H_{-(n-2)} & \cdots & H_{-1} & I_N & 0_{N\times N} \\ H_{-(n-3)} & \cdots & I_N & 0_{N\times N} & 0_{N\times N} \\ \vdots & \iddots & \vdots & \vdots & \vdots \\ I_N & \cdots & 0_{N\times N} & 0_{N\times N} & 0_{N\times N}\end{pmatrix},$$

respectively.

In [6] the proof of Theorem 1 is left as an exercise for the reader. For the convenience of the reader, we prove it in Appendix A. We now review a result on the characteristic polynomial of the matrix $\Phi_n(H)$ (see, e.g., [6, Theorem 4.23] and [7, p. 14]).

Theorem 2. *If $\{\mathsf{H}_k\}_{k\in\mathbb{Z}}$ is as in Theorem 1, $n\in\mathbb{N}$, and $\tau\in\mathbb{C}$, then*

$$\det(\tau I_{nN} - \Phi_n(H)) = \det\left(\tau^n I_N + \sum_{k=1}^n \tau^{n-k}\mathsf{H}_{-k}\right).$$

The following result provides two simple properties of the matrix $J_n \otimes I_N$, where J_n is the $n \times n$ backward identity matrix (that is, $J_n = (\delta_{j+k,n+1})_{j,k=1}^n$) and $\otimes$ denotes the Kronecker product, i.e.,

$$J_n \otimes I_N = \begin{pmatrix} 0_{N\times N} & \cdots & 0_{N\times N} & 0_{N\times N} & I_N \\ 0_{N\times N} & \cdots & 0_{N\times N} & I_N & 0_{N\times N} \\ 0_{N\times N} & \cdots & I_N & 0_{N\times N} & 0_{N\times N} \\ \vdots & \iddots & \vdots & \vdots & \vdots \\ I_N & \cdots & 0_{N\times N} & 0_{N\times N} & 0_{N\times N} \end{pmatrix}.$$

Lemma 3.

1. *$J_n \otimes I_N$ is involutory.*
2. *If $A \in \mathbb{C}^{nN\times nN}$ then*

$$[A(J_n \otimes I_N)]_{j,k} = [A]_{j,n+1-k} \in \mathbb{C}^{N\times N}$$

and

$$[(J_n \otimes I_N)A]_{j,k} = [A]_{n+1-j,k} \in \mathbb{C}^{N\times N}$$

for all $j,k \in \{1,\dots,n\}$.

Proof. (1) $(J_n \otimes I_N)^2 = J_n^2 \otimes I_N^2 = I_n \otimes I_N = I_{nN}$.
(2) If $j,k \in \{1,\dots,n\}$ then

$$\begin{aligned}[A(J_n \otimes I_N)]_{j,k} &= \sum_{h=1}^{n} [A]_{j,h}[J_n \otimes I_N]_{h,k} = \sum_{h=1}^{n} [A]_{j,h}[J_n]_{h,k} I_N \\ &= \sum_{h=1}^{n} [J_n]_{h,k}[A]_{j,h} I_N = \sum_{h=1}^{n} [J_n]_{h,k}[A]_{j,h} = [A]_{j,n+1-k}\end{aligned}$$

and

$$[(J_n \otimes I_N)A]_{j,k} = \sum_{h=1}^{n} [J_n \otimes I_N]_{j,h}[A]_{h,k} = \sum_{h=1}^{n} [J_n]_{j,h} I_N [A]_{h,k} = \sum_{h=1}^{n} [J_n]_{j,h}[A]_{h,k} = [A]_{n+1-j,k}.$$

□

We finish this section by presenting the Barnett factorization for block Toeplitz matrices, which will be proved by using Theorems 1 and 2, and Lemma 3.

Theorem 4. *If $\{\mathsf{H}_k\}_{k\in\mathbb{Z}}$ is as in Theorem 1 and $n \in \mathbb{N}$ then*

$$T_n(H)(\Psi_n(H))^n = B_n(H), \tag{1}$$

where

$$\Psi_n(H) = \left(\begin{array}{c|c} -\mathsf{H}_{-1} \ \dots \ -\mathsf{H}_{-(n-1)} & -\mathsf{H}_{-n} \\ \hline I_{(n-1)N} & 0_{(n-1)N\times N} \end{array}\right),$$

and $T_n(H)$ and $B_n(H)$ are the block Toeplitz matrices

$$T_n(H) = (\mathsf{H}_{j-k})_{j,k=1}^n$$

and

$$B_n(H) = \begin{pmatrix} -\mathsf{H}_{-n} & 0_{N\times N} & 0_{N\times N} & \cdots & 0_{N\times N} \\ -\mathsf{H}_{-(n-1)} & -\mathsf{H}_{-n} & 0_{N\times N} & \cdots & 0_{N\times N} \\ -\mathsf{H}_{-(n-2)} & -\mathsf{H}_{-(n-1)} & -\mathsf{H}_{-n} & \cdots & 0_{N\times N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -\mathsf{H}_{-1} & -\mathsf{H}_{-2} & -\mathsf{H}_{-3} & \cdots & -\mathsf{H}_{-n} \end{pmatrix},$$

respectively. Moreover,

$$\det(\tau I_{nN} - \Psi_n(H)) = \det\left(\tau^n I_N + \sum_{k=1}^{n} \tau^{n-k}\mathsf{H}_{-k}\right) \qquad \forall \tau \in \mathbb{C}. \tag{2}$$

Proof. From Theorem 1 and Lemma 3 we have

$$T_n(H)(\Psi_n(H))^n = T_n(H)(J_n \otimes I_N)((J_n \otimes I_N)\Psi_n(H)(J_n \otimes I_N))^n(J_n \otimes I_N)$$

$$
\begin{aligned}
&= K_n(H)\left((J_n \otimes I_N) \left(\begin{array}{c|c} -\mathrm{H}_{-n} & -\mathrm{H}_{-(n-1)} \ \dots \ -\mathrm{H}_{-1} \\ \hline 0_{(n-1)N\times N} & J_{n-1}\otimes I_N \end{array} \right) \right)^n (J_n \otimes I_N) \\
&= K_n(H)(\Phi_n(H))^n (J_n \otimes I_N) \\
&= Z_n(H)(J_n \otimes I_N) \\
&= B_n(H).
\end{aligned}
$$

As $\Phi_n(H) = (J_n \otimes I_N)\Psi_n(H)(J_n \otimes I_N)$ and $J_n \otimes I_N$ is involutory, $\Phi_n(H)$ and $\Psi_n(H)$ are similar. Consequently, $\Phi_n(H)$ and $\Psi_n(H)$ have the same characteristic polynomial. Therefore, applying Theorem 2 yields Equality (2). □

2.2. On the companion matrix

We begin this section by giving a result on natural powers of block lower triangular matrices.

Lemma 5. *Let $m, n \in \mathbb{N}$ with $n > m$. Consider $A \in \mathbb{C}^{m\times m}$, $B \in \mathbb{C}^{(n-m)\times m}$, and $C \in \mathbb{C}^{(n-m)\times(n-m)}$. If $q \in \mathbb{N}$ then*

$$
\left(\begin{array}{c|c} A & 0_{m\times(n-m)} \\ \hline B & C \end{array} \right)^q = \left(\begin{array}{c|c} A^q & 0_{m\times(n-m)} \\ \hline \sum_{k=0}^{q-1} C^k B A^{q-1-k} & C^q \end{array} \right). \tag{3}
$$

Proof. We proceed by induction on q. Since $C^0 B A^0 = I_{n-m} B I_m = B$, Equality (3) is true for $q = 1$. If we now assume that Equality (3) is true for certain $q \in \mathbb{N}$, then

$$
\begin{aligned}
\left(\begin{array}{c|c} A & 0_{m\times(n-m)} \\ \hline B & C \end{array} \right)^{q+1} &= \left(\begin{array}{c|c} A^q & 0_{m\times(n-m)} \\ \hline \sum_{k=0}^{q-1} C^k B A^{q-1-k} & C^q \end{array} \right) \left(\begin{array}{c|c} A & 0_{m\times(n-m)} \\ \hline B & C \end{array} \right) \\
&= \left(\begin{array}{c|c} A^{q+1} & 0_{m\times(n-m)} \\ \hline \sum_{k=0}^{q-1} C^k B A^{q-k} + C^q B & C^{q+1} \end{array} \right) \\
&= \left(\begin{array}{c|c} A^{q+1} & 0_{m\times(n-m)} \\ \hline \sum_{k=0}^{q} C^k B A^{q-k} & C^{q+1} \end{array} \right).
\end{aligned}
$$

□

If H is an $N \times N$ trigonometric polynomial of degree $p \in \mathbb{N}$ of the form

$$
H(\omega) = I_N + \sum_{k=1}^{p} \mathrm{e}^{-k\omega\mathrm{i}} \mathrm{H}_{-k} \qquad \forall \omega \in \mathbb{R}, \tag{4}
$$

the matrix

$$
\Psi_p(H) = \left(\begin{array}{c|c} -\mathrm{H}_{-1} \ \dots \ -\mathrm{H}_{-(p-1)} & -\mathrm{H}_{-p} \\ \hline I_{(p-1)N} & 0_{(p-1)N\times N} \end{array} \right)
$$

is called the companion matrix of H. We finish this section by presenting a result on the spectral radius of the companion matrix $\rho(\Psi_p(H))$, which will be proved by using Theorem 4 and Lemma 5.

Theorem 6. *If H is as in Equality (4), then the following assertions are equivalent:*

1. $\rho(\Psi_p(H)) < 1$.
2. $\det\left(I_N + \sum_{k=1}^{p} z^k \mathrm{H}_{-k}\right) \neq 0$ *when* $|z| \leq 1$.
3. *H is invertible (that is, $\det(H(\omega)) \neq 0$ for all $\omega \in \mathbb{R}$) and $\{\|(\Psi_n(H))^n\|_2\}$ is bounded with $\|\cdot\|_2$ being the spectral norm.*
4. *$\{\|(T_n(H))^{-1}\|_2\}$ is bounded.*
5. *H is invertible and $\{\|(T_n(H))^{-1} - T_n(H^{-1})\|_F\}$ is bounded with $\|\cdot\|_F$ being the Frobenius norm.*

Proof. From Equality (2) we obtain

$$
\det(\tau I_{pN} - \Psi_p(H)) = \det\left(\tau^p I_N + \sum_{k=1}^{p} \tau^{p-k} \mathrm{H}_{-k}\right) = (\tau^p)^N \det\left(I_N + \sum_{k=1}^{p} \tau^{-k} \mathrm{H}_{-k}\right) \tag{5}
$$

for all $\tau \in \mathbb{C} \setminus \{0\}$.

(1)⇔ (2) $\rho(\Psi_p(H)) < 1$ is equivalent to $\det(\tau I_{pN} - \Psi_p(H)) \neq 0$ when $|\tau| \geq 1$. Hence, applying Equality (5), $\rho(\Psi_p(H)) < 1$ if and only if $\det\left(I_N + \sum_{k=1}^{p} \tau^{-k}\mathrm{H}_{-k}\right) \neq 0$ when $|\tau| \geq 1$, or equivalently, $\det\left(I_N + \sum_{k=1}^{p} z^k \mathrm{H}_{-k}\right) \neq 0$ when $0 < |z| \leq 1$.

(1)⇒ (3) The proof falls into three parts.

Part 1: $\exists n_0 \in \mathbb{N}$ such that $\|(\Psi_p(H))^n\|_2 < \left(\frac{1+\rho(\Psi_p(H))}{2}\right)^n$ $\forall n \geq n_0$.

From [8, p. 299] we have

$$\rho(\Psi_p(H)) \leq \sqrt[n]{\|(\Psi_p(H))^n\|_2} \qquad \forall n \in \mathbb{N}$$

and

$$\lim_{n\to\infty} \sqrt[n]{\|(\Psi_p(H))^n\|_2} = \rho(\Psi_p(H)).$$

Consequently, there exists $n_0 \in \mathbb{N}$ such that

$$\sqrt[n]{\|(\Psi_p(H))^n\|_2} - \rho(\Psi_p(H)) = \left|\sqrt[n]{\|(\Psi_p(H))^n\|_2} - \rho(\Psi_p(H))\right| < \frac{1-\rho(\Psi_p(H))}{2} \qquad \forall n \geq n_0.$$

Therefore,

$$0 \leq \sqrt[n]{\|(\Psi_p(H))^n\|_2} < \frac{1+\rho(\Psi_p(H))}{2} \qquad \forall n \geq n_0.$$

Part 2: $\{\|(\Psi_n(H))^n\|_2\}$ is bounded.

If $n > p+1$ then $\Psi_n(H)$ can be written as

$$\Psi_n(H) = \left(\begin{array}{c|c|c} \multicolumn{2}{c|}{\Psi_p(H)} & 0_{pN\times(n-p)N} \\ \hline 0_{N\times(p-1)N} \mid I_N & \multicolumn{2}{c}{0_{N\times(n-p)N}} \\ \hline 0_{(n-p-1)N\times pN} & I_{(n-p-1)N} & 0_{(n-p-1)N\times N} \end{array}\right).$$

Thus, applying Lemma 5 yields

$$\begin{aligned}
\|(\Psi_n(H))^n\|_2 &\leq \|(\Psi_p(H))^n\|_2 + \left\|\left(\begin{array}{c|c} \multicolumn{2}{c}{0_{N\times(n-p)N}} \\ \hline I_{(n-p-1)N} & 0_{(n-p-1)N\times N} \end{array}\right)^n\right\|_2 \\
&\quad + \left\|\sum_{k=0}^{n-1}\left(\begin{array}{c|c} \multicolumn{2}{c}{0_{N\times(n-p)N}} \\ \hline I_{(n-p-1)N} & 0_{(n-p-1)N\times N} \end{array}\right)^k \left(\begin{array}{c|c} 0_{N\times(p-1)N} & I_N \\ \hline \multicolumn{2}{c}{0_{(n-p-1)N\times pN}} \end{array}\right)(\Psi_p(H))^{n-1-k}\right\|_2 \\
&\leq \|(\Psi_p(H))^n\|_2 + \left\|\left(\begin{array}{c|c} \multicolumn{2}{c}{0_{N\times(n-p)N}} \\ I_{(n-p-1)N} & 0_{(n-p-1)N\times N} \end{array}\right)\right\|_2^n \\
&\quad + \sum_{k=0}^{n-1}\left\|\left(\begin{array}{c|c} \multicolumn{2}{c}{0_{N\times(n-p)N}} \\ \hline I_{(n-p-1)N} & 0_{(n-p-1)N\times N} \end{array}\right)\right\|_2^k \left\|\left(\begin{array}{c|c} 0_{N\times(p-1)N} & I_N \\ \hline \multicolumn{2}{c}{0_{(n-p-1)N\times pN}} \end{array}\right)\right\|_2 \left\|(\Psi_p(H))^{n-1-k}\right\|_2 \\
&= \|(\Psi_p(H))^n\|_2 + 1 + \sum_{k=0}^{n-1}\left\|(\Psi_p(H))^{n-1-k}\right\|_2 \\
&= 1 + \sum_{k=0}^{n}\left\|(\Psi_p(H))^k\right\|_2 \\
&< 1 + \sum_{k=0}^{n_0-1}\left\|(\Psi_p(H))^k\right\|_2 + \sum_{k=n_0}^{n}\left(\frac{1+\rho(\Psi_p(H))}{2}\right)^k \\
&= 1 + \sum_{k=0}^{n_0-1}\left\|(\Psi_p(H))^k\right\|_2 + \frac{\left(\frac{1+\rho(\Psi_p(H))}{2}\right)^{n_0} - \left(\frac{1+\rho(\Psi_p(H))}{2}\right)^{n+1}}{1-\frac{1+\rho(\Psi_p(H))}{2}}
\end{aligned}$$

for all $n > \max\{p+1, n_0\}$. Since

$$\lim_{n\to\infty} 1 + \sum_{k=0}^{n_0-1}\left\|(\Psi_p(H))^k\right\|_2 + \frac{\left(\frac{1+\rho(\Psi_p(H))}{2}\right)^{n_0} - \left(\frac{1+\rho(\Psi_p(H))}{2}\right)^{n+1}}{1-\frac{1+\rho(\Psi_p(H))}{2}} = 1 + \sum_{k=0}^{n_0-1}\left\|(\Psi_p(H))^k\right\|_2 + \frac{\left(\frac{1+\rho(\Psi_p(H))}{2}\right)^{n_0}}{1-\frac{1+\rho(\Psi_p(H))}{2}},$$

$\left\{1 + \sum_{k=0}^{n_0-1}\left\|(\Psi_p(H))^k\right\|_2 + \frac{\left(\frac{1+\rho(\Psi_p(H))}{2}\right)^{n_0} - \left(\frac{1+\rho(\Psi_p(H))}{2}\right)^{n+1}}{1-\frac{1+\rho(\Psi_p(H))}{2}}\right\}$ is bounded, and hence, $\{\|(\Psi_n(H))^n\|_2\}$ is bounded.

Part 3: H is invertible.

As $|e^{\omega i}| = 1$ for all $\omega \in \mathbb{R}$, $e^{\omega i}$ is not an eigenvalue of $\Psi_p(H)$ with $\omega \in \mathbb{R}$, and consequently, from Equality (5) we obtain

$$0 \neq \det\left(e^{\omega i} I_{pN} - \Psi_p(H)\right) = \left(\left(e^{\omega i}\right)^p\right)^N \det\left(I_N + \sum_{k=1}^{p}\left(e^{\omega i}\right)^{-k} \mathrm{H}_{-k}\right) = e^{pN\omega i}\det(H(\omega))$$

for all $\omega \in \mathbb{R}$.

(3)⇒ (4) Since $\{\det(T_n(H))\} = \{1\}$, $T_n(H)$ is invertible for all $n \in \mathbb{N}$. As $H^{-1} : \mathbb{R} \to \mathbb{C}^{N\times N}$, which is defined as $H^{-1}(\omega) = (H(\omega))^{-1}$ for all $\omega \in \mathbb{R}$, is continuous and 2π-periodic, applying [3, Lemma 5.2], [3, Lemma 5.4], and Equality (1) yields

$$\begin{aligned}
\|(T_n(H))^{-1}\|_2 &\le \|(T_n(H))^{-1} - C_n(H^{-1})\|_2 + \|C_n(H^{-1})\|_2 \\
&\le \|(T_n(H))^{-1}(C_n(H^{-1}))^{-1} - I_{nN}\|_2 \|C_n(H^{-1})\|_2 + \|C_n(H^{-1})\|_2 \\
&= \|C_n(H^{-1})\|_2 (1 + \|(T_n(H))^{-1}C_n(H) - I_{nN}\|_2) \\
&\le \sigma_1(H^{-1})(1 + \|(T_n(H))^{-1}C_n(H) - I_{nN}\|_2) \\
&= \sigma_1(H^{-1})(1 + \|(T_n(H))^{-1}(C_n(H) - T_n(H))\|_2) \\
&= \sigma_1(H^{-1})(1 + \|(T_n(H))^{-1}(-B_n(H))\|_2) \\
&= \sigma_1(H^{-1})(1 + \|(T_n(H))^{-1}B_n(H)\|_2) \\
&= \sigma_1(H^{-1})(1 + \|(\Psi_n(H))^n\|_2) \qquad \forall n > 2p
\end{aligned}$$

with $\sigma_1(H^{-1}) = \sup_{\omega\in[0,2\pi]} \|(H(\omega))^{-1}\|_2 < \infty$ and

$$C_n(G) = (V_n \otimes I_N)\left(\delta_{j,k} G\left(\frac{2\pi(k-1)}{n}\right)\right)_{j,k=1}^{n} (V_n \otimes I_N)^*$$

for any continuous 2π-periodic function $G : \mathbb{R} \to \mathbb{C}^{N\times N}$, where V_n is the $n \times n$ Fourier unitary matrix:

$$V_n = \left(\frac{1}{\sqrt{n}} e^{-\frac{2\pi(j-1)(k-1)}{n}\mathrm{i}}\right)_{j,k=1}^{n}. \tag{6}$$

(4)⇒(1) The proof falls into three parts.

Part 1: $\rho(\Psi_p(H)) \le 1$.

If $n \ge p$, from Equality (2) we have

$$\begin{aligned}
\det(\tau I_{nN} - \Psi_n(H)) &= \det\left(\tau^n I_N + \sum_{k=1}^{n} \tau^{n-k} \mathrm{H}_{-k}\right) = \det\left(\tau^n I_N + \sum_{k=1}^{p} \tau^{n-k} \mathrm{H}_{-k}\right) \\
&= (\tau^{n-p})^N \det\left(\tau^p I_N + \sum_{k=1}^{p} \tau^{p-k} \mathrm{H}_{-k}\right) = (\tau^{n-p})^N \det(\tau I_{pN} - \Psi_p(H)) \qquad \forall \tau \in \mathbb{C},
\end{aligned}$$

and therefore, $\rho(\Psi_n(H)) = \rho(\Psi_p(H))$. Thus, applying Equality (1) yields

$$\begin{aligned}
(\rho(\Psi_p(H)))^n &= (\rho(\Psi_n(H)))^n \le \|(\Psi_n(H))^n\|_2 = \|(T_n(H))^{-1}B_n(H)\|_2 \\
&\le \|(T_n(H))^{-1}\|_2 \|B_n(H)\|_2 = \|(T_n(H))^{-1}\|_2 \|B_p(H)\|_2 \qquad \forall n \ge p.
\end{aligned}$$

Since $\{(\rho(\Psi_p(H)))^n\}$ is bounded, $\rho(\Psi_p(H)) \le 1$.

Part 2: If $\rho(\Psi_p(H)) = 1$ then $\exists\, \omega_0 \in \mathbb{R}$ such that $\det(H(\omega_0)) = 0$.

If $\rho(\Psi_p(H)) = 1$ then there exists $\omega_0 \in [0, 2\pi)$ such that $e^{\omega_0 \mathrm{i}}$ is an eigenvalue of $\Psi_p(H)$. Hence, from Equality (5) we obtain

$$0 = \det(e^{\omega_0\mathrm{i}} I_{pN} - \Psi_p(H)) = \left(\left(e^{\omega_0\mathrm{i}}\right)^p\right)^N \det\left(I_N + \sum_{k=1}^{p} \left(e^{\omega_0\mathrm{i}}\right)^{-k} \mathrm{H}_{-k}\right) = e^{pN\omega_0\mathrm{i}} \det(H(\omega_0)).$$

As $e^{pN\omega_0\mathrm{i}} \neq 0$, $\det(H(\omega_0)) = 0$.

Part 3: H is invertible.

From [3, Lemma 4.5], for each $n \in \mathbb{N}$ there is a matrix $A_n \in \mathbb{C}^{nN\times pN}$ such that

$$T_n(HH^*) = \left(T_n(H) \mid A_n\right)\left(T_n(H) \mid A_n\right)^* = T_n(H)(T_n(H))^* + A_n A_n^*,$$

where $H^* : \mathbb{R} \to \mathbb{C}^{N\times N}$ is the continuous 2π-periodic function defined as $H^*(\omega) = (H(\omega))^*$ for all $\omega \in \mathbb{R}$. Since $H(\omega)(H(\omega))^*$ is positive semidefinite for all $\omega \in \mathbb{R}$, applying [9, Proposition 3] we have that $T_n(HH^*)$ is positive semidefinite for all $n \in \mathbb{N}$ and

$$\inf_{n\in\mathbb{N}} \lambda_{nN}(T_n(HH^*)) = \min_{\omega\in[0,2\pi]} \lambda_N(H(\omega)(H(\omega))^*).$$

For each $n \in \mathbb{N}$ there is an eigenvector u_n of $T_n(HH^*)$ associated with $\lambda_{nN}(T_n(HH^*))$ satisfying that $\|u_n\|_F = 1$. Therefore,

$$\begin{aligned}
\lambda_{nN}(T_n(HH^*)) &= \lambda_{nN}(T_n(HH^*))\|u_n\|_F^2 = \lambda_{nN}(T_n(HH^*))u_n^* u_n = u_n^*(\lambda_{nN}(T_n(HH^*))u_n) = u_n^* T_n(HH^*)u_n \\
&= u_n^* T_n(H)(T_n(H))^* u_n + u_n^* A_n A_n^* u_n = \|(T_n(H))^* u_n\|_F^2 + \|A_n^* u_n\|_F^2 \ge \|(T_n(H))^* u_n\|_F^2 \\
&= \left(\frac{\|(T_n(H))^* u_n\|_F}{\|u_n\|_F}\right)^2 = \left(\frac{\|(T_n(H))^* u_n\|_F}{\|((T_n(H))^*)^{-1}(T_n(H))^* u_n\|_F}\right)^2 = \left(\frac{1}{\frac{\|((T_n(H))^{-1})^*(T_n(H))^* u_n\|_F}{\|(T_n(H))^* u_n\|_F}}\right)^2
\end{aligned}$$

$$\geq \left(\frac{1}{\|((T_n(H))^{-1})^*\|_2}\right)^2 = \frac{1}{\|(T_n(H))^{-1}\|_2^2} \geq \frac{1}{\left(\sup_{m\in\mathbb{N}}\|(T_m(H))^{-1}\|_2\right)^2} \qquad \forall n\in\mathbb{N},$$

and thus,

$$\min_{\omega\in\mathbb{R}}\lambda_N(H(\omega)(H(\omega))^*) = \min_{\omega\in[0,2\pi]}\lambda_N(H(\omega)(H(\omega))^*) = \inf_{n\in\mathbb{N}}\lambda_{nN}(T_n(HH^*)) \geq \frac{1}{\left(\sup_{m\in\mathbb{N}}\|(T_m(H))^{-1}\|_2\right)^2}.$$

Hence,

$$\begin{aligned}\det(H(\omega))\det((H(\omega))^*) &= \det(H(\omega)(H(\omega))^*) = \prod_{k=1}^{N}\lambda_k(H(\omega)(H(\omega))^*)\\ &\geq (\lambda_N(H(\omega)(H(\omega))^*))^N \geq \frac{1}{\left(\sup_{m\in\mathbb{N}}\|(T_m(H))^{-1}\|_2\right)^{2N}} > 0 \qquad \forall\omega\in\mathbb{R}.\end{aligned}$$

(4)⇒ (5) It is a direct consequence of [10, Lemma 3].
(5)⇒(4) From [3, Theorem 4.3] or [11, Theorem 4.1] we conclude that

$$\|(T_n(H))^{-1}\|_2 \leq \|(T_n(H))^{-1} - T_n(H^{-1})\|_2 + \|T_n(H^{-1})\|_2 \leq \|(T_n(H))^{-1} - T_n(H^{-1})\|_F + \sigma_1(H^{-1})$$

for all $n\in\mathbb{N}$. □

3. A sufficient condition for stochastic vector processes to be AWSS and to have an asymptotically stationary correlation structure

We first extend to vector processes the concept of asymptotically stationary correlation structure given by Berger in [4, p. 223].

Definition 7. A constant mean stochastic vector process $\{y_n\}$ has an asymptotically stationary correlation structure if $\{E(y_n y_{n+k}^*)\}$ is convergent for all $k\in\mathbb{Z}$, where E stands for expectation and $*$ denotes conjugate transpose.

We can now give a sufficient condition for stochastic vector processes to be AWSS and to have an asymptotically stationary correlation structure.

Theorem 8. *Consider a continuous 2π-periodic function $Y:\mathbb{R}\to\mathbb{C}^{N\times N}$ and a constant mean stochastic N-dimensional process $\{y_n\}$. Suppose that $\{\|E(y_{n:1}y_{n:1}^*) - T_n(Y)\|_F\}$ is bounded with*

$$y_{n:1} = \begin{pmatrix} y_n\\ y_{n-1}\\ y_{n-2}\\ \vdots\\ y_1\end{pmatrix} \qquad \forall n\in\mathbb{N}.$$

Then

1. *$\{y_n\}$ is AWSS with (asymptotic) power spectral density (PSD) Y, i.e., $\{\|E(y_{n:1}y_{n:1}^*)\|_2\}$ and $\{\|T_n(Y)\|_2\}$ are bounded, and*

$$\lim_{n\to\infty}\frac{\|E(y_{n:1}y_{n:1}^*) - T_n(Y)\|_F}{\sqrt{n}} = 0.$$

2. *$\{y_n\}$ has an asymptotically stationary correlation structure. In fact,*

$$\lim_{n\to\infty}E(y_n y_{n+k}^*) = \mathsf{Y}_k \qquad \forall k\in\mathbb{Z},$$

where $\{\mathsf{Y}_k\}_{k\in\mathbb{Z}}$ is the sequence of Fourier coefficients of Y.

Proof. (1) As $\{\|E(y_{n:1}y_{n:1}^*) - T_n(Y)\|_F\}$ is bounded,

$$\lim_{n\to\infty}\frac{\|E(y_{n:1}y_{n:1}^*) - T_n(Y)\|_F}{\sqrt{n}} = 0.$$

From [3, Theorem 4.3] or [11, Theorem 4.1], $\{\|T_n(Y)\|_2\}$ is bounded. Thus, since

$$\|E(y_{n:1}y_{n:1}^*)\|_2 \leq \|E(y_{n:1}y_{n:1}^*) - T_n(Y)\|_2 + \|T_n(Y)\|_2 \leq \|E(y_{n:1}y_{n:1}^*) - T_n(Y)\|_F + \|T_n(Y)\|_2$$

for all $n\in\mathbb{N}$, $\{\|E(y_{n:1}y_{n:1}^*)\|_2\}$ is also bounded.

(2) We first consider the case in which $k \geq 0$. If $n > k$ then

$$\begin{aligned}\sum_{h=k+1}^{n} \|E(y_{h-k}y_h^*) - \mathsf{Y}_k\|_F^2 &= \sum_{h=k+1}^{n} \|[E(y_{n:1}y_{n:1}^*)]_{k+n-h+1,n-h+1} - [T_n(Y)]_{k+n-h+1,n-h+1}\|_F^2 \\ &= \sum_{h=k+1}^{n} \|[E(y_{n:1}y_{n:1}^*) - T_n(Y)]_{k+n-h+1,n-h+1}\|_F^2 \\ &\leq \|E(y_{n:1}y_{n:1}^*) - T_n(Y)\|_F^2.\end{aligned}$$

Consequently, $\left\{\sum_{h=k+1}^{n} \left\|E\left(y_{h-k}y_h^*\right) - \mathsf{Y}_k\right\|_F^2\right\}_{n>k}$ is bounded. As $\left\{\sum_{h=k+1}^{n} \left\|E\left(y_{h-k}y_h^*\right) - \mathsf{Y}_k\right\|_F^2\right\}_{n>k}$ is bounded and monotonically increasing, it is convergent. Therefore,

$$\lim_{h\to\infty} \left\|E\left(y_{h-k}y_h^*\right) - \mathsf{Y}_k\right\|_F^2 = 0,$$

or equivalently,

$$\lim_{h\to\infty} \left\|E\left(y_{h-k}y_h^*\right) - \mathsf{Y}_k\right\|_F = 0.$$

Hence,

$$\lim_{n\to\infty} E\left(y_n y_{n+k}^*\right) = \lim_{h\to\infty} E\left(y_{h-k}y_h^*\right) = \mathsf{Y}_k.$$

We now consider the case in which $k < 0$. From [9, Proposition 2], $Y(\omega)$ is Hermitian for all $\omega \in \mathbb{R}$. Thus,

$$\begin{aligned}\lim_{n\to\infty} E\left(y_n y_{n+k}^*\right) &= \lim_{n\to\infty} (E(y_{n+k}y_n^*))^* = \left(\lim_{n\to\infty} E(y_{n+k}y_n^*)\right)^* = \left(\lim_{h\to\infty} E\left(y_h y_{h+(-k)}^*\right)\right)^* = \mathsf{Y}_{-k}^* \\ &= \left(\frac{1}{2\pi}\int_0^{2\pi} \mathrm{e}^{k\omega \mathrm{i}} Y(\omega) d\omega\right)^* = \frac{1}{2\pi}\int_0^{2\pi} \left(\mathrm{e}^{k\omega \mathrm{i}} Y(\omega)\right)^* d\omega = \frac{1}{2\pi}\int_0^{2\pi} \mathrm{e}^{-k\omega \mathrm{i}} Y(\omega) d\omega = \mathsf{Y}_k.\end{aligned}$$

□

In Section 4 we show that the boundedness of $\{\|E\left(y_{n:1}y_{n:1}^*\right) - T_n(Y)\|_F\}$ is a necessary and sufficient condition for an AR vector process $\{y_n\}$ to be AWSS with PSD Y. It should be mentioned that a necessary condition for a stochastic vector process $\{y_n\}$ to be AWSS with PSD Y is that $\{E\left(y_{n:1}y_{n:1}^*\right) - T_n(Y)\}$ be zero-distributed (see [12, Theorem 2.16]), but we will not use this necessary condition in the present paper.

4. Necessary and sufficient conditions for AR vector processes to be stationary

The following lemma shows that the correlation matrices of an AR vector process can be written in terms of block Toeplitz matrices.

Lemma 9. *Let $\{\mathsf{H}_k\}_{k\in\mathbb{Z}}$ be as in Theorem 1. Suppose that $\{w_n\}$ is a zero-mean N-dimensional process with $\left\{E\left(w_{n:1}w_{n:1}^*\right)\right\} = \{T_n(\Lambda)\}$, where Λ is an $N \times N$ positive definite matrix. Let $\{y_n\}$ be the zero-mean AR N-dimensional process given by*

$$y_n = w_n - \sum_{k=1}^{n-1} \mathsf{H}_{-k} y_{n-k} \qquad \forall n \in \mathbb{N}. \tag{7}$$

Then

1. $\left\{E\left(y_{n:1}y_{n:1}^*\right)\right\} = \left\{(T_n(H))^{-1}T_n(\Lambda)((T_n(H))^{-1})^*\right\}$.
2. $\inf_{n\in\mathbb{N}} \lambda_{nN}\left(E\left(y_{n:1}y_{n:1}^*\right)\right) > 0$ *with* $\lambda_{nN}\left(E\left(y_{n:1}y_{n:1}^*\right)\right)$ *being the smallest eigenvalue of* $E\left(y_{n:1}y_{n:1}^*\right)$ *for all* $n \in \mathbb{N}$.

Proof. (1) Equality (7) can be rewritten as

$$I_N y_n + \sum_{k=1}^{n-1} \mathsf{H}_{-k} y_{n-k} = w_n \qquad \forall n \in \mathbb{N}.$$

Consequently,

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

and therefore,

$$T_n(H) y_{n:1} y_{n:1}^* (T_n(H))^* = T_n(H) y_{n:1} (T_n(H) y_{n:1})^* = w_{n:1} w_{n:1}^* \qquad \forall n \in \mathbb{N}.$$

Hence,

$$\{T_n(H)E(y_{n:1}y_{n:1}^*)(T_n(H))^*\} = \{E(w_{n:1}w_{n:1}^*)\} = \{T_n(\Lambda)\}.$$

As $\{\det(T_n(H))\} = \{1\}$, $T_n(H)$ is invertible for all $n \in \mathbb{N}$, and thus,

$$\{E(y_{n:1}y_{n:1}^*)\} = \left\{(T_n(H))^{-1}T_n(\Lambda)((T_n(H))^*)^{-1}\right\} = \left\{(T_n(H))^{-1}T_n(\Lambda)((T_n(H))^{-1})^*\right\}.$$

(2) Applying [3, Theorem 4.3] or [11, Theorem 4.1] yields

$$\begin{aligned}\lambda_{nN}(E(y_{n:1}y_{n:1}^*)) &= \frac{1}{\lambda_1\left(\left(E\left(y_{n:1}y_{n:1}^*\right)\right)^{-1}\right)} = \frac{1}{\left\|\left(E\left(y_{n:1}y_{n:1}^*\right)\right)^{-1}\right\|_2} = \frac{1}{\|(T_n(H))^*(T_n(\Lambda))^{-1}T_n(H)\|_2} \\ &\geq \frac{1}{\|(T_n(H))^*\|_2\|(T_n(\Lambda))^{-1}\|_2\|T_n(H)\|_2} = \frac{\lambda_{nN}(T_n(\Lambda))}{\|T_n(H)\|_2^2} = \frac{\lambda_N(\Lambda)}{\|T_n(H)\|_2^2} \geq \frac{\lambda_N(\Lambda)}{(\sigma_1(H))^2} > 0\end{aligned}$$

for all $n \in \mathbb{N}$, where $\sigma_1(H) = \sup_{\omega\in[0,2\pi]}\|H(\omega)\|_2 < \infty$. $\square$

If H is a trigonometric polynomial of degree p, then $\{y_n\}$ is called an AR vector process of (finite) order p or an AR(p) vector process. In the literature an AR(p) vector process is said to be "*stationary*" if $\rho(\Psi_p(H)) < 1$, or equivalently, if $\det\left(I_N + \sum_{k=1}^{p} z^k \mathrm{H}_{-k}\right) \neq 0$ when $|z| \leq 1$ (see, e.g., [1, Section 2.2.1])). We now show that a necessary and sufficient condition for an AR(p) vector process to be "*stationary*" is to be AWSS.

Theorem 10. *Let $\{y_n\}$ be as in Lemma 9. Suppose that H is as in Equality (4). Then the following assertions are equivalent:*

1. $\rho(\Psi_p(H)) < 1$.
2. $\det\left(I_N + \sum_{k=1}^{p} z^k\mathrm{H}_{-k}\right) \neq 0$ *when* $|z| \leq 1$.
3. $\{\|(T_n(H))^{-1}\|_2\}$ *is bounded.*
4. *H is invertible and* $\{\|E\left(y_{n:1}y_{n:1}^*\right) - T_n(H^{-1}\Lambda(H^{-1})^*)\|_F\}$ *is bounded.*
5. *H is invertible and $\{y_n\}$ is AWSS with PSD* $H^{-1}\Lambda(H^{-1})^*$.
6. $\{y_n\}$ *is AWSS.*
7. $\{\|E\left(y_{n:1}y_{n:1}^*\right)\|_2\}$ *is bounded (that is,* $\sup_{n\in\mathbb{N}}\lambda_1\left(E\left(y_{n:1}y_{n:1}^*\right)\right) < \infty$*).*

Proof. From Theorem 6 the three first assertions are equivalent.

(3)⇒(4) By Theorem 6, H is invertible. Applying Lemma 9 and [3, Lemma 4.2] yields

$$\begin{aligned}&\|E(y_{n:1}y_{n:1}^*) - T_n(H^{-1}\Lambda(H^{-1})^*)\|_F \\ &= \|(T_n(H))^{-1}T_n(\Lambda)((T_n(H))^{-1})^* - T_n(H^{-1}\Lambda(H^{-1})^*)\|_F \\ &\leq \|(T_n(H))^{-1}\|_2\|T_n(\Lambda)((T_n(H))^{-1})^* - T_n(H)T_n(H^{-1}\Lambda(H^{-1})^*)\|_F \\ &\leq \|(T_n(H))^{-1}\|_2\left(\|T_n(\Lambda)((T_n(H))^{-1})^* - T_n(\Lambda(H^{-1})^*)\|_F + \|T_n(\Lambda(H^{-1})^*) - T_n(H)T_n(H^{-1}\Lambda(H^{-1})^*)\|_F\right) \\ &\leq \|(T_n(H))^{-1}\|_2\left(\|((T_n(H))^{-1})^*\|_2\|T_n(\Lambda) - T_n(\Lambda(H^{-1})^*)(T_n(H))^*\|_F \right.\\ &\quad\left. +\|T_n(\Lambda(H^{-1})^*) - T_n(H)T_n(H^{-1}\Lambda(H^{-1})^*)\|_F\right) \\ &= \|(T_n(H))^{-1}\|_2\left(\|(T_n(H))^{-1}\|_2\|T_n(\Lambda) - T_n(\Lambda(H^{-1})^*)T_n(H^*)\|_F \right.\\ &\quad\left. +\|T_n(\Lambda(H^{-1})^*) - T_n(H)T_n(H^{-1}\Lambda(H^{-1})^*)\|_F\right) \\ &= \|(T_n(H))^{-1}\|_2\left(\|(T_n(H))^{-1}\|_2\|T_n(\Lambda(H^{-1})^*)T_n(H^*) - T_n(\Lambda(H^{-1})^*H^*)\|_F \right.\\ &\quad\left. +\|T_n(H)T_n(H^{-1}\Lambda(H^{-1})^*) - T_n(HH^{-1}\Lambda(H^{-1})^*)\|_F\right)\end{aligned}$$

for all $n \in \mathbb{N}$. As H and H^* are trigonometric polynomials, from [10, Lemma 2] we obtain that $\{\|T_n(\Lambda(H^{-1})^*)T_n(H^*) - T_n(\Lambda(H^{-1})^*H^*)\|_F\}$ and $\{\|T_n(H)T_n(H^{-1}\Lambda(H^{-1})^*) - T_n(HH^{-1}\Lambda(H^{-1})^*)\|_F\}$ are bounded, and consequently, $\{\|E\left(y_{n:1}y_{n:1}^*\right) - T_n(H^{-1}\Lambda(H^{-1})^*)\|_F\}$ is bounded.

(4)⇒(5) It is direct from Theorem 8.

(5)⇒(6) It is obvious.

(6)⇒(7) It is direct from [3, Definition 7.1].

(7)⇒(3) Let $\Lambda = U(\delta_{j,k}\lambda_k(\Lambda))_{j,k=1}^N U^{-1}$ be an eigenvalue decomposition of the positive definite matrix Λ, where U is unitary. Therefore, $\sqrt{\Lambda} = U\left(\delta_{j,k}\sqrt{\lambda_k(\Lambda)}\right)_{j,k=1}^N U^{-1}$ is also positive definite, and hence,

$$\begin{aligned}\left\|(T_n(H))^{-1}\right\|_2 &= \left\|(T_n(H))^{-1}T_n(\sqrt{\Lambda}(\sqrt{\Lambda})^{-1})\right\|_2 \\ &= \left\|(T_n(H))^{-1}T_n(\sqrt{\Lambda})T_n((\sqrt{\Lambda})^{-1})\right\|_2 \\ &\leq \left\|(T_n(H))^{-1}T_n(\sqrt{\Lambda})\right\|_2\left\|T_n((\sqrt{\Lambda})^{-1})\right\|_2 \\ &= \sqrt{\lambda_1\left((T_n(H))^{-1}T_n(\sqrt{\Lambda})T_n((\sqrt{\Lambda})^*)((T_n(H))^{-1})^*\right)\lambda_1\left(T_n((\sqrt{\Lambda})^{-1})T_n(((\sqrt{\Lambda})^{-1})^*)\right)} \\ &= \sqrt{\lambda_1\left((T_n(H))^{-1}T_n(\Lambda)((T_n(H))^{-1})^*\right)\lambda_1(T_n((\sqrt{\Lambda})^{-1}((\sqrt{\Lambda})^*)^{-1}))}\end{aligned}$$

$$
\begin{aligned}
&= \sqrt{\lambda_1\left(E(y_{n:1}y_{n:1}^*)\right)\lambda_1(T_n(\Lambda^{-1}))} \\
&= \sqrt{\|E(y_{n:1}y_{n:1}^*)\|_2\lambda_1((T_n(\Lambda))^{-1})} \\
&= \sqrt{\frac{\|E(y_{n:1}y_{n:1}^*)\|_2}{\lambda_{nN}(T_n(\Lambda))}} \\
&= \sqrt{\frac{\|E(y_{n:1}y_{n:1}^*)\|_2}{\lambda_N(\Lambda)}} \qquad \forall n \in \mathbb{N}.
\end{aligned}
$$

□

Theorem 10 shows in which sense the correlation matrices of an AR(p) "*stationary*" vector process approach the correlation matrices of a WSS vector process. Specifically, Theorem 10 shows that a necessary and sufficient condition for an AR(p) vector process $\{y_n\}$ to be "*stationary*" is the boundedness of $\{\|E\left(y_{n:1}y_{n:1}^*\right) - T_n(H^{-1}\Lambda(H^{-1})^*)\|_F\}$.

We finish this section by showing that a necessary condition for an AR(p) vector process to be "*stationary*" is to have an asymptotically stationary correlation structure.

Theorem 11. *Let $\{y_n\}$ and H be as in Theorem 10. If $\rho(\Psi_p(H)) < 1$ then $\{y_n\}$ has an asymptotically stationary correlation structure. In fact,*

$$\lim_{n\to\infty} E\left(y_n y_{n+k}^*\right) = \frac{1}{2\pi}\int_0^{2\pi} \mathrm{e}^{-k\omega \mathrm{i}}(H(\omega))^{-1}\Lambda\,((H(\omega))^{-1})^* d\omega \qquad \forall k \in \mathbb{Z}.$$

Proof. It is a direct consequence of Theorems 8 and 10. □

Theorem 11 was given by Berger in [4, Eq. (6.3.10)] for real AR(1) scalar processes.

5. Applications

Theorem 10 finds application in practical situations involving computations with large correlation matrices of AR(p) "*stationary*" vector processes. As examples of such practical applications in this section we obtain a novel result in information theory and another one in statistical signal processing.

5.1. An application in information theory

Kolmogorov computed the differential entropy rate of real Gaussian WSS scalar processes (see, e.g., [13, Section 12.5]). We here compute the differential entropy rate of any proper[1] (complex) Gaussian AR(p) "*stationary*" vector process $\{y_n\}$, i.e., we compute $\lim_{n\to\infty}\frac{1}{n}h(y_{n:1})$, where $h(y_{n:1})$ denotes the differential entropy of $y_{n:1}$ for all $n \in \mathbb{N}$.

Theorem 12. *Let $\{y_n\}$ and H be as in Theorem 10. Suppose that $\{y_n\}$ is proper and Gaussian. If $\rho(\Psi_p(H)) < 1$ then*

$$\lim_{n\to\infty}\frac{1}{n}h(y_{n:1}) = N\log_2(\pi\mathrm{e}) + \log_2\det(\Lambda) - \frac{1}{\pi}\int_0^{2\pi}\log_2|\det(H(\omega))|d\omega.$$

Proof. By Theorem 10, H is invertible and $\{y_n\}$ is AWSS with PSD $H^{-1}\Lambda(H^{-1})^*$. Thus, applying Lemma 9 and [9, Proposition 2] yields

$$
\begin{aligned}
0 < \inf_{n\in\mathbb{N}}\lambda_{nN}(E(y_{n:1}y_{n:1}^*)) &\le \min_{\omega\in[0,2\pi]}\lambda_N((H(\omega))^{-1}\Lambda\,((H(\omega))^{-1})^*) \\
&\le \max_{\omega\in[0,2\pi]}\lambda_1((H(\omega))^{-1}\Lambda\,((H(\omega))^{-1})^*) \le \sup_{n\in\mathbb{N}}\lambda_1(E(y_{n:1}y_{n:1}^*)) < \infty.
\end{aligned}
$$

Consequently, from [3, Section 7.2] we obtain

$$
\begin{aligned}
\lim_{n\to\infty}\frac{1}{n}h(y_{n:1}) &= N\log_2(\pi\mathrm{e}) + \frac{1}{2\pi}\int_0^{2\pi}\log_2\det((H(\omega))^{-1}\Lambda\,((H(\omega))^{-1})^*)d\omega \\
&= N\log_2(\pi\mathrm{e}) + \frac{1}{2\pi}\int_0^{2\pi}\log_2(\det((H(\omega))^{-1})\det(\Lambda)\det(((H(\omega))^{-1})^*))d\omega \\
&= N\log_2(\pi\mathrm{e}) + \frac{1}{2\pi}\int_0^{2\pi}\log_2(\det((H(\omega))^{-1})\det(\Lambda)\overline{\det((H(\omega))^{-1})})d\omega \\
&= N\log_2(\pi\mathrm{e}) + \log_2\det(\Lambda) + \frac{1}{2\pi}\int_0^{2\pi}\log_2\left(|\det((H(\omega))^{-1})|^2\right)d\omega
\end{aligned}
$$

[1] We recall that a zero-mean stochastic N-dimensional process $\{y_n\}$ is proper if $\{E(y_{n:1}y_{n:1}^\top)\} = \{0_{nN\times nN}\}$, where $\top$ denotes transpose.

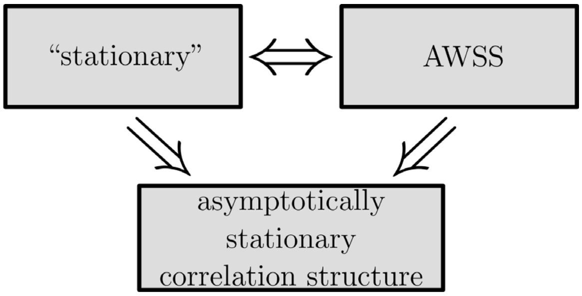

Fig. 1. Relation between the concepts of "*stationary*", AWSS, and asymptotically stationary correlation structure for AR(p) vector processes.

$$= N\log_2(\pi e) + \log_2\det(\Lambda) + \frac{1}{\pi}\int_0^{2\pi}\log_2\left|\frac{1}{\det(H(\omega))}\right|d\omega$$
$$= N\log_2(\pi e) + \log_2\det(\Lambda) + \frac{1}{\pi}\int_0^{2\pi}\log_2\frac{1}{|\det(H(\omega))|}d\omega.$$

□

5.2. An application in statistical signal processing

We now extend to AR(p) "*stationary*" vector processes the Pisarenko spectral estimation method given in [5, Theorem 2] for WSS scalar processes.

Theorem 13. *Let $\{y_n\}$ and H be as in Theorem 10. Consider a continuous strictly monotonic function $g:(0,\infty)\to\mathbb{R}$. If $\rho(\Psi_p(H))<1$ then*

$$\lim_{n\to\infty}\frac{1}{2\pi}\int_0^{2\pi}\left\|\widehat{P}_{E(y_{n:1}y_{n:1}^*),g}(\omega)-(H(\omega))^{-1}\Lambda\left((H(\omega))^{-1}\right)^*\right\|_F^2 d\omega = 0 \tag{8}$$

with $\widehat{P}_{E(y_{n:1}y_{n:1}^*),g}:\mathbb{R}\to\mathbb{C}^{N\times N}$ being the 2π-periodic step function defined as

$$\widehat{P}_{E(y_{n:1}y_{n:1}^*),g}(\omega)=\sum_{k=1}^{n}\chi_{\left[\frac{2\pi(k-1)}{n},\frac{2\pi k}{n}\right]}(\omega)g^{-1}\left([(V_n\otimes I_N)^* g(E(y_{n:1}y_{n:1}^*))(V_n\otimes I_N)]_{k,k}\right),\quad \omega\in[0,2\pi),\quad n\in\mathbb{N},$$

where χ denotes characteristic function and V_n is the $n\times n$ Fourier unitary matrix given in Equality (6).

Proof. By Theorem 10, H is invertible and $\{y_n\}$ is AWSS with PSD $H^{-1}\Lambda(H^{-1})^*$. Therefore, applying Lemma 9 and [9, Theorem 5] yields Equality (8). □

Observe that for each function g Theorem 13 provides a spectral estimation method, i.e., Theorem 13 provides a sequence of functions $\left\{\widehat{P}_{E(y_{n:1}y_{n:1}^*),g}\right\}$ that approximates (or estimates) the PSD $H^{-1}\Lambda(H^{-1})^*$ of the AR(p) "*stationary*" vector process $\{y_n\}$. In particular, by taking $g(x)=x$ and $g(x)=\frac{1}{x}$ in Theorem 13, we have extended to AR(p) "*stationary*" vector processes the (averaged) periodogram method and the Capon spectral estimation method, respectively.

6. Conclusions

In the literature three different definitions have been presented to describe when AR(p) vector processes behave similarly to WSS vector processes. In this paper we have found the relation between these three concepts for AR(p) vector processes. This relation, which had not been established until now, is shown in Fig. 1.

Data availability

Data will be made available on request.

Acknowledgements

This work was supported by the Basque Government through the research project PIBA_2020_1_0050.

Appendix A

Proof of Theorem 1. It falls into seven parts.

Part 1: $[(\Phi_n(H))^n]_{1,k} = [K_n(H)(\Phi_n(H))^n]_{n,k} \in \mathbb{C}^{N\times N}$ $\forall k \in \{1,\dots,n\}$.

As $[K_n(H)]_{n,k} = [I_{nN}]_{1,k}$ for all $k \in \{1,\dots,n\}$, we have

$$[K_n(H)(\Phi_n(H))^n]_{n,k} = \sum_{h=1}^{n} [K_n(H)]_{n,h}[(\Phi_n(H))^n]_{h,k} = \sum_{h=1}^{n} [I_{nN}]_{1,h}[(\Phi_n(H))^n]_{h,k} = [(\Phi_n(H))^n]_{1,k}$$

for all $k \in \{1,\dots,n\}$.

Part 2: $[(\Phi_n(H))^n]_{j,k} = [(\Phi_n(H))^{n+1}]_{j-1,k} \in \mathbb{C}^{N\times N}$ for all $j \in \{2,\dots,n\}$ and $k \in \{1,\dots,n\}$.

Since

$$[\Phi_n(H)]_{j,k} = [I_{nN}]_{j+1,k}, \qquad j \in \{1,\dots,n-1\}, \quad k \in \{1,\dots,n\}, \tag{A.1}$$

we obtain

$$[(\Phi_n(H))^{n+1}]_{j-1,k} = \sum_{h=1}^{n} [\Phi_n(H)]_{j-1,h}[(\Phi_n(H))^n]_{h,k} = \sum_{h=1}^{n} [I_{nN}]_{j,h}[(\Phi_n(H))^n]_{h,k} = [(\Phi_n(H))^n]_{j,k}$$

for all $j \in \{2,\dots,n\}$ and $k \in \{1,\dots,n\}$.

Part 3: $[K_n(H)(\Phi_n(H))^n]_{j,k} = \mathsf{H}_{-(n-j)}[K_n(H)(\Phi_n(H))^n]_{n,k} + \sum_{h=1}^{n}[K_n(H)(\Phi_n(H))^n]_{j+1,h}[\Phi_n(H)]_{h,k}$ for all $j \in \{1,\dots,n-1\}$ and $k \in \{1,\dots,n\}$.

As

$$[K_n(H)]_{j,k} = \begin{cases} \mathsf{H}_{j+k-n-1} & \text{if } j+k \le n+1, \\ 0_{N\times N} & \text{if } j+k > n+1, \end{cases}$$

for all $j,k \in \{1,\dots,n\}$, we have

$$\begin{aligned} [K_n(H)(\Phi_n(H))^n]_{j,k} &= \sum_{h=1}^{n} [K_n(H)]_{j,h}[(\Phi_n(H))^n]_{h,k} \\ &= \sum_{h=1}^{n-j+1} [K_n(H)]_{j,h}[(\Phi_n(H))^n]_{h,k} + \sum_{h=n-j+2}^{n} [K_n(H)]_{j,h}[(\Phi_n(H))^n]_{h,k} \\ &= \sum_{h=1}^{n-j+1} \mathsf{H}_{j+h-n-1}[(\Phi_n(H))^n]_{h,k} \qquad \forall j,k \in \{1,\dots,n\}. \end{aligned}$$

Consequently, from Parts 1 and 2 we obtain

$$\begin{aligned} &[K_n(H)(\Phi_n(H))^n]_{j,k} \\ &= \sum_{h=1}^{n-j+1} \mathsf{H}_{j+h-n-1}[(\Phi_n(H))^n]_{h,k} \\ &= \mathsf{H}_{j-n}[(\Phi_n(H))^n]_{1,k} + \sum_{h=2}^{n-j+1} \mathsf{H}_{j+h-n-1}[(\Phi_n(H))^n]_{h,k} \\ &= \mathsf{H}_{-(n-j)}[K_n(H)(\Phi_n(H))^n]_{n,k} + \sum_{h=2}^{n-j+1} \mathsf{H}_{j+h-n-1}[(\Phi_n(H))^{n+1}]_{h-1,k} \\ &= \mathsf{H}_{-(n-j)}[K_n(H)(\Phi_n(H))^n]_{n,k} + \sum_{h=2}^{n-j+1} \mathsf{H}_{j+h-n-1}\sum_{r=1}^{n}[(\Phi_n(H))^n]_{h-1,r}[\Phi_n(H)]_{r,k} \\ &= \mathsf{H}_{-(n-j)}[K_n(H)(\Phi_n(H))^n]_{n,k} + \sum_{s=1}^{n-j} \mathsf{H}_{j+s-n}\sum_{r=1}^{n}[(\Phi_n(H))^n]_{s,r}[\Phi_n(H)]_{r,k} \\ &= \mathsf{H}_{-(n-j)}[K_n(H)(\Phi_n(H))^n]_{n,k} + \sum_{s=1}^{n-(j+1)+1} \mathsf{H}_{(j+1)+s-n-1}\sum_{r=1}^{n}[(\Phi_n(H))^n]_{s,r}[\Phi_n(H)]_{r,k} \\ &= \mathsf{H}_{-(n-j)}[K_n(H)(\Phi_n(H))^n]_{n,k} + \sum_{r=1}^{n}\left(\sum_{s=1}^{n-(j+1)+1} \mathsf{H}_{(j+1)+s-n-1}[(\Phi_n(H))^n]_{s,r}\right)[\Phi_n(H)]_{r,k} \end{aligned}$$

$$= \mathsf{H}_{-(n-j)}[K_n(H)(\Phi_n(H))^n]_{n,k} + \sum_{r=1}^{n}[K_n(H)(\Phi_n(H))^n]_{j+1,r}[\Phi_n(H)]_{r,k}$$

for all $j \in \{1, \dots, n-1\}$ and $k \in \{1, \dots, n\}$.

Part 4: $[Z_n(H)]_{j,k} = \mathsf{H}_{-(n-j)}[Z_n(H)]_{n,k} + \sum_{h=1}^{n}[Z_n(H)]_{j+1,h}[\Phi_n(H)]_{h,k}$ for all $j \in \{1, \dots, n-1\}$ and $k \in \{1, \dots, n\}$.
Since

$$[Z_n(H)]_{j,k} = \begin{cases} -\mathsf{H}_{j+k-2n-1} & \text{if } j+k \geq n+1, \\ 0_{N\times N} & \text{if } j+k < n+1, \end{cases}$$

for all $j, k \in \{1, \dots, n\}$ and

$$[\Phi_n(H)]_{n,k} = [Z_n(H)]_{n,k} \qquad \forall k \in \{1, \dots, n\}, \tag{A.2}$$

applying Equality (A.1) yields

$$\begin{aligned}
&\mathsf{H}_{-(n-j)}[Z_n(H)]_{n,k} + \sum_{h=1}^{n}[Z_n(H)]_{j+1,h}[\Phi_n(H)]_{h,k} \\
&= \mathsf{H}_{-(n-j)}[Z_n(H)]_{n,k} + \sum_{h=1}^{n-1}[Z_n(H)]_{j+1,h}[I_{nN}]_{h+1,k} + [Z_n(H)]_{j+1,n}[\Phi_n(H)]_{n,k} \\
&= \mathsf{H}_{-(n-j)}[Z_n(H)]_{n,k} + \sum_{h=1}^{n-1}[Z_n(H)]_{j+1,h}[I_{nN}]_{h+1,k} - \mathsf{H}_{j-n}[Z_n(H)]_{n,k} \\
&= \sum_{h=1}^{n-1}[Z_n(H)]_{j+1,h}[I_{nN}]_{h+1,k} \\
&= \begin{cases} [Z_n(H)]_{j+1,k-1} & \text{if } k \neq 1, \\ 0_{N\times N} & \text{if } k = 1, \end{cases} \\
&= [Z_n(H)]_{j,k}, \qquad j \in \{1, \dots, n-1\}, \quad k \in \{1, \dots, n\}.
\end{aligned}$$

Part 5: If $q \in \{1, \dots, n-1\}$ then $[(\Phi_n(H))^q]_{1,k} = [I_{nN}]_{q+1,k} \in \mathbb{C}^{N\times N}$ $\forall k \in \{1, \dots, n\}$.
We proceed by induction on q. From Equality (A.1),

$$\left[(\Phi_n(H))^q\right]_{1,k} = [I_{nN}]_{q+1,k} \qquad \forall k \in \{1, \dots, n\} \tag{A.3}$$

is true for $q = 1$. We now assume that Equality (A.3) is true for certain $q \in \{1, \dots, n-2\}$, and applying Equality (A.1) we have

$$\begin{aligned}
[(\Phi_n(H))^{q+1}]_{1,k} &= [(\Phi_n(H))^q \Phi_n(H)]_{1,k} = \sum_{h=1}^{n}[(\Phi_n(H))^q]_{1,h}[\Phi_n(H)]_{h,k} \\
&= \sum_{h=1}^{n}[I_{nN}]_{q+1,h}[\Phi_n(H)]_{h,k} = [\Phi_n(H)]_{q+1,k} = [I_{nN}]_{q+2,k} \qquad \forall k \in \{1, \dots, n\}.
\end{aligned}$$

Part 6: $[K_n(H)(\Phi_n(H))^n]_{n,k} = [Z_n(H)]_{n,k} \in \mathbb{C}^{N\times N}$ $\forall k \in \{1, \dots, n\}$.
From Parts 1 and 5, and Equality (A.2) we obtain

$$\begin{aligned}
[K_n(H)(\Phi_n(H))^n]_{n,k} &= [(\Phi_n(H))^n]_{1,k} = \sum_{h=1}^{n}[(\Phi_n(H))^{n-1}]_{1,h}[\Phi_n(H)]_{h,k} \\
&= \sum_{h=1}^{n}[I_{nN}]_{n,h}[\Phi_n(H)]_{h,k} = [\Phi_n(H)]_{n,k} = [Z_n(H)]_{n,k} \qquad \forall k \in \{1, \dots, n\}.
\end{aligned}$$

Part 7: $K_n(H)(\Phi_n(H))^n = Z_n(H)$.
It is a direct consequence of Parts 3, 4, and 6. □

References

[1] G.C. Reinsel, Elements of Multivariate Time Series Analysis, Springer, 1993.
[2] R.M. Gray, Toeplitz and circulant matrices: a review, Foundations and Trends in Communications and Information Theory 2 (3) (2006) 155–239.
[3] J. Gutiérrez-Gutiérrez, P.M. Crespo, Block Toeplitz matrices: asymptotic results and applications, Foundations and Trends in Communications and Information Theory 8 (3) (2011) 179–257.
[4] T. Berger, Rate Distortion Theory: A Mathematical Basis for Data Compression, Prentice-Hall, 1971.
[5] V.F. Pisarenko, On the estimation of spectra by means of non-linear functions of the covariance matrix, Geophysical Jounal of the Royal Astronomical Society 28 (1972) 511–531.

[6] S. Barnett, Polynomials and Linear Control Systems, Marcel Dekker, 1983.
[7] I. Gohberg, P. Lancaster, L. Rodman, Matrix Polynomials, Academic Press, 1982.
[8] R.A. Horn, C.R. Johnson, Matrix Analysis, Cambridge University Press, 1990.
[9] J. Gutiérrez-Gutiérrez, A modified version of the Pisarenko method to estimate the power spectral density of any asymptotically wide sense stationary vector process, Appl Math Comput 362 (2019). article no. 124526
[10] J. Gutiérrez-Gutiérrez, M. Zárraga-Rodríguez, X. Insausti, On the asymptotic optimality of a low-complexity coding strategy for WSS, MA, and AR vector sources, Entropy 22 (2020). article no. 1378
[11] P. Tilli, Singular values and eigenvalues of non-Hermitian block Toeplitz matrices, Linear Algebra Appl 272 (1998) 59–89.
[12] G. Barbarino, C. Garoni, S. Serra-Capizzano, Block generalized locally Toeplitz sequences: theory and applications in the unidimensional case, Electron. Trans. Numer. Anal. 53 (2020) 28–112.
[13] T.M. Cover, J.A. Thomas, Elements of Information Theory, Wiley, 2006.