

Article

Rate-Distortion Function Upper Bounds for Gaussian Vectors and Their Applications in Coding AR Sources

Jesús Gutiérrez-Gutiérrez * , Marta Zárraga-Rodríguez and Fernando M. Villar-Rosety and Xavier Insausti 

Tecnun, University of Navarra, Paseo de Manuel Lardizábal 13, 20018 San Sebastián, Spain; mzarraga@tecnun.es (M.Z.-R.); fmdvillar@tecnun.es (F.M.V.-R.); xinsausti@tecnun.es (X.I.)

* Correspondence: jgutierrez@tecnun.es; Tel.: +34-94-3219877

Received: 16 April 2018 ; Accepted: 20 May 2018; Published: 23 May 2018



Abstract: In this paper, we give upper bounds for the rate-distortion function (RDF) of any Gaussian vector, and we propose coding strategies to achieve such bounds. We use these strategies to reduce the computational complexity of coding Gaussian asymptotically wide sense stationary (AWSS) autoregressive (AR) sources. Furthermore, we also give sufficient conditions for AR processes to be AWSS.

Keywords: source coding; rate-distortion function (RDF); Gaussian vector; autoregressive (AR) source; discrete Fourier transform (DFT)

1. Introduction

In 1956, Kolmogorov [1] gave a formula for the rate-distortion function (RDF) of Gaussian vectors and the RDF of Gaussian wide sense stationary (WSS) sources. Later, in 1970 Gray [2] obtained a formula for the RDF of Gaussian autoregressive (AR) sources.

In 1973, Pearl [3] gave an upper bound for the RDF of finite-length data blocks of Gaussian WSS sources, but he did not propose a coding strategy to achieve his bound for a given block length. In [4], we presented two tighter upper bounds for the RDF of finite-length data blocks of Gaussian WSS sources, and we proposed low-complexity coding strategies, based on the discrete Fourier transform (DFT), to achieve such bounds. Moreover, we proved that those two upper bounds tend to the RDF of the WSS source (computed by Kolmogorov in [1]) when the size of the data block grows.

In the present paper, we generalize the upper bounds and the two low-complexity coding strategies presented in [4] to any Gaussian vector. Therefore, in contrast to [4], here no assumption about the structure of the correlation matrix of the Gaussian vector has been made (observe that since the sources in [4] were WSS the correlation matrix of the Gaussian vectors there considered was Toeplitz). To obtain such generalization we start our analysis by first proving several new results on the DFT of random vectors. Although in [4] (Theorem 1) another new result on the DFT was presented, it cannot be used here, because such result and its proof rely on the power spectral density (PSD) of a WSS process and its properties.

The two low-complexity strategies here presented are applied in coding finite-length data blocks of Gaussian AR sources. Specifically, we prove that the rates (upper bounds) corresponding to these two strategies tend to the RDF of the AR source (computed by Gray in [2]) when the size of the data block grows and the AR source is asymptotically WSS (AWSS).

The definition of AWSS process was introduced by Gray in [5] (Chapter 6) and it is based on his concept of asymptotically equivalent sequences of matrices [6]. Sufficient conditions for AR processes to be AWSS can be found in [5] (Theorem 6.2) and [7] (Theorem 7). In this paper we present other sufficient conditions which make easier to check in practice whether an AR process is AWSS.

The paper is organized as follows. In Section 2 we obtain several new results on the DFT of random vectors which are used in Section 3. In Section 3 we give upper bounds for the RDF of Gaussian vectors, and we propose coding strategies to achieve such bounds. In Section 4 we apply the strategies proposed in Section 3 to reduce the computational complexity of coding Gaussian AWSS AR sources. In Section 5 we give sufficient conditions for AR processes to be AWSS. We finish the paper with a numerical example and conclusions.

2. Several New Results on the DFT of Random Vectors

We begin by introducing some notation. \mathbb{C} denotes the set of (finite) complex numbers, i is the imaginary unit, Re and Im denote real and imaginary parts, respectively. $*$ stands for conjugate transpose, \top denotes transpose, and $\lambda_k(A), k \in \{1, \dots, n\}$, are the eigenvalues of an $n \times n$ Hermitian matrix A arranged in decreasing order. E stands for expectation, and V_n is the $n \times n$ Fourier unitary matrix, i.e.,

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

If $z \in \mathbb{C}$ then \hat{z} denotes the real (column) vector

$$\hat{z} = \begin{pmatrix} \text{Re}(z) \\ \text{Im}(z) \end{pmatrix}.$$

If $z_k \in \mathbb{C}$ for all $k \in \{1, \dots, n\}$ then $z_{n:1}$ is the n -dimensional vector given by

$$z_{n:1} = \begin{pmatrix} z_n \\ z_{n-1} \\ z_{n-2} \\ \vdots \\ z_1 \end{pmatrix}.$$

In this section, we give several new results on the DFT of random vectors in two theorems and one lemma.

Theorem 1. Let $y_{n:1}$ be the DFT of an n -dimensional random vector $x_{n:1}$, that is, $y_{n:1} = V_n^* x_{n:1}$.

1. If $k \in \{1, \dots, n\}$ then

$$\lambda_n(E(x_{n:1} x_{n:1}^*)) \leq E(|x_k|^2) \leq \lambda_1(E(x_{n:1} x_{n:1}^*)) \tag{1}$$

and

$$\lambda_n(E(x_{n:1} x_{n:1}^*)) \leq E(|y_k|^2) \leq \lambda_1(E(x_{n:1} x_{n:1}^*)). \tag{2}$$

2. If the random vector $x_{n:1}$ is real and $k \in \{1, \dots, n-1\} \setminus \{\frac{n}{2}\}$ then

$$\frac{\lambda_n(E(x_{n:1} x_{n:1}^\top))}{2} \leq E((\text{Re}(y_k))^2) \leq \frac{\lambda_1(E(x_{n:1} x_{n:1}^\top))}{2}, \tag{3}$$

and

$$\frac{\lambda_n(E(x_{n:1} x_{n:1}^\top))}{2} \leq E((\text{Im}(y_k))^2) \leq \frac{\lambda_1(E(x_{n:1} x_{n:1}^\top))}{2}. \tag{4}$$

Proof. (1) We first prove that if W_n is an $n \times n$ unitary matrix then

$$\lambda_n(E(x_{n:1} x_{n:1}^*)) \leq [W_n \text{diag}_{1 \leq j \leq n}(\lambda_j(E(x_{n:1} x_{n:1}^*))) W_n^*]_{n-k+1, n-k+1} \leq \lambda_1(E(x_{n:1} x_{n:1}^*)). \tag{5}$$

We have

$$\begin{aligned}
 [W_n \text{diag}_{1 \leq j \leq n} (\lambda_j (E(x_{n:1} x_{n:1}^*))) W_n^*]_{k_1, k_2} &= \sum_{h=1}^n [W_n]_{k_1, h} [\text{diag}_{1 \leq j \leq n} (\lambda_j (E(x_{n:1} x_{n:1}^*))) W_n^*]_{h, k_2} \\
 &= \sum_{h=1}^n [W_n]_{k_1, h} \sum_{l=1}^n [\text{diag}_{1 \leq j \leq n} (\lambda_j (E(x_{n:1} x_{n:1}^*))) W_n^*]_{h, l} [W_n^*]_{l, k_2} \\
 &= \sum_{h=1}^n [W_n]_{k_1, h} \lambda_h (E(x_{n:1} x_{n:1}^*)) \overline{[W_n]_{k_2, h}}
 \end{aligned} \tag{6}$$

for all $k_1, k_2 \in \{1, \dots, n\}$, and hence,

$$[W_n \text{diag}_{1 \leq j \leq n} (\lambda_j (E(x_{n:1} x_{n:1}^*))) W_n^*]_{n-k+1, n-k+1} = \sum_{h=1}^n \lambda_h (E(x_{n:1} x_{n:1}^*)) |[W_n]_{n-k+1, h}|^2.$$

Consequently,

$$\begin{aligned}
 \lambda_n (E(x_{n:1} x_{n:1}^*)) \sum_{h=1}^n |[W_n]_{n-k+1, h}|^2 &\leq [W_n \text{diag}_{1 \leq j \leq n} (\lambda_j (E(x_{n:1} x_{n:1}^*))) W_n^*]_{n-k+1, n-k+1} \\
 &\leq \lambda_1 (E(x_{n:1} x_{n:1}^*)) \sum_{h=1}^n |[W_n]_{n-k+1, h}|^2,
 \end{aligned}$$

and applying

$$\sum_{h=1}^n |[W_n]_{n-k+1, h}|^2 = \sum_{h=1}^n [W_n]_{n-k+1, h} [W_n^*]_{h, n-k+1} = [W_n W_n^*]_{n-k+1, n-k+1} = [I_n]_{n-k+1, n-k+1} = 1,$$

where I_n denotes the $n \times n$ identity matrix, we obtain Equation (5).

Let $E(x_{n:1} x_{n:1}^*) = U_n \text{diag}_{1 \leq j \leq n} (\lambda_j (E(x_{n:1} x_{n:1}^*))) U_n^{-1}$ be a diagonalization of $E(x_{n:1} x_{n:1}^*)$ where the eigenvector matrix U_n is unitary. As

$$E(|x_k|^2) = [E(x_{n:1} x_{n:1}^*)]_{n-k+1, n-k+1} = [U_n \text{diag}_{1 \leq j \leq n} (\lambda_j (E(x_{n:1} x_{n:1}^*))) U_n^*]_{n-k+1, n-k+1},$$

Equation (1) follows directly by taking $W_n = U_n$ in Equation (5).

Since

$$\begin{aligned}
 E(|y_k|^2) &= [E(y_{n:1} y_{n:1}^*)]_{n-k+1, n-k+1} \\
 &= [E(V_n^* x_{n:1} x_{n:1}^* (V_n^*)^*)]_{n-k+1, n-k+1} \\
 &= [V_n^* E(x_{n:1} x_{n:1}^*) (V_n^*)^*]_{n-k+1, n-k+1} \\
 &= [V_n^* U_n \text{diag}_{1 \leq j \leq n} (\lambda_j (E(x_{n:1} x_{n:1}^*))) U_n^* (V_n^*)^*]_{n-k+1, n-k+1} \\
 &= [V_n^* U_n \text{diag}_{1 \leq j \leq n} (\lambda_j (E(x_{n:1} x_{n:1}^*))) (V_n^* U_n)^*]_{n-k+1, n-k+1},
 \end{aligned} \tag{7}$$

taking $W_n = V_n^* U_n$ in Equation (5) we obtain Equation (2).

(2) Applying [4] (Equation (10)) and taking $W_n = U_n$ in Equation (6) yields

$$\begin{aligned}
 E((\text{Re}(y_k))^2) &= \frac{1}{n} \sum_{k_1, k_2=1}^n \cos \frac{2\pi(1-k_1)k}{n} \cos \frac{2\pi(1-k_2)k}{n} E(x_{n-k_1+1} x_{n-k_2+1}) \\
 &= \frac{1}{n} \sum_{k_1, k_2=1}^n \cos \frac{2\pi(1-k_1)k}{n} \cos \frac{2\pi(1-k_2)k}{n} [E(x_{n:1} x_{n:1}^\top)]_{k_1, k_2}
 \end{aligned}$$

$$\begin{aligned}
 &= \frac{1}{n} \sum_{k_1, k_2=1}^n \cos \frac{2\pi(1-k_1)k}{n} \cos \frac{2\pi(1-k_2)k}{n} \left[U_n \text{diag}_{1 \leq j \leq n} (\lambda_j (E(x_{n:1} x_{n:1}^\top))) U_n^* \right]_{k_1, k_2} \\
 &= \frac{1}{n} \sum_{k_1, k_2=1}^n \cos \frac{2\pi(1-k_1)k}{n} \cos \frac{2\pi(1-k_2)k}{n} \sum_{h=1}^n [U_n]_{k_1, h} \lambda_h (E(x_{n:1} x_{n:1}^\top)) \overline{[U_n]_{k_2, h}} \\
 &= \frac{1}{n} \sum_{h=1}^n \lambda_h (E(x_{n:1} x_{n:1}^\top)) \left(\sum_{k_1=1}^n \cos \frac{2\pi(1-k_1)k}{n} [U_n]_{k_1, h} \right) \left(\sum_{k_2=1}^n \cos \frac{2\pi(1-k_2)k}{n} \overline{[U_n]_{k_2, h}} \right) \\
 &= \frac{1}{n} \sum_{h=1}^n \lambda_h (E(x_{n:1} x_{n:1}^\top)) \left| \sum_{l=1}^n \cos \frac{2\pi(1-l)k}{n} [U_n]_{l, h} \right|^2,
 \end{aligned}$$

and therefore,

$$\begin{aligned}
 \lambda_n (E(x_{n:1} x_{n:1}^\top)) \frac{1}{n} \sum_{h=1}^n \left| \sum_{l=1}^n \cos \frac{2\pi(1-l)k}{n} [U_n]_{l, h} \right|^2 \\
 \leq E((\text{Re}(y_k))^2) \leq \lambda_1 (E(x_{n:1} x_{n:1}^\top)) \frac{1}{n} \sum_{h=1}^n \left| \sum_{l=1}^n \cos \frac{2\pi(1-l)k}{n} [U_n]_{l, h} \right|^2.
 \end{aligned}$$

Analogously, it can be proved that

$$\begin{aligned}
 \lambda_n (E(x_{n:1} x_{n:1}^\top)) \frac{1}{n} \sum_{h=1}^n \left| \sum_{l=1}^n \sin \frac{2\pi(1-l)k}{n} [U_n]_{l, h} \right|^2 \\
 \leq E((\text{Im}(y_k))^2) \leq \lambda_1 (E(x_{n:1} x_{n:1}^\top)) \frac{1}{n} \sum_{h=1}^n \left| \sum_{l=1}^n \sin \frac{2\pi(1-l)k}{n} [U_n]_{l, h} \right|^2.
 \end{aligned}$$

To finish the proof we only need to show that

$$\frac{1}{n} \sum_{h=1}^n \left| \sum_{l=1}^n \cos \frac{2\pi(1-l)k}{n} [U_n]_{l, h} \right|^2 = \frac{1}{n} \sum_{h=1}^n \left| \sum_{l=1}^n \sin \frac{2\pi(1-l)k}{n} [U_n]_{l, h} \right|^2 = \frac{1}{2}. \tag{8}$$

If b_1, \dots, b_n are n real numbers then

$$\begin{aligned}
 \frac{1}{n} \sum_{h=1}^n \left| \sum_{l=1}^n b_l [U_n]_{l, h} \right|^2 &= \frac{1}{n} \sum_{h=1}^n \left(\sum_{k_1=1}^n b_{k_1} [U_n]_{k_1, h} \right) \left(\sum_{k_2=1}^n b_{k_2} \overline{[U_n]_{k_2, h}} \right) = \frac{1}{n} \sum_{k_1, k_2=1}^n b_{k_1} b_{k_2} \sum_{h=1}^n [U_n]_{k_1, h} \overline{[U_n]_{k_2, h}} \\
 &= \frac{1}{n} \sum_{k_1, k_2=1}^n b_{k_1} b_{k_2} [U_n U_n^*]_{k_1, k_2} = \frac{1}{n} \sum_{k_1, k_2=1}^n b_{k_1} b_{k_2} [I_n]_{k_1, k_2} = \frac{1}{n} \sum_{l=1}^n b_l^2,
 \end{aligned} \tag{9}$$

and thus,

$$\begin{aligned}
 \frac{1}{n} \sum_{h=1}^n \left| \sum_{l=1}^n \sin \frac{2\pi(1-l)k}{n} [U_n]_{l, h} \right|^2 &= \frac{1}{n} \sum_{l=1}^n \left(\sin \frac{2\pi(1-l)k}{n} \right)^2 = \frac{1}{n} \sum_{l=1}^n \left(1 - \left(\cos \frac{2\pi(1-l)k}{n} \right)^2 \right) \\
 &= 1 - \frac{1}{n} \sum_{l=1}^n \left(\cos \frac{2\pi(1-l)k}{n} \right)^2 = 1 - \frac{1}{n} \sum_{h=1}^n \left| \sum_{l=1}^n \cos \frac{2\pi(1-l)k}{n} [U_n]_{l, h} \right|^2.
 \end{aligned}$$

Equation (8) now follows directly from [4] (Equation (15)). \square

Lemma 1. Let $y_{n:1}$ be the DFT of an n -dimensional random vector $x_{n:1}$. If $k \in \{1, \dots, n\}$ then

1. $E(|y_k|^2) = [V_n^* E(x_{n:1} x_{n:1}^*) V_n]_{n-k+1, n-k+1}$.
2. $E(y_k^2) = [V_n^* E(x_{n:1} x_{n:1}^\top) \overline{V_n}]_{n-k+1, n-k+1}$.
3. $E(\text{Re}(y_k) \text{Im}(y_k)) = \frac{1}{2} \text{Im}(E(y_k^2))$.

$$\begin{aligned}
 4. \quad E \left((\operatorname{Re}(y_k))^2 \right) &= \frac{E(|y_k|^2) + \operatorname{Re}(E(y_k^2))}{2}. \\
 5. \quad E \left((\operatorname{Im}(y_k))^2 \right) &= \frac{E(|y_k|^2) - \operatorname{Re}(E(y_k^2))}{2}.
 \end{aligned}$$

Proof. (1) It is a direct consequence of Equation (7).

(2) We have

$$\begin{aligned}
 E \left(y_k^2 \right) &= \left[E \left(y_{n:1} y_{n:1}^\top \right) \right]_{n-k+1, n-k+1} = \left[E \left(V_n^* x_{n:1} x_{n:1}^\top (V_n^*)^\top \right) \right]_{n-k+1, n-k+1} \\
 &= \left[E \left(V_n^* x_{n:1} x_{n:1}^\top \bar{V}_n \right) \right]_{n-k+1, n-k+1} = \left[V_n^* E \left(x_{n:1} x_{n:1}^\top \right) \bar{V}_n \right]_{n-k+1, n-k+1}.
 \end{aligned}$$

(3) Observe that

$$\begin{aligned}
 E \left(y_k^2 \right) &= E \left((\operatorname{Re}(y_k))^2 - (\operatorname{Im}(y_k))^2 + 2\operatorname{Re}(y_k)\operatorname{Im}(y_k)i \right) \\
 &= E \left((\operatorname{Re}(y_k))^2 \right) - E \left((\operatorname{Im}(y_k))^2 \right) + 2E \left(\operatorname{Re}(y_k)\operatorname{Im}(y_k) \right) i,
 \end{aligned} \tag{10}$$

and hence,

$$\operatorname{Im} \left(E \left(y_k^2 \right) \right) = 2E \left(\operatorname{Re}(y_k)\operatorname{Im}(y_k) \right).$$

(4) and (5) From Equation (10) we obtain

$$\operatorname{Re} \left(E \left(y_k^2 \right) \right) = E \left((\operatorname{Re}(y_k))^2 \right) - E \left((\operatorname{Im}(y_k))^2 \right). \tag{11}$$

Furthermore,

$$E \left(|y_k|^2 \right) = E \left((\operatorname{Re}(y_k))^2 + (\operatorname{Im}(y_k))^2 \right) = E \left((\operatorname{Re}(y_k))^2 \right) + E \left((\operatorname{Im}(y_k))^2 \right). \tag{12}$$

(4) and (5) follow directly from Equations (11) and (12). □

Theorem 2. Let $y_{n:1}$ be the DFT of a real n -dimensional random vector $x_{n:1}$. If $k \in \{1, \dots, n-1\} \setminus \{\frac{n}{2}\}$ then

$$\frac{\lambda_n(E(x_{n:1}x_{n:1}^\top))}{2} \leq \lambda_2 \left(E \left(\hat{y}_k (\hat{y}_k)^\top \right) \right) \leq \lambda_1 \left(E \left(\hat{y}_k (\hat{y}_k)^\top \right) \right) \leq \frac{\lambda_1(E(x_{n:1}x_{n:1}^\top))}{2}.$$

Proof. Fix $r \in \{1, 2\}$ and consider a real unit eigenvector $v = (v_1, v_2)^\top$ corresponding to $\lambda_r \left(E \left(\hat{y}_k (\hat{y}_k)^\top \right) \right)$. We have

$$\lambda_r \left(E \left(\hat{y}_k (\hat{y}_k)^\top \right) \right) = \lambda_r \left(E \left(\hat{y}_k (\hat{y}_k)^\top \right) \right) v^\top v = v^\top \left(\lambda_r \left(E \left(\hat{y}_k (\hat{y}_k)^\top \right) \right) v \right) = v^\top E \left(\hat{y}_k (\hat{y}_k)^\top \right) v.$$

From [4] (Equation (10)) we obtain

$$\begin{aligned}
 &E \left(\hat{y}_k (\hat{y}_k)^\top \right) \\
 &= \frac{1}{n} \sum_{k_1, k_2=1}^n \begin{pmatrix} \cos \frac{2\pi(1-k_1)k}{n} \cos \frac{2\pi(1-k_2)k}{n} E(x_{n-k_1+1}x_{n-k_2+1}) & \cos \frac{2\pi(1-k_1)k}{n} \sin \frac{2\pi(1-k_2)k}{n} E(x_{n-k_1+1}x_{n-k_2+1}) \\ \sin \frac{2\pi(1-k_1)k}{n} \cos \frac{2\pi(1-k_2)k}{n} E(x_{n-k_1+1}x_{n-k_2+1}) & \sin \frac{2\pi(1-k_1)k}{n} \sin \frac{2\pi(1-k_2)k}{n} E(x_{n-k_1+1}x_{n-k_2+1}) \end{pmatrix} \\
 &= \frac{1}{n} \sum_{k_1, k_2=1}^n \left[E \left(x_{n:1} x_{n:1}^\top \right) \right]_{k_1, k_2} w_{k_1} w_{k_2}^\top
 \end{aligned}$$

with

$$w_l = \begin{pmatrix} \cos \frac{2\pi(1-l)k}{n} \\ \sin \frac{2\pi(1-l)k}{n} \end{pmatrix}, \quad l \in \{1, \dots, n\},$$

and consequently,

$$\begin{aligned} \lambda_r \left(E \left(\hat{y}_k (\hat{y}_k)^\top \right) \right) &= \frac{1}{n} \sum_{k_1, k_2=1}^n \left[E \left(x_{n:1} x_{n:1}^\top \right) \right]_{k_1, k_2} v^\top w_{k_1} w_{k_2}^\top v \\ &= \frac{1}{n} \sum_{k_1, k_2=1}^n \sum_{h=1}^n [U_n]_{k_1, h} \lambda_h \left(E \left(x_{n:1} x_{n:1}^\top \right) \right) \overline{[U_n]_{k_2, h}} v^\top w_{k_1} w_{k_2}^\top v \\ &= \frac{1}{n} \sum_{k_1, k_2=1}^n \left(w_{k_1}^\top v \right)^\top \sum_{h=1}^n [U_n]_{k_1, h} \lambda_h \left(E \left(x_{n:1} x_{n:1}^\top \right) \right) \overline{[U_n]_{k_2, h}} w_{k_2}^\top v \\ &= \frac{1}{n} \sum_{h=1}^n \lambda_h \left(E \left(x_{n:1} x_{n:1}^\top \right) \right) \left(\sum_{k_1=1}^n w_{k_1}^\top v [U_n]_{k_1, h} \right) \left(\sum_{k_2=1}^n w_{k_2}^\top v \overline{[U_n]_{k_2, h}} \right) \\ &= \frac{1}{n} \sum_{h=1}^n \lambda_h \left(E \left(x_{n:1} x_{n:1}^\top \right) \right) \left| \sum_{l=1}^n w_l^\top v [U_n]_{l, h} \right|^2 \end{aligned}$$

with $E \left(x_{n:1} x_{n:1}^\top \right) = U_n \text{diag}_{1 \leq j \leq n} \left(\lambda_j \left(E \left(x_{n:1} x_{n:1}^\top \right) \right) \right) U_n^{-1}$ being a diagonalization of $E \left(x_{n:1} x_{n:1}^\top \right)$ where the eigenvector matrix U_n is unitary. Therefore,

$$\lambda_n \left(E \left(x_{n:1} x_{n:1}^\top \right) \right) \frac{1}{n} \sum_{h=1}^n \left| \sum_{l=1}^n w_l^\top v [U_n]_{l, h} \right|^2 \leq \lambda_r \left(E \left(\hat{y}_k (\hat{y}_k)^\top \right) \right) \leq \lambda_1 \left(E \left(x_{n:1} x_{n:1}^\top \right) \right) \frac{1}{n} \sum_{h=1}^n \left| \sum_{l=1}^n w_l^\top v [U_n]_{l, h} \right|^2.$$

To finish the proof we only need to show that

$$\frac{1}{n} \sum_{h=1}^n \left| \sum_{l=1}^n w_l^\top v [U_n]_{l, h} \right|^2 = \frac{1}{2}.$$

Applying Equation (9) and [4] (Equations (14) and (15)) yields

$$\begin{aligned} &\frac{1}{n} \sum_{h=1}^n \left| \sum_{l=1}^n w_l^\top v [U_n]_{l, h} \right|^2 \\ &= \frac{1}{n} \sum_{l=1}^n \left(w_l^\top v \right)^2 = \frac{1}{n} \sum_{l=1}^n \left(\cos \frac{2\pi(1-l)k}{n} v_1 + \sin \frac{2\pi(1-l)k}{n} v_2 \right)^2 \\ &= v_1^2 \frac{1}{n} \sum_{l=1}^n \left(\cos \frac{2\pi(1-l)k}{n} \right)^2 + v_2^2 \frac{1}{n} \sum_{l=1}^n \left(\sin \frac{2\pi(1-l)k}{n} \right)^2 + 2v_1 v_2 \frac{1}{n} \sum_{l=1}^n \cos \frac{2\pi(1-l)k}{n} \sin \frac{2\pi(1-l)k}{n} \\ &= v_1^2 \frac{1}{n} \sum_{l=1}^n \left(\cos \frac{2\pi(1-l)k}{n} \right)^2 + \frac{v_2^2}{2} + v_1 v_2 \frac{1}{n} \sum_{l=1}^n \sin \frac{4\pi(1-l)k}{n} \\ &= v_1^2 \frac{1}{n} \sum_{l=1}^n \left(1 - \left(\sin \frac{2\pi(1-l)k}{n} \right)^2 \right) + \frac{v_2^2}{2} - v_1 v_2 \frac{1}{n} \sum_{l=1}^n \sin \frac{4\pi(l-1)k}{n} \\ &= v_1^2 \left(1 - \frac{1}{n} \sum_{l=1}^n \left(\sin \frac{2\pi(1-l)k}{n} \right)^2 \right) + \frac{v_2^2}{2} - v_1 v_2 \frac{1}{n} \sum_{l=1}^n \text{Im} \left(e^{\frac{4\pi(l-1)k}{n} i} \right) \\ &= \frac{v_1^2}{2} + \frac{v_2^2}{2} - v_1 v_2 \frac{1}{n} \text{Im} \left(\sum_{l=1}^n e^{\frac{4\pi(l-1)k}{n} i} \right) = \frac{1}{2} v^\top v = \frac{1}{2}. \end{aligned}$$

□

3. RDF Upper Bounds for Real Gaussian Vectors

We first review the formula for the RDF of a real Gaussian vector given by Kolmogorov in [1].

Theorem 3. If $x_{n:1}$ is a real zero-mean Gaussian n -dimensional vector with positive definite correlation matrix, its RDF is given by

$$R_{x_{n:1}}(D) = \frac{1}{n} \sum_{k=1}^n \max \left\{ 0, \frac{1}{2} \ln \frac{\lambda_k (E(x_{n:1}x_{n:1}^\top))}{\theta} \right\} \quad \forall D \in \left(0, \frac{\text{tr}(E(x_{n:1}x_{n:1}^\top))}{n} \right],$$

where tr denotes trace and θ is a real number satisfying

$$D = \frac{1}{n} \sum_{k=1}^n \min \left\{ \theta, \lambda_k (E(x_{n:1}x_{n:1}^\top)) \right\}.$$

We recall that $R_{x_{n:1}}(D)$ can be thought of as the minimum rate (measured in nats) at which one must encode (compress) $x_{n:1}$ in order to be able to recover it with a mean square error (MSE) per dimension not larger than D , that is:

$$\frac{E \left(\|x_{n:1} - \widetilde{x}_{n:1}\|_2^2 \right)}{n} \leq D,$$

where $\widetilde{x}_{n:1}$ denotes the estimation of $x_{n:1}$ and $\|\cdot\|_2$ is the spectral norm.

The following result provides an optimal coding strategy for $x_{n:1}$ in order to achieve $R_{x_{n:1}}(D)$ whenever $D \leq \lambda_n (E(x_{n:1}x_{n:1}^\top))$. Observe that if $D \leq \lambda_n (E(x_{n:1}x_{n:1}^\top))$ then

$$R_{x_{n:1}}(D) = \frac{1}{2n} \sum_{k=1}^n \ln \frac{\lambda_k (E(x_{n:1}x_{n:1}^\top))}{D} = \frac{1}{2n} \ln \frac{\det(E(x_{n:1}x_{n:1}^\top))}{D^n}. \tag{13}$$

Corollary 1. Suppose that $x_{n:1}$ is as in Theorem 3. Let $E(x_{n:1}x_{n:1}^\top) = U_n \text{diag}_{1 \leq k \leq n}(\lambda_k (E(x_{n:1}x_{n:1}^\top))) U_n^{-1}$ be a diagonalization of $E(x_{n:1}x_{n:1}^\top)$ where the eigenvector matrix U_n is real and orthogonal. If $D \in (0, \lambda_n (E(x_{n:1}x_{n:1}^\top))]$ then

$$R_{x_{n:1}}(D) = \frac{1}{n} \sum_{k=1}^n R_{z_k}(D) = \frac{1}{2n} \sum_{k=1}^n \ln \frac{E(z_k^2)}{D} \tag{14}$$

with $z_{n:1} = U_n^\top x_{n:1}$.

Proof. We encode z_1, \dots, z_n separately with $E(\|z_k - \widetilde{z}_k\|_2^2) \leq D$ for all $k \in \{1, \dots, n\}$. Let $\widetilde{x}_{n:1} := U_n \widetilde{z}_{n:1}$, where

$$\widetilde{z}_{n:1} := \begin{pmatrix} \widetilde{z}_n \\ \vdots \\ \widetilde{z}_1 \end{pmatrix}.$$

As U_n^\top is unitary (in fact, it is a real orthogonal matrix) and the spectral norm is unitarily invariant, we have

$$\begin{aligned} \frac{E \left(\|x_{n:1} - \widetilde{x}_{n:1}\|_2^2 \right)}{n} &= \frac{E \left(\|U_n^\top x_{n:1} - U_n^\top \widetilde{x}_{n:1}\|_2^2 \right)}{n} = \frac{E \left(\|z_{n:1} - \widetilde{z}_{n:1}\|_2^2 \right)}{n} \\ &= \frac{E \left(\sum_{k=1}^n (z_k - \widetilde{z}_k)^2 \right)}{n} = \frac{\sum_{k=1}^n E \left((z_k - \widetilde{z}_k)^2 \right)}{n} = \frac{\sum_{k=1}^n E \left(\|z_k - \widetilde{z}_k\|_2^2 \right)}{n} \leq D, \end{aligned}$$

and thus,

$$R_{x_{n:1}}(D) \leq \frac{1}{n} \sum_{k=1}^n R_{z_k}(D).$$

To finish the proof we show Equation (14). Since

$$E(z_{n:1}z_{n:1}^\top) = E(U_n^\top x_{n:1}x_{n:1}^\top U_n) = U_n^\top E(x_{n:1}x_{n:1}^\top) U_n = \text{diag}_{1 \leq k \leq n}(\lambda_k(E(x_{n:1}x_{n:1}^\top))),$$

we obtain

$$E(z_k^2) = [E(z_{n:1}z_{n:1}^\top)]_{n-k+1, n-k+1} = \lambda_{n-k+1}(E(x_{n:1}x_{n:1}^\top)) \geq \lambda_n(E(x_{n:1}x_{n:1}^\top)) \geq D > 0.$$

Hence, applying Equation (13) yields

$$\begin{aligned} \frac{1}{n} \sum_{k=1}^n R_{z_k}(D) &= \frac{1}{n} \sum_{k=1}^n \frac{1}{2} \ln \frac{E(z_k^2)}{D} = \frac{1}{2n} \sum_{k=1}^n \ln \frac{\lambda_{n-k+1}(E(x_{n:1}x_{n:1}^\top))}{D} \\ &= \frac{1}{2n} \sum_{k=1}^n \ln \frac{\lambda_k(E(x_{n:1}x_{n:1}^\top))}{D} = R_{x_{n:1}}(D). \end{aligned}$$

□

Corollary 1 shows that an optimal coding strategy for $x_{n:1}$ is to encode z_1, \dots, z_n separately.

We now give two coding strategies for $x_{n:1}$ based on the DFT whose computational complexity is lower than the computational complexity of the optimal coding strategy provided in Corollary 1.

Theorem 4. Let $x_{n:1}$ be as in Theorem 3. Suppose that $y_{n:1}$ is the DFT of $x_{n:1}$ and $D \in (0, \lambda_n(E(x_{n:1}x_{n:1}^\top))]$. Then

$$R_{x_{n:1}}(D) \leq \tilde{R}_{x_{n:1}}(D) \leq \check{R}_{x_{n:1}}(D) \leq \frac{1}{2n} \sum_{k=1}^n \ln \frac{E(|y_k|^2)}{D} \tag{15}$$

$$\leq R_{x_{n:1}}(D) + \frac{1}{2} \ln \left(1 + \frac{\|E(x_{n:1}x_{n:1}^\top) - V_n \text{diag}_{1 \leq k \leq n}([V_n^* E(x_{n:1}x_{n:1}^\top) V_n]_{k,k}) V_n^*\|_F}{\sqrt{n} \lambda_n(E(x_{n:1}x_{n:1}^\top))} \right), \tag{16}$$

where $\|\cdot\|_F$ is the Frobenius norm,

$$\tilde{R}_{x_{n:1}}(D) := \begin{cases} \frac{R_{y_{\frac{n}{2}}}(D) + 2 \sum_{k=\frac{n}{2}+1}^{n-1} R_{y_k}(\frac{D}{2}) + R_{y_n}(D)}{n} & \text{if } n \text{ is even,} \\ \frac{2 \sum_{k=\frac{n+1}{2}}^{n-1} R_{y_k}(\frac{D}{2}) + R_{y_n}(D)}{n} & \text{if } n \text{ is odd,} \end{cases}$$

and

$$\check{R}_{x_{n:1}}(D) := \begin{cases} \frac{R_{y_{\frac{n}{2}}}(D) + \sum_{k=\frac{n}{2}+1}^{n-1} (R_{\text{Re}(y_k)}(\frac{D}{2}) + R_{\text{Im}(y_k)}(\frac{D}{2})) + R_{y_n}(D)}{n} & \text{if } n \text{ is even,} \\ \frac{\sum_{k=\frac{n+1}{2}}^{n-1} (R_{\text{Re}(y_k)}(\frac{D}{2}) + R_{\text{Im}(y_k)}(\frac{D}{2})) + R_{y_n}(D)}{n} & \text{if } n \text{ is odd.} \end{cases}$$

Proof. Equations (15) and (16) were presented in [4] (Equations (16) and (20)) for the case where the correlation matrix $E(x_{n:1}x_{n:1}^\top)$ is Toeplitz. They were proved by using a result on the DFT of random vectors with Toeplitz correlation matrix, namely, ref. [4] (Theorem 1). The proof of Theorem 4 is similar to the proof of [4] (Equations (16) and (20)) but using Theorem 1 instead of [4] (Theorem 1). Observe that in Theorems 1 and 4 no assumption about the structure of $E(x_{n:1}x_{n:1}^\top)$ has been made. □

Theorem 4 shows that a coding strategy for $x_{n:1}$ is to encode $y_{\lceil \frac{n}{2} \rceil}, \dots, y_n$ separately, where $\lceil \frac{n}{2} \rceil$ denotes the smallest integer higher than or equal to $\frac{n}{2}$. Theorem 4 also shows that another coding strategy for $x_{n:1}$ is to encode separately the real part and the imaginary part of y_k instead of encoding y_k when $k \in \{\lceil \frac{n}{2} \rceil, \dots, n-1\} \setminus \{\frac{n}{2}\}$. The computational complexity of these two coding strategies based on the DFT is lower than the computational complexity of the optimal coding strategy provided in

Corollary 1. Specifically, the complexity of computing the DFT ($y_{n:1} = V_n^* x_{n:1}$) is $O(n \log n)$ whenever the fast Fourier transform (FFT) algorithm is used, while the complexity of computing $z_{n:1} = U_n^\top x_{n:1}$ is $O(n^2)$. Moreover, when the coding strategies based on the DFT are used, we do not need to compute a real orthogonal eigenvector matrix U_n of $E(x_{n:1} x_{n:1}^\top)$. It should also be mentioned that for these coding strategies based on the DFT the knowledge of $E(x_{n:1} x_{n:1}^\top)$ is not even required, in fact, for them we only need to know $E(\hat{y}_k (\hat{y}_k)^\top)$ with $k \in \{\lceil \frac{n}{2} \rceil, \dots, n\}$.

The rates corresponding to the two coding strategies given in Theorem 4, $\tilde{R}_{x_{n:1}}(D)$ and $\check{R}_{x_{n:1}}(D)$, can be written in terms of $E(x_{n:1} x_{n:1}^\top)$ and V_n by using Lemma 1 and the following lemma.

Lemma 2. Let $y_{n:1}$ and D be as in Theorem 4. Then

1. $R_{y_k}(D) = \frac{1}{2} \ln \frac{E(y_k^2)}{D}$ for all $k \in \{1, \dots, n\} \cap \{\frac{n}{2}, n\}$.
2. $R_{\hat{y}_k}(\frac{D}{2}) = \frac{1}{4} \ln \frac{E((\text{Re}(y_k))^2)E((\text{Im}(y_k))^2) - (E(\text{Re}(y_k)\text{Im}(y_k)))^2}{(\frac{D}{2})^2}$ for all $k \in \{1, \dots, n-1\} \setminus \{\frac{n}{2}\}$.
3. $R_{\text{Re}(y_k)}(\frac{D}{2}) = \frac{1}{2} \ln \frac{E((\text{Re}(y_k))^2)}{\frac{D}{2}}$ for all $k \in \{1, \dots, n-1\} \setminus \{\frac{n}{2}\}$.
4. $R_{\text{Im}(y_k)}(\frac{D}{2}) = \frac{1}{2} \ln \frac{E((\text{Im}(y_k))^2)}{\frac{D}{2}}$ for all $k \in \{1, \dots, n-1\} \setminus \{\frac{n}{2}\}$.

Proof. (1) Applying Equation (2) and [4] (Lemma 1) yields

$$0 < D \leq \lambda_n(E(x_{n:1} x_{n:1}^\top)) \leq E(|y_k|^2) = E(y_k^2).$$

Assertion (1) now follows directly from Equation (13).

(2) Applying Theorem 2 we have

$$0 < \frac{D}{2} \leq \frac{\lambda_n(E(x_{n:1} x_{n:1}^\top))}{2} \leq \lambda_2(E(\hat{y}_k (\hat{y}_k)^\top)).$$

Consequently, from Equation (13) we obtain

$$R_{\hat{y}_k}(\frac{D}{2}) = \frac{1}{4} \ln \frac{\det(E(\hat{y}_k (\hat{y}_k)^\top))}{(\frac{D}{2})^2} = \frac{1}{4} \ln \frac{\det \begin{pmatrix} E((\text{Re}(y_k))^2) & E(\text{Re}(y_k)\text{Im}(y_k)) \\ E(\text{Im}(y_k)\text{Re}(y_k)) & E((\text{Im}(y_k))^2) \end{pmatrix}}{(\frac{D}{2})^2}.$$

Assertions (3) and (4) Applying Equations (3) and (4) yields

$$0 < \frac{D}{2} \leq \frac{\lambda_n(E(x_{n:1} x_{n:1}^\top))}{2} \leq E((\text{Re}(y_k))^2).$$

and

$$0 < \frac{D}{2} \leq \frac{\lambda_n(E(x_{n:1} x_{n:1}^\top))}{2} \leq E((\text{Im}(y_k))^2).$$

Assertions (3) and (4) now follow directly from Equation (13). \square

We end this section with a result that is a direct consequence of Lemma 2. This result shows when the rates corresponding to the two coding strategies given in Theorem 4, $\tilde{R}_{x_{n:1}}(D)$ and $\check{R}_{x_{n:1}}(D)$, are equal.

Lemma 3. Let $x_{n:1}$, $y_{n:1}$, and D be as in Theorem 4. Then the two following assertions are equivalent:

1. $\tilde{R}_{x_{n:1}}(D) = \check{R}_{x_{n:1}}(D)$.
2. $E(\text{Re}(y_k)\text{Im}(y_k)) = 0$ for all $k \in \{\lceil \frac{n}{2} \rceil, \dots, n-1\} \setminus \{\frac{n}{2}\}$.

Proof. Fix $k \in \{\lceil \frac{n}{2} \rceil, \dots, n-1\} \setminus \{\frac{n}{2}\}$. From Lemma 2 we have

$$\begin{aligned} 2R_{\hat{y}_k} \left(\frac{D}{2} \right) &= \frac{1}{2} \ln \frac{E \left((\operatorname{Re}(y_k))^2 \right) E \left((\operatorname{Im}(y_k))^2 \right) - (E(\operatorname{Re}(y_k) \operatorname{Im}(y_k)))^2}{\left(\frac{D}{2} \right)^2} \\ &\leq \frac{1}{2} \ln \frac{E \left((\operatorname{Re}(y_k))^2 \right) E \left((\operatorname{Im}(y_k))^2 \right)}{\left(\frac{D}{2} \right)^2} \\ &= \frac{1}{2} \ln \frac{E \left((\operatorname{Re}(y_k))^2 \right)}{\frac{D}{2}} + \frac{1}{2} \ln \frac{E \left((\operatorname{Im}(y_k))^2 \right)}{\frac{D}{2}} \\ &= R_{\operatorname{Re}(y_k)} \left(\frac{D}{2} \right) + R_{\operatorname{Im}(y_k)} \left(\frac{D}{2} \right). \end{aligned}$$

□

4. Low-Complexity Coding Strategies for Gaussian AWSS AR Sources

We begin by introducing some notation. The symbols \mathbb{N} , \mathbb{Z} , and \mathbb{R} denote the set of positive integers, integers, and (finite) real numbers, respectively. If $f: \mathbb{R} \rightarrow \mathbb{C}$ is continuous and 2π -periodic, we denote by $T_n(f)$ the $n \times n$ Toeplitz matrix given by

$$[T_n(f)]_{j,k} = t_{j-k},$$

where $\{t_k\}_{k \in \mathbb{Z}}$ is the sequence of Fourier coefficients of f , i.e.,

$$t_k = \frac{1}{2\pi} \int_0^{2\pi} f(\omega) e^{-k\omega i} d\omega \quad \forall k \in \mathbb{Z}.$$

If A_n and B_n are $n \times n$ matrices for all $n \in \mathbb{N}$, we write $\{A_n\} \sim \{B_n\}$ if the sequences $\{A_n\}$ and $\{B_n\}$ are asymptotically equivalent, that is, $\{\|A_n\|_2\}$ and $\{\|B_n\|_2\}$ are bounded and $\lim_{n \rightarrow \infty} \frac{\|A_n - B_n\|_F}{\sqrt{n}} = 0$ (see [5] (Section 2.3) and [6]).

We now review the definitions of AWSS processes and AR processes.

Definition 1. A random process $\{x_n\}$ is said to be AWSS if it has constant mean (i.e., $E(x_j) = E(x_k)$ for all $j, k \in \mathbb{N}$) and there exists a continuous 2π -periodic function $f: \mathbb{R} \rightarrow \mathbb{C}$ such that $\{E(x_{n:1} x_{n:1}^*)\} \sim \{T_n(f)\}$. The function f is called (asymptotic) PSD of $\{x_n\}$.

Definition 2. A real zero-mean random process $\{x_n\}$ is said to be AR if

$$x_n = w_n - \sum_{k=1}^{n-1} a_{-k} x_{n-k} \quad \forall n \in \mathbb{N},$$

or equivalently,

$$\sum_{k=0}^{n-1} a_{-k} x_{n-k} = w_n \quad \forall n \in \mathbb{N}, \tag{17}$$

where $a_0 = 1$, $a_{-k} \in \mathbb{R}$ for all $k \in \mathbb{N}$, and $\{w_n\}$ is a real zero-mean random process satisfying that $E(w_j w_k) = \delta_{j,k} \sigma^2$ for all $j, k \in \mathbb{N}$ with $\sigma^2 > 0$ and $\delta_{j,k}$ being the Kronecker delta (i.e., $\delta_{j,k} = 1$ if $j = k$, and it is zero otherwise).

The AR process $\{x_n\}$ in Equation (17) is of finite order if there exists $p \in \mathbb{N}$ such that $a_{-k} = 0$ for all $k > p$. In this case, $\{x_n\}$ is called an AR(p) process.

The following theorem shows that if $x_{n:1}$ is a large enough data block of a Gaussian AWSS AR source, the rate does not increase whenever we encode it using the two coding strategies based on the DFT presented in Section 3, instead of encoding $x_{n:1}$ using an eigenvector matrix of its correlation matrix.

Theorem 5. Let $\{x_n\}$ be as in Definition 2. Suppose that $\{a_k\}_{k \in \mathbb{Z}}$, with $a_k = 0$ for all $k \in \mathbb{N}$, is the sequence of Fourier coefficients of a function $a: \mathbb{R} \rightarrow \mathbb{C}$ which is continuous and 2π -periodic. Then

1. $\inf_{n \in \mathbb{N}} \lambda_n (E(x_{n:1} x_{n:1}^\top)) \geq \frac{\sigma^2}{\max_{\omega \in [0, 2\pi]} |a(\omega)|^2} > 0$.
2. Consider $D \in (0, \inf_{n \in \mathbb{N}} \lambda_n (E(x_{n:1} x_{n:1}^\top))]$.

(a) If $\{x_n\}$ is Gaussian,

$$\frac{1}{2} \ln \frac{\sigma^2}{D} = R_{x_{n:1}}(D) \leq \tilde{R}_{x_{n:1}}(D) \leq \check{R}_{x_{n:1}}(D) \leq K_1(n, D) \leq K_2(n, D) \leq K_3(n, D) \quad \forall n \in \mathbb{N}, \quad (18)$$

where $K_1(n, D)$ is given by Equation (16), and $K_2(n, D)$ and $K_3(n, D)$ are obtained by replacing $\lambda_n (E(x_{n:1} x_{n:1}^\top))$ in Equation (16) by $\inf_{n \in \mathbb{N}} \lambda_n (E(x_{n:1} x_{n:1}^\top))$ and $\frac{\sigma^2}{\max_{\omega \in [0, 2\pi]} |a(\omega)|^2}$, respectively.

(b) If $\{x_n\}$ is Gaussian and AWSS,

$$\lim_{n \rightarrow \infty} R_{x_{n:1}}(D) = \lim_{n \rightarrow \infty} \tilde{R}_{x_{n:1}}(D) = \lim_{n \rightarrow \infty} \check{R}_{x_{n:1}}(D) = \lim_{n \rightarrow \infty} K_3(n, D). \quad (19)$$

Proof. (1) Equation (17) can be rewritten as

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

Consequently,

$$T_n(a)E(x_{n:1} x_{n:1}^\top) (T_n(a))^\top = E(T_n(a)x_{n:1} (T_n(a)x_{n:1})^\top) = E(w_{n:1} w_{n:1}^\top) = \sigma^2 I_n \quad \forall n \in \mathbb{N}.$$

As $\det(T_n(a)) = 1$, $T_n(a)$ is invertible, and therefore,

$$\begin{aligned} E(x_{n:1} x_{n:1}^\top) &= \sigma^2 (T_n(a))^{-1} ((T_n(a))^\top)^{-1} = \sigma^2 ((T_n(a))^\top T_n(a))^{-1} = \sigma^2 ((T_n(a))^* T_n(a))^{-1} \\ &= \sigma^2 (N_n \text{diag}_{1 \leq k \leq n} ((\sigma_k(T_n(a)))^2) N_n^*)^{-1} = N_n \text{diag}_{1 \leq k \leq n} \left(\frac{\sigma^2}{(\sigma_k(T_n(a)))^2} \right) N_n^* \end{aligned} \quad (20)$$

for all $n \in \mathbb{N}$, where $T_n(a) = M_n \text{diag}_{1 \leq k \leq n} (\sigma_k(T_n(a))) N_n^*$ is a singular value decomposition of $T_n(a)$. Thus, applying [8] (Theorem 4.3) yields

$$\lambda_n (E(x_{n:1} x_{n:1}^\top)) = \frac{\sigma^2}{(\sigma_1(T_n(a)))^2} \geq \frac{\sigma^2}{\max_{\omega \in [0, 2\pi]} |a(\omega)|^2} > 0 \quad \forall n \in \mathbb{N}.$$

(2a) From Equation (13) we have

$$\begin{aligned} R_{x_{n:1}}(D) &= \frac{1}{2n} \ln \frac{\det(E(x_{n:1} x_{n:1}^\top))}{D^n} = \frac{1}{2n} \ln \frac{\det(\sigma^2 (T_n(a))^{-1} ((T_n(a))^\top)^{-1})}{D^n} \\ &= \frac{1}{2n} \ln \frac{(\sigma^2)^n}{D^n \det(T_n(a)) \det((T_n(a))^\top)} = \frac{1}{2n} \ln \frac{(\sigma^2)^n}{D^n} = \frac{1}{2} \ln \frac{\sigma^2}{D} \quad \forall n \in \mathbb{N}. \end{aligned}$$

Assertion (2a) now follows from Theorem 4 and Assertion (1).

(2b) From Assertion (2a) we only need to show that

$$\lim_{n \rightarrow \infty} \frac{\|E(x_{n:1}x_{n:1}^\top) - V_n \text{diag}_{1 \leq k \leq n}([V_n^* E(x_{n:1}x_{n:1}^\top) V_n]_{k,k}) V_n^*\|_F}{\sqrt{n}} = 0. \tag{21}$$

As the Frobenius norm is unitarily invariant we obtain

$$\begin{aligned} 0 &\leq \frac{\|E(x_{n:1}x_{n:1}^\top) - V_n \text{diag}_{1 \leq k \leq n}([V_n^* E(x_{n:1}x_{n:1}^\top) V_n]_{k,k}) V_n^*\|_F}{\sqrt{n}} \\ &\leq \frac{\|E(x_{n:1}x_{n:1}^\top) - T_n(f)\|_F}{\sqrt{n}} + \frac{\|T_n(f) - \widehat{C}_n(f)\|_F}{\sqrt{n}} + \frac{\|V_n \text{diag}_{1 \leq k \leq n}([V_n^* E(x_{n:1}x_{n:1}^\top) V_n]_{k,k}) V_n^* - \widehat{C}_n(f)\|_F}{\sqrt{n}} \\ &= \frac{\|E(x_{n:1}x_{n:1}^\top) - T_n(f)\|_F}{\sqrt{n}} + \frac{\|T_n(f) - \widehat{C}_n(f)\|_F}{\sqrt{n}} + \frac{\|V_n \text{diag}_{1 \leq k \leq n}([V_n^* (E(x_{n:1}x_{n:1}^\top) - T_n(f)) V_n]_{k,k}) V_n^*\|_F}{\sqrt{n}} \\ &= \frac{\|E(x_{n:1}x_{n:1}^\top) - T_n(f)\|_F}{\sqrt{n}} + \frac{\|T_n(f) - \widehat{C}_n(f)\|_F}{\sqrt{n}} + \frac{\|\text{diag}_{1 \leq k \leq n}([V_n^* (E(x_{n:1}x_{n:1}^\top) - T_n(f)) V_n]_{k,k})\|_F}{\sqrt{n}} \\ &\leq \frac{\|E(x_{n:1}x_{n:1}^\top) - T_n(f)\|_F}{\sqrt{n}} + \frac{\|T_n(f) - \widehat{C}_n(f)\|_F}{\sqrt{n}} + \frac{\|V_n^* (E(x_{n:1}x_{n:1}^\top) - T_n(f)) V_n\|_F}{\sqrt{n}} \\ &= 2 \frac{\|E(x_{n:1}x_{n:1}^\top) - T_n(f)\|_F}{\sqrt{n}} + \frac{\|T_n(f) - \widehat{C}_n(f)\|_F}{\sqrt{n}}, \end{aligned}$$

where f is (asymptotic) PSD of $\{x_n\}$ and $\widehat{C}_n(f) = V_n \text{diag}_{1 \leq k \leq n}([V_n^* T_n(f) V_n]_{k,k}) V_n^*$. Assertion (2b) now follows from $\{E(x_{n:1}x_{n:1}^\top)\} \sim \{T_n(f)\}$ and [9] (Lemma 4.2). \square

If $\sum_{k=-\infty}^0 |a_k| < \infty$, there always exists such function a and it is given by $a(\omega) = \sum_{k=-\infty}^0 a_k e^{k\omega i}$ for all $\omega \in \mathbb{R}$ (see, e.g., [8] (Appendix B)). In particular, if $\{x_n\}$ is an AR(p) process, $a(\omega) = \sum_{k=-p}^0 a_k e^{k\omega i}$ for all $\omega \in \mathbb{R}$.

5. Sufficient Conditions for AR Processes to be AWSS

In the following two results we give sufficient conditions for AR processes to be AWSS.

Theorem 6. Let $\{x_n\}$ be as in Definition 2. Suppose that $\{a_k\}_{k \in \mathbb{Z}}$, with $a_k = 0$ for all $k \in \mathbb{N}$, is the sequence of Fourier coefficients of a function $a: \mathbb{R} \rightarrow \mathbb{C}$ which is continuous and 2π -periodic. Then the following assertions are equivalent:

1. $\{x_n\}$ is AWSS.
2. $\{\|E(x_{n:1}x_{n:1}^\top)\|_2\}$ is bounded.
3. $\{T_n(a)\}$ is stable (that is, $\{\|(T_n(a))^{-1}\|_2\}$ is bounded).
4. $a(\omega) \neq 0$ for all $\omega \in \mathbb{R}$ and $\{x_n\}$ is AWSS with (asymptotic) PSD $\frac{\sigma^2}{|a|^2}$.

Proof. (1) \Rightarrow (2) This is a direct consequence of the definition of AWSS process, i.e., of Definition 1.

(2) \Leftrightarrow (3) From Equation (20) we have

$$\|E(x_{n:1}x_{n:1}^\top)\|_2 = \frac{\sigma^2}{(\sigma_n(T_n(a)))^2} = \sigma^2 \left\| N_n \text{diag}_{1 \leq k \leq n} \left(\frac{1}{\sigma_k(T_n(a))} \right) M_n^* \right\|_2^2 = \sigma^2 \|(T_n(a))^{-1}\|_2^2$$

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

(3) \Rightarrow (4) It is well known that if $f: \mathbb{R} \rightarrow \mathbb{C}$ is continuous and 2π -periodic, and $\{T_n(f)\}$ is stable then $f(\omega) \neq 0$ for all $\omega \in \mathbb{R}$. Hence, $a(\omega) \neq 0$ for all $\omega \in \mathbb{R}$.

Applying [8] (Lemma 4.2.1) yields $\{(T_n(a))^\top\} = \{(T_n(a))^*\} = \{T_n(\bar{a})\}$. Consequently, from [7] (Theorem 3) we obtain

$$\{(T_n(a))^\top T_n(a)\} = \{T_n(\bar{a})T_n(a)\} \sim \{T_n(\bar{a}a)\} = \{T_n(|a|^2)\}.$$

Observe that the sequence

$$\left\{ \left\| \left((T_n(a))^\top T_n(a) \right)^{-1} \right\|_2 \right\} = \left\{ \left\| \frac{1}{\sigma^2} E \left(x_{n:1} x_{n:1}^\top \right) \right\|_2 \right\} = \left\{ \frac{1}{\sigma^2} \left\| E \left(x_{n:1} x_{n:1}^\top \right) \right\|_2 \right\}$$

is bounded. As the function $|a|^2$ is real, applying [8] (Theorem 4.4) we have that $T_n(|a|^2)$ is Hermitian and $0 < \min_{\omega \in [0, 2\pi]} |a(\omega)|^2 \leq \lambda_n(T_n(|a|^2))$ for all $n \in \mathbb{N}$, and therefore,

$$\left\| \left(T_n(|a|^2) \right)^{-1} \right\|_2 = \frac{1}{\lambda_n(T_n(|a|^2))} \leq \frac{1}{\min_{\omega \in [0, 2\pi]} |a(\omega)|^2} \quad \forall n \in \mathbb{N}.$$

Thus, from [5] (Theorem 1.4) we obtain

$$\left\{ \frac{1}{\sigma^2} E \left(x_{n:1} x_{n:1}^\top \right) \right\} = \left\{ \left((T_n(a))^\top T_n(a) \right)^{-1} \right\} \sim \left\{ \left(T_n(|a|^2) \right)^{-1} \right\}.$$

Hence, applying [10] (Theorem 4.2) and [5] (Theorem 1.2) yields

$$\left\{ \frac{1}{\sigma^2} E \left(x_{n:1} x_{n:1}^\top \right) \right\} \sim \left\{ T_n \left(\frac{1}{|a|^2} \right) \right\}.$$

Consequently, from [8] (Lemma 3.1.3) and [8] (Lemma 4.2.3) we have

$$\left\{ E \left(x_{n:1} x_{n:1}^\top \right) \right\} \sim \left\{ \sigma^2 T_n \left(\frac{1}{|a|^2} \right) \right\} = \left\{ T_n \left(\frac{\sigma^2}{|a|^2} \right) \right\}.$$

(4) \Rightarrow (1) It is obvious. \square

Corollary 2. Let $\{x_n\}$ be as in Definition 2 with $\sum_{k=-\infty}^0 |a_k| < \infty$. If $\sum_{k=0}^\infty a_{-k} z^k \neq 0$ for all $|z| \leq 1$ then $\{x_n\}$ is AWSS.

Proof. It is well known that if a sequence of complex numbers $\{t_k\}_{k \in \mathbb{Z}}$ satisfies that $\sum_{k=-\infty}^\infty |t_k| < \infty$ and that $\sum_{k=-\infty}^\infty t_k z^k \neq 0$ for all $|z| \leq 1$ then $\{T_n(f)\}$ is stable with $f(\omega) = \sum_{k=-\infty}^\infty t_k e^{k\omega i}$ for all $\omega \in \mathbb{R}$. Therefore, $\{T_n(b)\}$ is stable with $b(\omega) = \sum_{k=0}^\infty a_{-k} e^{k\omega i}$ for all $\omega \in \mathbb{R}$. Thus,

$$\left\{ \left\| (T_n(a))^{-1} \right\|_2 \right\} = \left\{ \left\| \left((T_n(a))^{-1} \right)^\top \right\|_2 \right\} = \left\{ \left\| \left((T_n(a))^\top \right)^{-1} \right\|_2 \right\} = \left\{ \left\| (T_n(b))^{-1} \right\|_2 \right\}$$

is bounded with $a(\omega) = \sum_{k=-\infty}^0 a_k e^{k\omega i}$ for all $\omega \in \mathbb{R}$. As $\{T_n(a)\}$ is stable, from Theorem 6 we conclude that $\{x_n\}$ is AWSS. \square

6. Numerical Example and Conclusions

6.1. Example

Let $\{x_n\}$ be as in Definition 2 with $a_{-k} = 0$ for all $k > 1$. Observe that $\frac{\sigma^2}{\max_{\omega \in [0, 2\pi]} |a(\omega)|^2} = \frac{\sigma^2}{(1+|a_{-1}|)^2}$. If $|a_{-1}| < 1$ from Corollary 2 we obtain that the AR(1) process $\{x_n\}$ is AWSS. Figure 1 shows $R_{x_{n:1}}(D)$, $\tilde{R}_{x_{n:1}}(D)$, and $\check{R}_{x_{n:1}}(D)$ by assuming that $\{x_n\}$ is Gaussian, $a_{-1} = -\frac{1}{2}$, $\sigma^2 = 1$, $D = \frac{\sigma^2}{(1+|a_{-1}|)^2} = \frac{4}{9}$, and $n \leq 100$. Figure 1 also shows the highest upper bound of $R_{x_{n:1}}(D)$ presented in Theorem 5,

namely, $K_3(n, D)$. Observe that the figure bears evidence of the equalities and inequalities given in Equations (18) and (19).

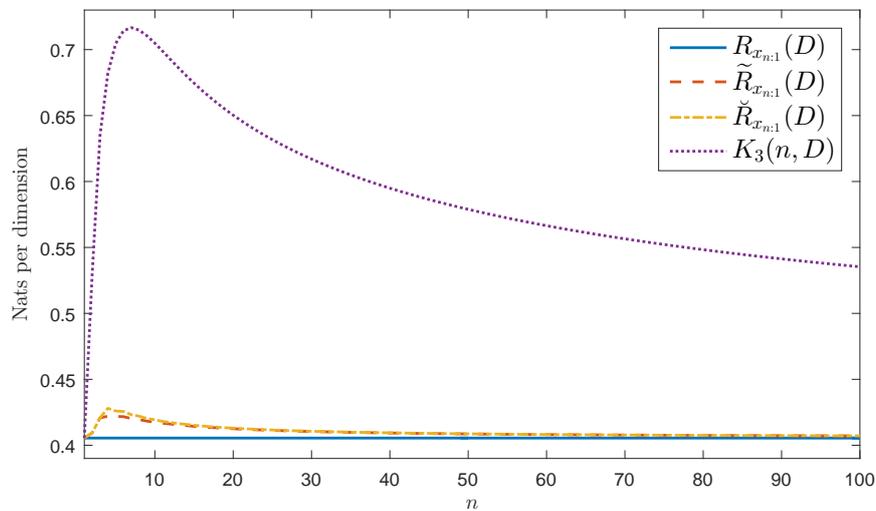


Figure 1. Considered rates for a Gaussian AWSS AR(1) source.

6.2. Conclusions

The computational complexity of coding finite-length data blocks of Gaussian sources can be reduced by using any of the two low-complexity coding strategies here presented instead of the optimal coding strategy. Moreover, the rate does not increase if we use those strategies instead of the optimal one whenever the Gaussian source is AWSS and AR, and the considered data block is large enough.

Author Contributions: Authors are listed in order of their degree of involvement in the work, with the most active contributors listed first. All authors have read and approved the final manuscript.

Funding: This work was supported in part by the Spanish Ministry of Economy and Competitiveness through the CARMEN project (TEC2016-75067-C4-3-R).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Kolmogorov, A.N. On the Shannon theory of information transmission in the case of continuous signals. *IRE Trans. Inf. Theory* **1956**, *2*, 102–108. [\[CrossRef\]](#)
2. Gray, R.M. Information rates of autoregressive processes. *IEEE Trans. Inf. Theory* **1970**, *16*, 412–421. [\[CrossRef\]](#)
3. Pearl, J. On coding and filtering stationary signals by discrete Fourier transforms. *IEEE Trans. Inf. Theory* **1973**, *19*, 229–232. [\[CrossRef\]](#)
4. Gutiérrez-Gutiérrez, J.; Zárrega-Rodríguez, M.; Insausti, X. Upper bounds for the rate distortion function of finite-length data blocks of Gaussian WSS sources. *Entropy* **2017**, *19*, 554. [\[CrossRef\]](#)
5. Gray, R.M. Toeplitz and circulant matrices: A review. *Found. Trends Commun. Inf. Theory* **2006**, *2*, 155–239. [\[CrossRef\]](#)
6. Gray, R.M. On the asymptotic eigenvalue distribution of Toeplitz matrices. *IEEE Trans. Inf. Theory* **1972**, *18*, 725–730. [\[CrossRef\]](#)
7. Gutiérrez-Gutiérrez, J.; Crespo, P.M. Asymptotically equivalent sequences of matrices and multivariate ARMA processes. *IEEE Trans. Inf. Theory* **2011**, *57*, 5444–5454. [\[CrossRef\]](#)
8. Gutiérrez-Gutiérrez, J.; Crespo, P.M. Block Toeplitz matrices: Asymptotic results and applications. *Found. Trends Commun. Inf. Theory* **2011**, *8*, 179–257. [\[CrossRef\]](#)

9. 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](#)]
10. Gutiérrez-Gutiérrez, J.; Crespo, P.M. Asymptotically equivalent sequences of matrices and Hermitian block Toeplitz matrices with continuous symbols: Applications to MIMO systems. *IEEE Trans. Inf. Theory* **2008**, *54*, 5671–5680. [[CrossRef](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).