

Article

# On the Rate-Distortion Function of Sampled Cyclostationary Gaussian Processes

Emeka Abakasanga <sup>1</sup>, Nir Shlezinger <sup>2</sup> , and Ron Dabora <sup>1,\*</sup> 

<sup>1</sup> Department of Electrical and Computer Engineering, Ben-Gurion University, Be'er-Sheva 8410501, Israel; abakasan@post.bgu.ac.il

<sup>2</sup> Faculty of Mathematics and Computer Science, Weizmann Institute of Science, Rehovot 7610001, Israel; nirshlezinger1@gmail.com

\* Correspondence: ron@ee.bgu.ac.il.

Received: 14 February 2020; Accepted: 10 March 2020; Published: 17 March 2020



**Abstract:** Man-made communications signals are typically modelled as continuous-time (CT) wide-sense cyclostationary (WSCS) processes. As modern processing is digital, it is applied to discrete-time (DT) processes obtained by sampling the CT processes. When sampling is applied to a CT WSCS process, the statistics of the resulting DT process depends on the relationship between the sampling interval and the period of the statistics of the CT process: When these two parameters have a common integer factor, then the DT process is WSCS. This situation is referred to as synchronous sampling. When this is not the case, which is referred to as asynchronous sampling, the resulting DT process is wide-sense almost cyclostationary (WSACS). The sampled CT processes are commonly encoded using a source code to facilitate storage or transmission over wireless networks, e.g., using compress-and-forward relaying. In this work, we study the fundamental tradeoff between rate and distortion for source codes applied to sampled CT WSCS processes, characterized via the rate-distortion function (RDF). We note that while RDF characterization for the case of synchronous sampling directly follows from classic information-theoretic tools utilizing ergodicity and the law of large numbers, when sampling is asynchronous, the resulting process is not information stable. In such cases, the commonly used information-theoretic tools are inapplicable to RDF analysis, which poses a major challenge. Using the information-spectrum framework, we show that the RDF for asynchronous sampling in the low distortion regime can be expressed as the limit superior of a sequence of RDFs in which each element corresponds to the RDF of a synchronously sampled WSCS process (yet their limit is not guaranteed to exist). The resulting characterization allows us to introduce novel insights on the relationship between sampling synchronization and the RDF. For example, we demonstrate that, differently from stationary processes, small differences in the sampling rate and the sampling time offset can notably affect the RDF of sampled CT WSCS processes.

**Keywords:** wide-sense cyclostationary; wide-sense almost cyclostationary; rate-distortion function; information spectrum

## 1. Introduction

Man-made signals are typically generated using a repetitive procedure, which takes place at fixed intervals. The resulting signals are thus commonly modeled as continuous-time (CT) random processes exhibiting periodic statistical properties [1–3], which are referred to as wide-sense cyclostationary (WSCS) processes. In digital communications, where the transmitted waveforms commonly obey the WSCS model [3], the received CT signal is first sampled to obtain a discrete-time (DT) received signal. In the event that the sampling interval is commensurate with the period of the statistics of the CT WSCS signal, cyclostationarity is preserved in DT ([3] Section 3.9). In this work, we refer to

this situation as synchronous sampling. However, it is practically common to encounter scenarios in which the sampling rate at the receiver and the symbol rate of the received CT WSCS process are incommensurate, which is referred to as asynchronous sampling. The resulting sampled process in such cases is a DT wide-sense almost cyclostationary (WSACS) stochastic process ([3] Section 3.9).

This research aims at investigating lossy source coding for asynchronously sampled CT WSCS processes. In the source coding problem, every sequence of information symbols from the source is mapped into a sequence of code symbols, referred to as codewords, taken from a predefined codebook. In lossy source coding, the source sequence is recovered up to a predefined distortion constraint, within an arbitrary small tolerance of error. The figure-of-merit for lossy source coding is the rate-distortion function (RDF) which characterizes the minimum number of bits per source symbol required to compress the source sequence such that it can be reconstructed at the decoder within the specified maximal distortion [4]. For an independent and identically distributed (IID) random source process, the RDF can be expressed as the minimum mutual information between the source variable and the reconstruction variable, such that with the corresponding conditional distribution of the reconstruction symbol given the source symbol, the distortion constraint is satisfied ([5] Chapter 10). The source coding problem has been further studied in multiple different scenarios, including the reconstruction of a single source at multiple destinations [6] and the reconstruction of multiple correlated stationary Gaussian sources at a single destination [7–9].

For sampled stationary source processes, ergodicity theory and the asymptotic equipartition property (AEP) ([5] Chapter 3) were utilized for characterizing the RDF in different scenarios ([10] Chapter 9), ([4] Section I), [11]. However, as in a broad range of applications, including digital communications networks, the CT signals are WSCS processes, the sampling operation results in DT source signals whose statistics depends on the relationship between the sampling rate and the period of the statistics of the source signal. When sampling is synchronous, the resulting DT source signal is WSCS ([3] Section 3.9). The RDF for lossy compression of DT WSCS Gaussian sources with memory was studied in [12]. The work [12] used the fact that any WSCS signal can be transformed into a set of stationary subprocess [2]; thereby facilitating the application of information-theoretic results obtained for multivariate stationary sources to the derivation of the RDF; Nonetheless, in many digital communications scenarios, the sampling rate and the symbol rate of the CT WSCS process are not related in any way, and are possibly incommensurate, resulting in a sampled process which is a DT WSACS stochastic process. Such situations can occur as a result of the a-priori determined values of the sampling interval and the symbol duration of the WSCS source signal, as well as due to sampling clock jitters resulting from hardware impairments. A comprehensive review of trends and applications for almost cyclostationary signals can be found in [13]. Despite of their apparent frequent occurrences, the RDF for lossy compression of WSACS sources has not been characterized, which is the motivation for the current research. A major challenge associated with characterizing fundamental limits for asynchronously sampled WSCS processes stems from the fact that the resulting processes are not information stable, in the sense that their conditional distributions are not ergodic ([14] Page X), [15,16]. As a result, the standard information-theoretic tools cannot be employed, making the characterization of the RDF a very challenging problem.

Our recent study in [17] on channel coding reveals that for the case of additive CT WSCS Gaussian noise, capacity varies significantly with sampling rates, whether the Nyquist criterion is satisfied or not. In particular, it was observed that the capacity can change dramatically with minor variations in the sampling rate, causing it to switch from synchronous sampling to asynchronous sampling. This is in direct contrast to the results obtained for wide-sense stationary noise for which the capacity remains unchanged for any sampling rate above the Nyquist rate [18]. A natural fundamental question that arises from this result is how the RDF of a sampled Gaussian source process varies with the sampling rate. As a motivating example, one may consider compress-and-forward (CF) relaying, in which the relay samples at a rate which can be incommensurate with the symbol rate of the incoming communications signal.

In this work, we employ the information-spectrum framework [14] for characterizing the RDF of asynchronously sampled memoryless Gaussian WSCS processes, as this framework is applicable to the information-theoretic analysis of non information-stable processes ([14] Page VII). We further note that while rate characterizations obtained using information spectrum tools and its associated quantities may be difficult to evaluate ([14] Remark 1.7.3), here we obtain a numerically computable characterization of the RDF. In particular, we focus on the mean squared error (MSE) distortion measure in the low distortion regime, namely, source codes for which the average square of the difference between the source and the reproduction process is not larger than the minimal source variance. The results of this research lead to accurate modelling of signal compression in current and future digital communications systems. The derived RDF, which characterizes the fundamental performance limits in encoding sampled CT WSCS Gaussian processes into a digital representation, allows to evaluate source coding schemes associated with different levels of complexity in terms of their gap from optimality, when applied to this important class of signals.

Furthermore, we utilize our characterization of the RDF to examine how the RDF for a sampled CT WSCS Gaussian source varies with different sampling rates and sampling time offsets. We demonstrate that, differently from stationary sources, when applying a lossy source code to a sampled WSCS process, the achievable rate-distortion tradeoff can be significantly affected by minor variations in the sampling time offset and the sampling rate. Our results thus allow identifying the sampling rate and sampling time offsets which minimize the RDF in systems involving sampled WSCS processes.

The rest of this work is organised as follows: Section 2 provides a scientific background on cyclostationary processes and on rate-distortion analysis of DT WSCS Gaussian sources. Section 3 presents the problem formulation and auxiliary results, and Section 4 details the main result of RDF characterization for sampled WSCS Gaussian process. Numerical examples and discussions are provided in Section 5, and Section 6 concludes the paper.

## 2. Preliminaries and Background

In the following we review the main tools and framework used in this work: In Section 2.1 we detail the notations, and in Section 2.2 we review the basics of cyclostationary processes and the statistical properties of a DT process resulting from sampling a CT WSCS process. In Section 2.3 we recall some preliminaries of rate-distortion theory as well as the RDF for DT WSCS Gaussian source processes. This background creates a premise for the statement of the main result provided in Section 4 of this paper.

### 2.1. Notations

In this paper, random vectors are denoted by boldface uppercase letters, e.g.,  $\mathbf{X}$ ; boldface lowercase letters denote deterministic column vectors, e.g.,  $\mathbf{x}$ . Scalar RVs and deterministic values are denoted via standard uppercase and lowercase fonts respectively, e.g.,  $X$  and  $x$ . Scalar random processes are denoted with  $X(t)$ ,  $t \in \mathcal{R}$  for CT and with  $X[n]$ ,  $n \in \mathcal{Z}$  for DT. Uppercase Sans-Serif fonts represent matrices, e.g.,  $A$ , and the element at the  $i^{\text{th}}$  row and the  $l^{\text{th}}$  column of  $A$  is denoted with  $(A)_{i,l}$ . We use  $|\cdot|$  to denote the absolute value,  $\lfloor d \rfloor$ ,  $d \in \mathcal{R}$ , to denote the floor function and  $d^+$ ,  $d \in \mathcal{R}$ , to denote the  $\max\{0, d\}$ .  $\delta[\cdot]$  denotes the Kronecker delta function:  $\delta[n] = 1$  for  $n = 0$  and  $\delta[n] = 0$  otherwise, and  $\mathbb{E}\{\cdot\}$  denotes the stochastic expectation. The sets of positive integers, integers, rational numbers, real numbers, positive numbers, and complex numbers are denoted by  $\mathcal{N}$ ,  $\mathcal{Z}$ ,  $\mathcal{Q}$ ,  $\mathcal{R}$ ,  $\mathcal{R}^{++}$ , and  $\mathcal{C}$ , respectively. The cumulative distribution function (CDF) is denoted by  $F_X(x) \triangleq \Pr(X \leq x)$  and the probability density function (PDF) of a CT random variable (RV) is denoted by  $p_X(x)$ . We represent a real Gaussian distribution with mean  $\mu$  and variance  $\sigma^2$  by the notation  $\mathcal{N}(\mu, \sigma^2)$ . All logarithms are taken to base-2, and  $j = \sqrt{-1}$ . Lastly, for any sequence  $y[i]$ ,  $i \in \mathcal{N}$ , and positive integer  $k \in \mathcal{N}$ ,  $\mathbf{y}^{(k)}$  denotes the column vector  $(y[1], \dots, y[k])^T$ .

## 2.2. Wide-Sense Cyclostationary Random Processes

Here, we review some preliminaries from the theory of cyclostationarity. We begin by recalling the definition of wide-sense cyclostationary processes for CT and for DT:

**Definition 1** (CT wide-sense cyclostationary processes ([3] Section 3.2.1)). *A scalar stochastic process  $\{S(t)\}_{t \in \mathcal{R}}$  is called WSCS if both its first-order and its second-order moments are periodic with respect to  $t \in \mathcal{R}$  with some period  $T_p \in \mathcal{R}$ .*

**Definition 2** (DT wide-sense cyclostationary processes ([2] Section 17.2)). *A scalar stochastic process  $\{S[n]\}_{n \in \mathcal{Z}}$  is called WSCS if both its first-order and its second-order moments are periodic with respect to  $n \in \mathcal{Z}$  with some period  $N_p \in \mathcal{Z}$ .*

WSCS signal are thus random processes whose first and second-order moments are periodic functions with the same period. To define WSACS signals, we first recall the definition of almost-periodic functions:

**Definition 3** (Almost-periodic functions ([19] Definition 2.1)). *A DT function  $x[n]$ ,  $n \in \mathcal{Z}$ , is called an almost-periodic function if for every  $\epsilon > 0$  there exists a number  $l(\epsilon) \in \mathcal{N}$  with the property that for any  $n \in \mathcal{Z}$  and any  $\alpha \in \mathcal{Z}$ ,  $\exists \Delta \in [\alpha, \alpha + l(\epsilon)]$ , such that*

$$|x[\Delta] - x[n]| < \epsilon.$$

**Definition 4** (DT wide-sense almost-cyclostationary processes ([2] Section 17.2)). *A scalar stochastic process  $\{S[n]\}_{n \in \mathcal{Z}}$  is called WSACS if its first and its second order moments are almost-periodic functions with respect to  $n \in \mathcal{Z}$ .*

**Remark 1.** *Note that when the mean and the autocorrelation function are each periodic with periods which are incommensurate, the resulting processes is WSACS. We note that in many practical cases the mean is zero, see e.g., ([2] Section 17.2), hence the classification of the process will be determined by the periodicity of the autocorrelation function.*

The DT WSCS model is commonly used in the communications literature, as it facilitates the the analysis of many problems of interest, such as fundamental rate limits analysis [20–22], channel identification [23], synchronization [24], and noise mitigation [25]. However, in many scenarios, the considered signals are WSACS rather than WSCS. To see how the WSACS model is obtained in the context of sampled signals, we briefly recall the discussion in [17] on sampled WSCS processes (please refer to ([17] Section II.B) for more details): Consider a CT WSCS random process  $S(t)$ , which is sampled uniformly with a sampling interval of  $T_s$  and sampling time offset  $\phi$ , resulting in a DT random process  $S[i] = S(i \cdot T_s + \phi)$ . It is well known that contrary to stationary processes, which have a time-invariant statistical characteristics, the values of  $T_s$  and  $\phi$  have a significant effect on the statistics of sampled WSCS processes ([17] Section II.B). To demonstrate this point, consider a CT WSCS process with variance  $\sigma_s^2(t) = \frac{1}{2} \cdot \sin(2\pi t / T_{\text{sym}}) + 2$  for some  $T_{\text{sym}} > 0$ . The sampled process for  $\phi = 0$  (no sampling time offset) and  $T_s = \frac{T_{\text{sym}}}{3}$  has a variance function whose period is  $N_p = 3$ :  $\sigma_s^2(iT_s) = \{2, 2.433, 1.567, 2, 2.433, 1.567, \dots\}$ , for  $i = 0, 1, 2, 3, 4, 5, \dots$ ; while the DT process obtained with the same sampling interval and with a sampling time offset of  $\phi = \frac{T_s}{2\pi}$  has a periodic variance with  $N_p = 3$  with values  $\sigma_s^2(iT_s + \phi) = \{2.155, 2.335, 1.510, 2.155, 2.335, 1.510, \dots\}$ , for  $i = 0, 1, 2, 3, 4, 5, \dots$ , which are different from the values of the DT variance for  $\phi = 0$ . It follows that both variances are periodic in discrete-time with the same period  $N_p = 3$ , although with different values within the period, which is a result of the sampling time offset, yet, both DT processes correspond to two instances

of synchronous sampling. Lastly, consider the sampled variance obtained by sampling without a time offset (i.e.,  $\phi = 0$ ) at a sampling interval of  $T_s = (1 + \frac{1}{2\pi}) \frac{T_{\text{sym}}}{3}$ . For this case,  $T_s$  is not an integer divisor of  $T_{\text{sym}}$  or of any of its integer multiples (i.e.,  $\frac{T_{\text{sym}}}{T_s} = 2 + \frac{2\pi-2}{2\pi+1} \equiv 2 + \epsilon$ ; where  $\epsilon \notin \mathbb{Q}$  and  $\epsilon \in [0, 1)$ ) resulting in the variance values  $\sigma_s^2(iT_s) = \{2, 2.335, 1.5027, 2.405, 1.896, 1.75, \dots\}$ , for  $i = 0, 1, 2, 3, 4, 5 \dots$ . For this scenario, the DT variance is not periodic but is almost-periodic, corresponding to asynchronous sampling and the resulting DT process is not WSCS but WSACS ([3] Section 3.2). The example above demonstrates that the statistical properties of sampled WSCS processes depend on the sampling rate and the sampling time offset, implying that the RDF of such processes should also depend on these quantities, as we demonstrate in the sequel.

### 2.3. The Rate-Distortion Function for DT WSCS Processes

In this subsection we review the source coding problem and the existing results on the RDF of WSCS processes. We begin by recalling the definition of a source coding scheme, see, e.g., ([26] Chapter 30), ([5] Chapter 10):

**Definition 5** (Source coding scheme). *A source coding scheme with blocklength  $l$  consists of (see Figure 1):*

1. *An encoder  $f_S$  which maps a block of  $l$  source samples  $\{S[i]\}_{i=1}^l$  into an index from a set of  $M = 2^{lR}$  indexes,  $f_S : \{S[i]\}_{i=1}^l \mapsto \{1, 2, \dots, M\}$ .*
2. *A decoder  $g_S$  which maps the received index into a reconstructed sequence of length  $l$ ,  $\{\hat{S}[i]\}_{i=1}^l$ ,  $g_S : \{1, 2, \dots, M\} \mapsto \{\hat{S}[i]\}_{i=1}^l$*

The encoder-decoder pair is referred to as an  $(R, l)$  source code, where  $R$  is the rate of the code in bits per source symbol, defined as:

$$R = \frac{1}{l} \log_2 M \tag{1}$$

The RDF characterizes the minimal average number of bits per source symbol, denoted  $R(D)$ , that can be used to encode a source process such that it can be reconstructed from its encoded representation with a recovery distortion not larger than  $D > 0$  ([5] Section 10.2). In the current work, we use the MSE distortion measure, which measures the distortion due to decoding a source symbol  $S$  into  $\hat{S}$  via  $d(S, \hat{S}) = (S - \hat{S})^2$ . The distortion for a sequence of source samples  $\mathbf{S}^{(l)}$  decoded into a reproduction sequence  $\hat{\mathbf{S}}^{(l)}$  is given by  $d(\mathbf{S}^{(l)}, \hat{\mathbf{S}}^{(l)}) = \frac{1}{l} \sum_{i=1}^l (S[i] - \hat{S}[i])^2$  and the average distortion in decoding a random source sequence  $\mathbf{S}^{(l)}$  into a random reproduction sequence  $\hat{\mathbf{S}}^{(l)}$  is defined as:

$$\bar{d}(\mathbf{S}^{(l)}, \hat{\mathbf{S}}^{(l)}) \triangleq \mathbb{E} \left\{ d(\mathbf{S}^{(l)}, \hat{\mathbf{S}}^{(l)}) \right\} = \frac{1}{l} \sum_{i=1}^l \mathbb{E} \left\{ (S[i] - \hat{S}[i])^2 \right\}, \tag{2}$$

where the expectation in Equation (2) is taken with respect to the joint probability distributions on the source  $S[i]$  and its reproduction  $\hat{S}[i]$ . Using Definition 5 we can now define the achievable rate-distortion pair for a source  $S[i]$ , as stated in the following definition ([10] Pg. 471):

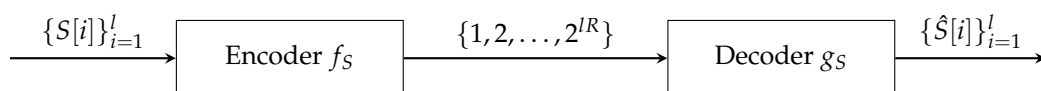


Figure 1. Source coding block diagram.

**Definition 6** (Achievable rate-distortion pair). *A rate-distortion pair  $(R, D)$  is achievable for a process  $\{S[i]\}_{i \in \mathcal{N}}$  if for any  $\eta > 0$  and for all sufficiently large  $l$  it is possible to construct an  $(R_s, l)$  source code such that*

$$R_s \leq R + \eta. \tag{3}$$

and

$$\bar{d}(\mathbf{S}^{(l)}, \hat{\mathbf{S}}^{(l)}) \leq D + \eta. \tag{4}$$

**Definition 7.** The rate-distortion function  $R(D)$  is defined as the infimum of all achievable rates  $R$  for a given maximum allowed distortion  $D$ .

Definition 6 defines a rate-distortion pair to be achievable if the rate and the distortion constraints are satisfied using source codes with any sufficiently large blocklength. In the following lemma, which will be used to characterize the RDF of DT WSCS signals, we state that it is sufficient to consider only source codes whose blocklength is an integer multiple of some fixed positive integer:

**Lemma 1.** Consider the process  $\{S[i]\}_{i \in \mathcal{N}}$  with a finite and bounded variance. For a given maximum allowed distortion  $D$ , the optimal reproduction process  $\{\hat{S}[i]\}_{i \in \mathcal{N}}$  is also the optimal reproduction process when restricted to using source codes whose blocklengths are integer multiples of some fixed positive integer  $r$ .

**Proof.** The proof of the lemma is detailed in Appendix A.  $\square$

This lemma facilitates switching between multivariate and scalar representations of the source and the reproduction processes.

The RDF obviously depends on the distribution of the source  $\{S[i]\}_{i \in \mathcal{N}}$ . Thus, statistically different sources have different RDFs. However, when a source is scaled by some positive constant, the RDF of the scaled process with the MSE distortion can be inferred from that of the original source process, as stated in the following theorem:

**Theorem 1.** Let  $\{S[i]\}_{i \in \mathcal{N}}$  be a source process for which the rate-distortion pair  $(R, D)$  is achievable under the MSE distortion. Then, for every  $\alpha \in \mathcal{R}^{++}$ , it holds that the rate-distortion pair  $(R, \alpha^2 \cdot D)$  is achievable for the scaled source  $\{\alpha \cdot S[i]\}_{i \in \mathcal{N}}$ .

**Proof.** The proof of the theorem is detailed in Appendix B.  $\square$

Lastly, in the proof of our main result, we make use of the RDF for DT WSCS sources derived in ([12] Theorem 1), repeated below for ease of reference. Prior to the statement of the theorem, we recall that for blocklengths which are integer multiples of  $N_p$ , a WSCS process  $S[i]$  with period  $N_p > 0$  can be represented as an equivalent  $N_p$ -dimensional process  $\mathbf{S}^{(N_p)}[i]$  via the decimated component decomposition (DCD) ([2] Section 17.2). The power spectral density (PSD) of the process  $\mathbf{S}^{(N_p)}$  is defined as ([12] Section II):

$$\left( \rho_{\mathbf{S}} \left( e^{j2\pi f} \right) \right)_{u,v} = \sum_{\Delta \in \mathcal{Z}} \left( R_{\mathbf{S}}[\Delta] \right)_{u,v} e^{-j2\pi f \Delta} \quad -\frac{1}{2} \leq f \leq \frac{1}{2}, \quad u, v \in \{1, 2, \dots, N_p\} \tag{5}$$

where  $R_{\mathbf{S}}[\Delta] \triangleq \mathbb{E} \left\{ \mathbf{S}^{(N_p)}[i] \cdot \mathbf{S}^{(N_p)}[i + \Delta] \right\}$  ([2] Section 17.2). We now proceed to the statement of ([12] Theorem 1):

**Theorem 2.** ([12] Theorem 1) Consider a zero-mean real DT WSCS Gaussian source  $S[i], i \in \mathcal{N}$  with memory, and let  $N_p \in \mathcal{N}$  denote the period of its statistics. The RDF is expressed as:

$$R(D) = \frac{1}{2N_p} \sum_{m=1}^{N_p} \int_{f=-0.5}^{0.5} \left( \log \left( \frac{\lambda_m \left( e^{j2\pi f} \right)}{\theta} \right) \right)^+ df, \tag{6a}$$

where  $\lambda_m \left( e^{j2\pi f} \right)$ ,  $m = 1, 2, \dots, N_p$  denote the eigenvalues of the PSD matrix of the process  $\mathbf{S}^{(N_p)}[i]$ , which is obtained from  $S[i]$  by applying the  $N_p$ -dimensional DCD, and  $\theta$  is selected such that

$$D = \frac{1}{N_p} \sum_{m=1}^{N_p} \int_{f=-0.5}^{0.5} \min \left\{ \lambda_m \left( e^{j2\pi f} \right), \theta \right\} df. \tag{6b}$$

We note that  $\mathbf{S}^{(N_p)}[i]$  corresponds to a vector of stationary processes whose elements are not identically distributed; hence the variance function is different for each scalar stationary process. Using ([12] Theorem 1), we can directly obtain the RDF for the special case of a DT memoryless WSCS Gaussian process. This is stated in the following corollary:

**Corollary 1.** Let  $\{S[i]\}_{i \in \mathcal{N}}$  be a zero-mean DT memoryless real WSCS Gaussian source with period  $N_p \in \mathcal{N}$ , and set  $\sigma_m^2 = \mathbb{E}\{S^2[m]\}$  for  $m = 1, 2, \dots, N_p$ . The RDF for compression of  $S[i]$  is expressed as:

$$R(D) = \begin{cases} \frac{1}{2N_p} \sum_{m=1}^{N_p} \log \left( \frac{\sigma_m^2}{D_m} \right) & D \leq \frac{1}{N_p} \sum_{m=1}^{N_p} \sigma_m^2 \\ 0 & D > \frac{1}{N_p} \sum_{m=1}^{N_p} \sigma_m^2 \end{cases} \tag{7a}$$

where  $D_m \triangleq \min \{ \sigma_m^2, \theta \}$ , and  $\theta$  is defined such that

$$D = \frac{1}{N_p} \sum_{m=1}^{N_p} D_m. \tag{7b}$$

**Proof.** Applying Equations (6a) and (6b) to our specific case of a memoryless WSCS source, we obtain Equations (7a) and (7b) as follows: First, note that the corresponding DCD components for a zero-mean memoryless WSCS process are also zero-mean and memoryless; hence the PSD matrix for the multivariate process  $\mathbf{S}^{(N_p)}[i]$  is a diagonal matrix, whose eigenvalues are the constant diagonal elements such that the  $m$ th diagonal element is equal to the variance  $\sigma_m^2$ :  $\lambda_m \left( e^{j2\pi f} \right) = \sigma_m^2$ . Now, writing Equation (6a) for this case we obtain:

$$\begin{aligned} R(D) &= \frac{1}{2N_p} \sum_{m=1}^{N_p} \int_{f=-0.5}^{0.5} \left( \log \left( \frac{\lambda_m \left( e^{j2\pi f} \right)}{\theta} \right) \right)^+ df \\ &= \frac{1}{2N_p} \sum_{m=1}^{N_p} \left( \log \left( \frac{\sigma_m^2}{\theta} \right) \right)^+. \end{aligned} \tag{8}$$

Since  $\left( \log \left( \frac{\sigma_m^2}{\theta} \right) \right)^+ = \max \left\{ 0, \log \left( \frac{\sigma_m^2}{\theta} \right) \right\} \equiv \log \left( \frac{\sigma_m^2}{D_m} \right)$  it follows that (8) coincides with (7a). Next, expressing Equation (6b) for the memoryless source process, we obtain:

$$D = \frac{1}{N_p} \sum_{m=1}^{N_p} \int_{f=-0.5}^{0.5} \min \left\{ \lambda_m \left( e^{j2\pi f} \right), \theta \right\} df = \frac{1}{N_p} \sum_{m=1}^{N_p} \min \left\{ \sigma_m^2, \theta \right\}, \tag{9}$$

proving Equation (7b).  $\square$

Now, from Lemma 1, we conclude that the RDF for compression of source sequences whose blocklength is an integer multiple of  $N_p$  is the same as the RDF for compressing source sequences whose blocklength is arbitrary. We recall that from ([5] Chapter 10.3.3) it follows that for the zero-mean memoryless Gaussian DCD vector source process  $\mathbf{S}^{(N_p)}[i]$  the optimal reproduction process which achieves the RDF is an  $N_p \times 1$  memoryless process whose covariance matrix is diagonal with

non-identically distributed elements. From [2], we can apply the inverse DCD to obtain a WSCS process. Hence, from Lemma 1 we can conclude that the optimal reproduction process for the DT WSCS Gaussian source is a DT WSCS Gaussian process.

### 3. Problem Formulation and Auxiliary Results

Our objective is to characterize the RDF for compression of asynchronously sampled CT WSCS Gaussian sources when the sampling interval is larger than the memory of the source. In particular, we focus on the minimal rate required to achieve a high fidelity reproduction, representing the RDF curve for distortion values not larger than the variance of the source. Such characterization of the RDF for asynchronous sampling is essential for comprehending the relationship between the minimal required number of bits and the sampling rate at a given distortion. Our analysis constitutes an important step towards constructing joint source-channel coding schemes for scenarios in which the symbol rate of the transmitter is not necessarily synchronized with the sampling rate of the source to be transmitted. Such scenarios arise, for example, when recording a communications signal for storage or processing, or in compress-and-forward relaying ([26] Chapter 16.7), [27]) in which the relay compresses the sampled received signal, which is then forwarded to the assisted receiver. As the relay operates with its own sampling clock, which need not necessarily be synchronized with the symbol rate of the assisted transmitter, sampling at the relay may result in a DT WSACS source signal. In the following we first characterize the sampled source model in Section 3.1. Then, as a preliminary step towards our characterization the RDF for asynchronously sampled CT WSCS Gaussian processes stated in Section 4, we recall in Section 3.2 the definitions of some information-spectrum quantities used in this study. Finally, in Section 3.3, we recall an auxiliary result relating the information-spectrum quantities of a collection of sequences of RVs to the information-spectrum quantities of its limit sequence of RVs. This result will be applied in the derivation of the RDF with asynchronous sampling.

#### 3.1. Source Model

Consider a real CT, zero-mean WSCS Gaussian random process  $S_c(t)$  with period  $T_{ps}$ . Let the variance function of  $S_c(t)$  be defined as  $\sigma_{S_c}^2(t) \triangleq \mathbb{E}\{S_c^2(t)\}$ , and assume it is both upper bounded and lower bounded away from zero, and that it is continuous in  $t \in \mathcal{R}$ . Let  $\tau_m > 0$  denote the maximal correlation length of  $S_c(t)$ , i.e.,  $r_{S_c}(t, \tau) \triangleq \mathbb{E}\{S_c(t)S_c(t - \tau)\} = 0, \forall |\tau| > \tau_m$ . By the cyclostationarity of  $S_c(t)$ , we have that  $\sigma_{S_c}^2(t) = \sigma_{S_c}^2(t + T_{ps}), \forall t \in \mathcal{R}$ . Let  $S_c(t)$  be sampled uniformly with the sampling interval  $T_s > 0$  such that  $T_{ps} = (p + \epsilon) \cdot T_s$  for  $p \in \mathcal{N}$  and  $\epsilon \in [0, 1)$  yielding  $S_\epsilon[i] \triangleq S_c(i \cdot T_s)$ , where  $i \in \mathcal{Z}$ . The variance of  $S_\epsilon[i]$  is given by  $\sigma_{S_\epsilon}^2[i] \triangleq r_{S_\epsilon}[i, 0] = \sigma_{S_c}^2\left(\frac{i \cdot T_{ps}}{p + \epsilon}\right)$ .

In this work, as in [17], we assume that the duration of temporal correlation of the CT signal is shorter than the sampling interval  $T_s$ , namely,  $\tau_m < T_s$ . Consequently, the DT Gaussian process  $S_\epsilon[i]$  is a memoryless zero-mean Gaussian process and its autocorrelation function is given by:

$$\begin{aligned} r_{S_\epsilon}[i, \Delta] &= \mathbb{E}\left\{S_\epsilon[i]S_\epsilon[i + \Delta]\right\} \\ &= \mathbb{E}\left\{S_c\left(\frac{i \cdot T_{ps}}{p + \epsilon}\right) \cdot S_c\left(\frac{(i + \Delta) \cdot T_{ps}}{p + \epsilon}\right)\right\} = \sigma_{S_c}^2\left(\frac{i \cdot T_{ps}}{p + \epsilon}\right) \cdot \delta[\Delta] = \sigma_{S_\epsilon}^2[i] \cdot \delta[\Delta]. \end{aligned} \quad (10)$$

While we do not explicitly account for sampling time offsets in our definition of the sampled process  $S_\epsilon[i]$ , it can be incorporated by replacing  $\sigma_{S_c}^2(t)$  with a time-shifted version, i.e.,  $\sigma_{S_c}^2(t - \phi)$ , see also ([17] Section II.C).

It can be noted from (10) that if  $\epsilon$  is a rational number, i.e.,  $\exists u, v \in \mathcal{N}$ ,  $u$  and  $v$  are relatively prime, such that  $\epsilon = \frac{u}{v}$ , then  $\{S_\epsilon[i]\}_{i \in \mathcal{Z}}$  is a DT memoryless WSCS process with the period  $p_{u,v} = p \cdot v + u \in \mathcal{N}$  ([17] Section II.C). For this class of processes, the RDF can be obtained from ([12] Theorem 1) as stated in Corollary 1. On the other hand, if  $\epsilon$  is an irrational number, then sampling becomes asynchronous and leads to a WSACS process whose RDF has not been characterized to date.



### 3.2. Definitions of Relevant Information-Spectrum Quantities

Conventional information theoretic tools for characterizing RDFs are based on an underlying ergodicity of the source. Consequently, these techniques cannot be applied to characterize the RDF of asynchronously sampled WSCS processes. To tackle this challenge, we use the information-spectrum framework, as this framework [14] can be utilized to obtain general formulas for rate limits for any arbitrary class of processes. The resulting expressions are not restricted to specific statistical models of the considered processes, and in particular, do not require information stability or stationarity. In the following, we recall the definitions of several information-spectrum quantities used in this study, see also ([14] Definitions 1.3.1 and 1.3.2):

**Definition 8.** *The limit-inferior in probability of a sequence of real RVs  $\{Z_k\}_{k \in \mathcal{N}}$  is defined as*

$$p\text{-}\liminf_{k \rightarrow \infty} Z_k \triangleq \sup \left\{ \alpha \in \mathcal{R} \mid \lim_{k \rightarrow \infty} \Pr(Z_k < \alpha) = 0 \right\} \triangleq \alpha_0. \tag{11}$$

Hence,  $\alpha_0$  is the largest real number satisfying that  $\forall \tilde{\alpha} < \alpha_0$  and  $\forall \mu > 0$  there exists  $k_0(\mu, \tilde{\alpha}) \in \mathcal{N}$  such that  $\Pr(Z_k < \tilde{\alpha}) < \mu, \forall k > k_0(\mu, \tilde{\alpha})$ .

**Definition 9.** *The limit-superior in probability of a sequence of real RVs  $\{Z_k\}_{k \in \mathcal{N}}$  is defined as*

$$p\text{-}\limsup_{k \rightarrow \infty} Z_k \triangleq \inf \left\{ \beta \in \mathcal{R} \mid \lim_{k \rightarrow \infty} \Pr(Z_k > \beta) = 0 \right\} \triangleq \beta_0. \tag{12}$$

Hence,  $\beta_0$  is the smallest real number satisfying that  $\forall \tilde{\beta} > \beta_0$  and  $\forall \mu > 0$ , there exists  $k_0(\mu, \tilde{\beta}) \in \mathcal{N}$ , such that  $\Pr(Z_k > \tilde{\beta}) < \mu, \forall k > k_0(\mu, \tilde{\beta})$ .

The notion of uniform integrability of a sequence of RVs is a basic property in probability ([28] Chapter 12), which is not directly related to information spectrum methods. However, since it plays an important role in the information spectrum characterization of RDFs, we include its statement in the following definition:

**Definition 10** (Uniform integrability ([28] Definition 12.1), ([14] Equation (5.3.2))). *The sequence of real-valued random variables  $\{Z_k\}_{k=1}^\infty$ , is said to satisfy uniform integrability if*

$$\lim_{u \rightarrow \infty} \sup_{k \geq 1} \int_{z: |z| \geq u} p_{Z_k}(z) |z| dz = 0 \tag{13}$$

The aforementioned quantities facilitate characterizing the RDF of arbitrary sources. Consider a general source process  $\{S[i]\}_{i=1}^\infty$  (stationary or non-stationary) taking values from the source alphabet  $S[i] \in \mathcal{S}$  and a reproduction process  $\{\hat{S}[i]\}_{i=1}^\infty$  with values from the reproduction alphabet  $\hat{S}[i] \in \hat{\mathcal{S}}$ . It follows from ([14] Section 5.5) that for a distortion measure which satisfies the uniform integrability criterion, i.e., that there exists a deterministic sequence  $\{r[i]\}_{i=1}^\infty$  such that the sequence of RVs  $\{d(S^{(k)}, r^{(k)})\}_{k=1}^\infty$  satisfies Definition 10 ([14] Page 336), then the RDF is expressed as ([14] Equation (5.4.2)):

$$R(D) = \inf_{F_{\mathcal{S}, \hat{\mathcal{S}}}: \bar{d}_{\mathcal{S}}(S^{(k)}, \hat{S}^{(k)}) \leq D} \bar{I}(S^{(k)}; \hat{S}^{(k)}), \tag{14}$$

where  $\bar{d}_S(\mathbf{S}^{(k)}, \hat{\mathbf{S}}^{(k)}) = \limsup_{k \rightarrow \infty} \mathbb{E} \left\{ d \left( \mathbf{S}^{(k)}, \hat{\mathbf{S}}^{(k)} \right) \right\}$ ,  $F_{S, \hat{S}}$  denotes the joint CDF of  $\{S[i]\}_{i=1}^\infty$  and  $\{\hat{S}[i]\}_{i=1}^\infty$ , and  $\bar{I} \left( \mathbf{S}^{(k)} : \hat{\mathbf{S}}^{(k)} \right)$  represents the limit superior in probability of the mutual information rate of  $\mathbf{S}^{(k)}$  and  $\hat{\mathbf{S}}^{(k)}$ , given by:

$$\bar{I} \left( \mathbf{S}^{(k)} ; \hat{\mathbf{S}}^{(k)} \right) \triangleq \text{p-} \limsup_{k \rightarrow \infty} \frac{1}{k} \log \frac{p_{S^{(k)} | \hat{S}^{(k)}} \left( \mathbf{S}^{(k)} | \hat{\mathbf{S}}^{(k)} \right)}{p_{S^{(k)}} \left( \mathbf{S}^{(k)} \right)} \tag{15}$$

In order to use the RDF characterization in (14), the distortion measure must satisfy the uniform integrability criterion. For the considered class of sources detailed in Section 3.1, the MSE distortion satisfies this criterion, as stated in the following lemma:

**Lemma 2.** For any real memoryless zero-mean Gaussian source  $\{S[i]\}_{i=1}^\infty$  with bounded variance, i.e.,  $\exists \sigma_{\max}^2 < \infty$  such that  $\mathbb{E}\{S^2[i]\} \leq \sigma_{\max}^2$  for all  $i \in \mathcal{N}$ , the MSE distortion satisfies the uniform integrability criterion.

**Proof.** Set the deterministic sequence  $\{r[i]\}_{i=1}^\infty$  to be the all-zero sequence. Under this setting and the MSE distortion, it holds that  $d(\mathbf{S}^{(k)}, \mathbf{r}^{(k)}) = \frac{1}{k} \sum_{i=1}^k S^2[i]$ . To prove the lemma, we show that the sequence of RVs  $\{d(\mathbf{S}^{(k)}, \mathbf{r}^{(k)})\}_{k=1}^\infty$  has a bounded  $\ell_2$  norm, which implies that it is uniformly integrable by ([28] Corollary 12.8). The  $\ell_2$  norm of  $d(\mathbf{S}^{(k)}, \mathbf{r}^{(k)})$  satisfies

$$\begin{aligned} \mathbb{E}\{d(\mathbf{S}^{(k)}, \mathbf{r}^{(k)})^2\} &= \frac{1}{k^2} \mathbb{E} \left\{ \sum_{i=1}^k S^2[i] \sum_{j=1}^k S^2[j] \right\} \\ &= \frac{1}{k^2} \sum_{i=1}^k \sum_{j=1}^k \mathbb{E} \left\{ S^2[i] S^2[j] \right\} \stackrel{(a)}{\leq} \frac{1}{k^2} \sum_{i=1}^k \sum_{j=1}^k 3\sigma_{\max}^4 = 3\sigma_{\max}^4, \end{aligned} \tag{16}$$

where (a) follows since  $\mathbb{E}\{S^2[i]S^2[j]\} = \mathbb{E}\{S^2[i]\}\mathbb{E}\{S^2[j]\} = \sigma_{\max}^4$  for  $i \neq j$  while  $\mathbb{E}\{S^4[i]\} = 3\sigma_{\max}^4$  ([29] Chapter 5.4). Equation (16) proves that  $d(\mathbf{S}^{(k)}, \mathbf{r}^{(k)})$  is  $\ell_2$ -bounded by  $3\sigma_{\max}^4 < \infty$  for all  $k \in \mathcal{N}$ , which in turn implies that the MSE distortion is uniformly integrable for the source  $\{S[i]\}_{i=1}^\infty$ .  $\square$

Since, as detailed in Section 3.1, we focus in the following on memoryless zero-mean Gaussian sources, Lemma 2 implies that the RDF of the source can be characterized using (14). However, (14) is in general difficult to evaluate, and thus does not lead to a meaningful understanding of how the RDF of sampled WSCS sources behaves, motivating our analysis in Section 4.

### 3.3. Information Spectrum Limits

The following theorem originally stated in ([17] Theorem 1) presents a fundamental result which is directly useful for the derivation of the RDF:

**Theorem 3.** ([17] Theorem 1) Let  $\{\tilde{Z}_{k,n}\}_{n,k \in \mathcal{N}}$  be a set of sequences of real scalar RVs satisfying two assumptions:

- AS1 For every fixed  $n \in \mathcal{N}$ , every convergent subsequence of  $\{\tilde{Z}_{k,n}\}_{k \in \mathcal{N}}$  converges in distribution, as  $k \rightarrow \infty$ , to a finite deterministic scalar. Each subsequence may converge to a different scalar.
- AS2 For every fixed  $k \in \mathcal{N}$ , the sequence  $\{\tilde{Z}_{k,n}\}_{n \in \mathcal{N}}$  converges uniformly in distribution, as  $n \rightarrow \infty$ , to a scalar real-valued RV  $Z_k$ . Specifically, letting  $\tilde{F}_{k,n}(\alpha)$  and  $F_k(\alpha)$ ,  $\alpha \in \mathcal{R}$ , denote the CDFs of  $\tilde{Z}_{k,n}$  and of  $Z_k$ , respectively, then by AS2 it follows that  $\forall \eta > 0$ , there exists  $n_0(\eta)$  such that for every  $n > n_0(\eta)$

$$|\tilde{F}_{k,n}(\alpha) - F_k(\alpha)| < \eta,$$

for each  $\alpha \in \mathcal{R}$ ,  $k \in \mathcal{N}$ .

Then, for  $\{\tilde{Z}_{k,n}\}_{n,k \in \mathcal{N}}$  it holds that

$$p\text{-}\liminf_{k \rightarrow \infty} Z_k = \lim_{n \rightarrow \infty} \left( p\text{-}\liminf_{k \rightarrow \infty} \tilde{Z}_{k,n} \right), \tag{17a}$$

$$p\text{-}\limsup_{k \rightarrow \infty} Z_k = \lim_{n \rightarrow \infty} \left( p\text{-}\limsup_{k \rightarrow \infty} \tilde{Z}_{k,n} \right). \tag{17b}$$

**Proof.** In Appendix C we explicitly prove Equation (17b). This complements the proof in ([17] Appendix A) which explicitly considers only (17a).  $\square$

#### 4. Rate-Distortion Characterization for Sampled CT WSCS Gaussian Sources

##### 4.1. Main Result

Using the information-spectrum based characterization of the RDF (14) combined with the characterization of the limit of a sequence of information spectrum quantities in Theorem 3, we now analyze the RDF of asynchronously sampled WSCS processes. Our analysis is based on constructing a sequence of synchronously sampled WSCS processes, whose RDF is given in Corollary 1. Then, we show that the RDF of the asynchronously sampled process can be obtained as the limit superior of the computable RDFs of the sequence of synchronously sampled processes. We begin by letting  $\epsilon_n \triangleq \frac{\lfloor n \cdot \epsilon \rfloor}{n}$  for  $n \in \mathcal{N}$  and defining a Gaussian source process  $S_n[i] = S_c\left(\frac{i \cdot T_{ps}}{p + \epsilon_n}\right)$ . From the discussion in Section 3.1 (see also ([17] Section II.C)), it follows that since  $\epsilon_n$  is rational,  $S_n[i]$  is a WSCS process and its period is given by  $p_n = p \cdot n + \lfloor n \cdot \epsilon \rfloor$ . Accordingly, the periodic correlation function of  $S_n[i]$  can be obtained similarly to (10) as:

$$r_{S_n}[i, \Delta] = \mathbb{E} \left\{ S_n[i] S_n[i + \Delta] \right\} = \sigma_{S_c}^2 \left( \frac{i \cdot T_{ps}}{p + \epsilon_n} \right) \cdot \delta[\Delta]. \tag{18}$$

Due to cyclostationarity of  $S_n[i]$ , we have that  $r_{S_n}[i, \Delta] = r_{S_n}[i + p_n, \Delta]$ ,  $\forall i, \Delta \in \mathcal{Z}$ , and we let  $\sigma_{S_n}^2[i] \triangleq r_{S_n}[i, 0]$  denote its periodic variance.

We next restate Corollary 1 in terms of  $\epsilon_n$  as follows:

**Proposition 1.** Consider a DT, memoryless, zero-mean, WSCS Gaussian random process  $S_n[i]$  with a variance  $\sigma_{S_n}^2[i]$ , obtained from  $S_c(t)$  by sampling with a sampling interval of  $T_s(n) = \frac{T_{ps}}{p + \epsilon_n}$ . Let  $S_n^{(p_n)}[i]$  denote the memoryless stationary multivariate random process obtained by applying the DCD to  $S_n[i]$  and let  $\sigma_{S_n}^2[m]$ ,  $m = 1, 2, \dots, p_n$ , denote the variance of the  $m^{\text{th}}$  component of  $S_n^{(p_n)}[i]$ . The rate-distortion function is given by:

$$R_n(D) = \begin{cases} \frac{1}{2^{p_n}} \sum_{m=1}^{p_n} \left( \log \left( \frac{\sigma_{S_n}^2[m]}{D_n[m]} \right) \right) & D \leq \frac{1}{p_n} \sum_{m=1}^{p_n} \sigma_{S_n}^2[m] \\ 0 & D > \frac{1}{p_n} \sum_{m=1}^{p_n} \sigma_{S_n}^2[m] \end{cases}, \tag{19a}$$

where for  $D \leq \frac{1}{p_n} \sum_{m=1}^{p_n} \sigma_{S_n}^2[m]$  we let  $D_n[m] \triangleq \min \{ \sigma_{S_n}^2[m], \theta_n \}$ , and  $\theta_n$  is selected such that

$$D = \frac{1}{p_n} \sum_{m=1}^{p_n} D_n[m]. \tag{19b}$$

We recall that the RDF of  $S_n[i]$  is characterized in Proposition 1 via the RDF of the multivariate stationary process  $S_n^{(p_n)}[i]$  obtained via a  $p_n$ -dimensional DCD applied to  $S_n[i]$ . Next, we recall that the relationship between the source process  $S_n^{(p_n)}[i]$  and the optimal reconstruction process, denoted

by  $\hat{S}_n^{(p_n)}[i]$ , is characterized in ([5] Chapter 10.3.3) via a linear, multivariate, time-invariant backward channel with a  $p_n \times 1$  additive vector noise process  $W_n^{(p_n)}[i]$ , and is given by:

$$S_n^{(p_n)}[i] = \hat{S}_n^{(p_n)}[i] + W_n^{(p_n)}[i], \quad i \in \mathcal{N}. \tag{20}$$

It also follows from ([5] Section 10.3.3) that for the IID Gaussian multivariate process whose entries are independent and distributed via  $(S_n^{(p_n)}[i])_m \sim \mathcal{N}(0, \sigma_{S_n}^2[m])$ ,  $m \in \{1, 2, \dots, p_n\}$ , the optimal reconstruction vector process  $\hat{S}_n^{(p_n)}[i]$  and the corresponding noise vector process  $W_n^{(p_n)}[i]$  each follow a multivariate Gaussian distribution:

$$\hat{S}_n^{(p_n)}[i] \sim \mathcal{N} \left( \mathbf{0}, \begin{bmatrix} \sigma_{\hat{S}_n}^2[1] \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_{\hat{S}_n}^2[p_n] \end{bmatrix} \right) \quad \text{and} \quad W_n^{(p_n)}[i] \sim \mathcal{N} \left( \mathbf{0}, \begin{bmatrix} D_n[1] \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & D_n[p_n] \end{bmatrix} \right),$$

where  $D_n[m] \triangleq \min \{ \sigma_{S_n}^2[m], \theta_n \}$ ;  $\theta_n$  denotes the reverse waterfilling threshold defined in Prop. 1 for the index  $n$ , and is selected such that  $D = \frac{1}{p_n} \sum_{m=1}^{p_n} D_n[m]$ . The optimal reconstruction process,  $\hat{S}_n^{(p_n)}[i]$  and the noise process  $W_n^{(p_n)}[i]$  are mutually independent, and for each  $m \in \{1, 2, \dots, p_n\}$  it holds that  $\mathbb{E} \left\{ \left( S_n^{(p_n)}[i] - \hat{S}_n^{(p_n)}[i] \right)_m^2 \right\} = D_n[m]$ , see ([5] Chapters 10.3.2 and 10.3.3). The multivariate relationship between stationary processes in (20) can be transformed into an equivalent linear relationship between cyclostationary Gaussian memoryless processes via the inverse DCD transformation ([2] Sec 17.2) applied to each of the processes, resulting in:

$$S_n[i] = \hat{S}_n[i] + W_n[i], \quad i \in \mathcal{N}. \tag{21}$$

We are now ready to state our main result, which is the RDF of asynchronously sampled DT sources  $S_\epsilon[i], \epsilon \notin \mathcal{Q}$ , in the low MSE regime, i.e., when the distortion  $D$  is not larger than the source variance. The RDF is stated in the following theorem, which applies to both synchronous sampling as well as to asynchronous sampling:

**Theorem 4.** Consider a DT source  $\{S_\epsilon[i]\}_{i=1}^\infty$  obtained by sampling a CT WSCS source, whose period of statistics is  $T_{ps}$ , at intervals  $T_s$ . Then, for any distortion constraint  $D$  such that  $D < \min_{0 \leq t \leq T_{ps}} \sigma_{S_\epsilon}^2(t)$  and any  $\epsilon \in [0, 1)$ , the RDF  $R_\epsilon(D)$  for compressing  $\{S_\epsilon[i]\}_{i=1}^\infty$  can be obtained as the limit:

$$R_\epsilon(D) = \limsup_{n \rightarrow \infty} R_n(D), \tag{22}$$

where  $R_n(D)$  is defined Prop. 1.

**Proof.** The detailed proof is provided in Appendix D. Here, we give a brief outline: The derivation of the RDF with asynchronous sampling follows three steps: First, we note that sampling at a rate of  $T_s(n) = \frac{T_{ps}}{p+\epsilon_n}$  results in a sequence of DT WSCS sources  $\{S_n[i]\}_{i \in \mathcal{N}, n \in \mathcal{N}}$  whose sampling interval  $T_s(n)$  asymptotically approaches, as  $n \rightarrow \infty$ , the sampling interval for irrational  $\epsilon$  given by  $T_s = \frac{T_{ps}}{p+\epsilon}$ . We define a sequence of rational numbers  $\epsilon_n$  s.t.  $\epsilon_n \rightarrow \epsilon$  as  $n \rightarrow \infty$ ; Building upon this insight, we prove that the RDF with  $T_s$  can be stated as a double limit where the outer limit is with respect to the blocklength and the inner limit is with respect to  $\epsilon_n$ . Lastly, we use Theorem 3 to show that the order of the limits can be exchanged, obtaining a limit of expressions which are computable.  $\square$

**Remark 2.** Theorem 4 focuses on the low distortion regime, defined as the values of  $D$  satisfying  $D < \min_{0 \leq t \leq T_{ps}} \sigma_{S_c}^2(t)$ . This implies that  $\theta_n$  has to be smaller than  $\min_{0 \leq t \leq T_{ps}} \sigma_{S_c}^2(t)$ ; hence, from Prop. 1 it follows that for the corresponding stationary noise vector  $\mathbf{W}_n^{(p_n)}[i]$  in (20),  $D_n[m] = \min \{ \sigma_{S_n}^2[m], \theta_n \} = \theta_n$  and  $D = \frac{1}{p_n} \sum_{m=1}^{p_n} D_n[m] = \theta_n = D_n[m]$ . We note that since every element of the vector  $(\mathbf{W}_n^{(p_n)}[i])_m$  has the same variance  $D_n[m] = D$  for all  $n \in \mathcal{N}$  and  $m = 1, 2, \dots, p_n$  then by applying the inverse DCD to  $\mathbf{W}_n^{(p_n)}[i]$ , the resulting scalar DT process  $W_n[i]$  is wide sense stationary; and in fact IID with  $\mathbb{E} \{ (W_n[i])^2 \} = D$ .

#### 4.2. Discussion and Relationship with Capacity Derivation in Reference 17

Theorem 4 provides a meaningful and computable characterization for the RDF of sampled WSCS signals. We note that the proof of the main theorem uses some of the steps used in our recent study on the capacity of memoryless channels with sampled CT WSCS Gaussian noise [17]. It should be emphasized, however, that there are several fundamental differences between the two studies, which require the introduction of new treatments and derivations original to the current work. First, it is important to note that in the study on capacity, a physical channel model exists, and therefore the conditional PDF of the output signal given the input signal can be characterized explicitly for both synchronous sampling and asynchronous sampling for every input distribution. For the current study of the RDF we note that the relationship (21), commonly referred to as the backward channel [30], ([5] Chapter 10.3.2), characterizes the relationship between the source process and the optimal reproduction process, and hence is valid only for synchronous sampling and for the optimal reproduction process. Consequently, in the RDF analysis the limiting relationship (21) as  $n \rightarrow \infty$  is not even known to exist and, in fact, we can show it exists under a rather strict condition on the distortion (namely, the condition  $D < \min_{0 \leq t \leq T_{ps}} \sigma_{S_c}^2(t)$  stated in Theorem 4). In particular, to prove the statement in Theorem 4, we had to show that from the backward channel (21), we can define an asymptotic relationship, as  $n \rightarrow \infty$ , which corresponds to the asynchronously sampled source process, denoted by  $S_\epsilon[i]$ , and relates  $S_\epsilon[i]$  with its optimal reconstruction process  $\hat{S}_\epsilon[i]$ . This is done by showing that the PDFs for the reproduction process  $\hat{S}_n[i]$  and noise process  $W_n[i]$  from (21), each converge uniformly as  $n \rightarrow \infty$  to a respective limiting PDF, which has to be defined as well. This enabled us to relate the RDFs for the synchronous sampling and for the asynchronous sampling cases using Theorem 3, eventually leading to (22). Accordingly, in our detailed proof of Theorem 4 given in Appendix D, Lemmas D.4 and D.6 as well as a significant part of Lemma D.2 are largely new, addressing the special aspects of the proof arising from the fundamental differences between current setup and the setup in [17], while the derivations of Lemmas D.1 and D.5 follow similarly to ([17] Lemma B.1) and ([17] Lemma B.5), respectively, and parts of Lemma D.2 coincide with ([17] Lemma B.2).

## 5. Numerical Examples

In this section we demonstrate the insights arising from our RDF characterization via numerical examples. Recalling that Theorem 4 states the RDF for asynchronously sampled CT WSCS Gaussian process,  $R_\epsilon(D)$ , as the limit supremum of a sequence of RDFs corresponding to DT memoryless WSCS Gaussian source processes  $\{R_n(D)\}_{n \in \mathcal{N}}$ , we first consider the convergence of  $\{R_n(D)\}_{n \in \mathcal{N}}$  in Section 5.1. Next, in Section 5.2 we study the variation of the RDF of the sampled CT process due to changes in the sampling rate and in the sampling time offset.

Similarly to ([17] Section IV), define a periodic continuous pulse function, denoted by  $\Pi_{t_{dc}, t_{rf}}(t)$ , with equal rise/fall time  $t_{rf} = 0.01$ , duty cycle  $t_{dc} \in [0, 0.98]$ , and period of 1, i.e.,  $\Pi_{t_{dc}, t_{rf}}(t + 1) = \Pi_{t_{dc}, t_{rf}}(t)$  for all  $t \in \mathcal{R}$ . Specifically, for  $t \in [0, 1)$  the function  $\Pi_{t_{dc}, t_{rf}}(t)$  is given by

$$\Pi_{t_{dc}, t_{rf}}(t) = \begin{cases} \frac{t}{t_{rf}} & t \in [0, t_{rf}] \\ 1 & t \in (t_{rf}, t_{dc} + t_{rf}) \\ 1 - \frac{t - t_{dc} - t_{rf}}{t_{rf}} & t \in [t_{dc} + t_{rf}, t_{dc} + 2 \cdot t_{rf}] \\ 0 & t \in (t_{dc} + 2 \cdot t_{rf}, 1). \end{cases} \quad (23)$$

In the following, we model the time varying variance of the WSCS source  $\sigma_{S_c}^2(t)$  to be a linear periodic function of  $\Pi_{t_{dc}, t_{rf}}(t)$ . To that aim, we define a time offset between the first sample and the rise start time of the periodic continuous pulse function; we denote the time offset by  $\phi \in [0, 1)$ . This corresponds to the sampling time offset normalized to the period  $T_{ps}$ . The variance of  $S_c(t)$  is a periodic function with period  $T_{ps}$  which is defined as

$$\sigma_{S_c}^2(t) = 0.2 + 4.8 \cdot \Pi_{t_{dc}, t_{rf}}\left(\frac{t}{T_{ps}} - \phi\right), \quad t \in [0, T_{ps}), \quad (24)$$

with a period of  $T_{ps} = 5 \mu\text{secs}$ .

### 5.1. Convergence of $R_n(D)$ in $n$

From Theorem 4 it follows that if the distortion satisfies  $D < \min_{0 \leq t \leq T_{ps}} \sigma_{S_c}^2(t)$ , the RDF of the asynchronously sampled CT WSCS Gaussian process is given by the limit superior of the sequence  $\{R_n(D)\}_{n \in \mathcal{N}}$ ; where  $R_n(D)$  is defined in Proposition 1. In this subsection, we study the sequence of RDFs  $\{R_n(D)\}_{n \in \mathcal{N}}$  as  $n$  increases. For this evaluation setup, we fixed the distortion constraint at  $D = 0.18$  and set  $\epsilon = \frac{\pi}{7}$  and  $p = 2$ . Let the variance of the CT WSCS Gaussian source process  $\sigma_{S_c}^2(t)$  be modelled by Equation (24) for two sampling time offsets  $\phi = \{0, \frac{1}{16}\}$ . For each offset  $\phi$ , four duty cycle values were considered:  $t_{dc} = [20, 45, 75, 98]\%$ . For each  $n$  we obtain the synchronous sampling mismatch  $\epsilon_n \triangleq \frac{\lfloor n \cdot \epsilon \rfloor}{n}$ , which approaches  $\epsilon$  as  $n \rightarrow \infty$ , where  $n \in \mathcal{N}$ . Since  $\epsilon_n$  is a rational number, corresponding to a sampling period of  $T_s(n) = \frac{T_{ps}}{p + \epsilon_n}$ , then for each  $n$ , the resulting DT process is WSCS with the period  $p_n = p \cdot n + \lfloor n \cdot \epsilon \rfloor$  and its RDF follows from Proposition 1.

Figures 2 and 3 depict  $R_n(D)$  for  $n \in [1, 500]$  with the specified duty cycles and sampling time offsets, where in Figure 2 there is no sampling time offset, i.e.,  $\phi = 0$ , and in Figure 3 the sampling time offset is set to  $\phi = \frac{1}{16}$ . We observe that in both figures the RDF values are higher for higher  $t_{dc}$ . This can be explained by noting that for higher  $t_{dc}$  values, the resulting time-averaged variance of the DT source process increases, hence, a higher number of bits per source sample is required to encode the source process maintaining the same distortion value. Moreover, in all configurations,  $R_n(D)$  varies significantly for smaller values of  $n$ . Comparing Figures 2 and 3, we see that the pattern of these variations depends on the sampling time offset  $\phi$ . For example, when  $t_{dc} = 45\%$  at  $n \in [4, 15]$ , then for  $\phi = 0$  the RDF varies in the range [1.032, 1.143] bits per source sample, while for  $\phi = \frac{1}{16}$  the RDF varies in the range [1.071, 1.237] bits per source sample. However, as  $n$  increases above 230, the variations in  $R_n(D)$  become smaller and are less dependent on the sampling time offset, and the resulting values of  $R_n(D)$  are approximately in the same range for each  $t_{dc}$  in both Figures 2 and 3 for  $n \geq 230$ . This behaviour can be explained by noting that as  $n$  varies, the period  $p_n$  also varies and hence the statistics of the DT variance differs over its respective period. This consequently affects the resulting RDF (especially for small periods). As  $n$  increases  $\epsilon_n$  approaches the asynchronous sampling mismatch  $\epsilon$  and the period  $p_n$  takes a sufficiently large value such that the samples of the DT variance over the period are identically distributed irrespective of the value of  $\phi$ ; leading to a negligible variation in the RDF as seen in the above figures.

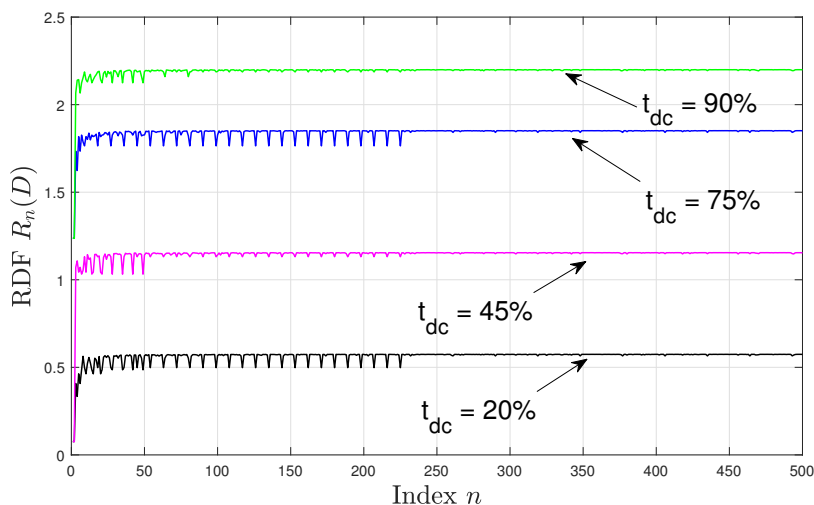


Figure 2.  $R_n(D)$  versus  $n$ ; offset  $\phi = 0$ .

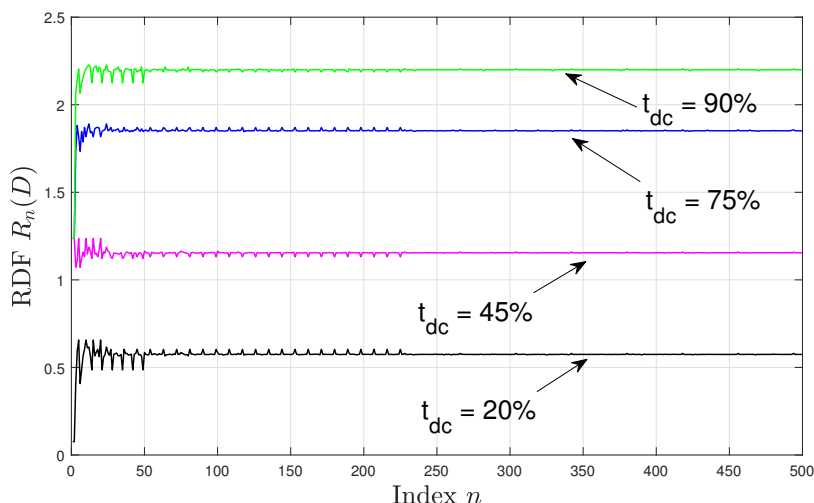


Figure 3.  $R_n(D)$  versus  $n$ ; offset  $\phi = \frac{1}{16}$ .

### 5.2. The Variation of the RDF with the Sampling Rate

Next, we observe the dependence of the RDF for the sampled memoryless WSCS Gaussian process on the value of the sampling interval  $T_s$ . For this setup, we fix the distortion constraint to  $D = 0.18$  and set the duty cycle in the source process (24) to  $t_{dc} = [45, 75]\%$ . Figures 4 and 5 demonstrate the numerically evaluated values for  $R_n(D)$  at sampling intervals in the range  $2 < \frac{T_{ps}}{T_s} < 4$  with the sampling time offsets  $\phi = 0$  and  $\phi = \frac{1}{16}$ , respectively. We note that while the discussion which follows focuses on this range, as it corresponds to relatively low sampling rates—which are typically preferable in practice, the statements and observations regarding the relationship between the denominator of  $\frac{T_{ps}}{T_s}$  and the value of  $R_n(D)$ , and regarding the continuity the RDF in the parameter  $\frac{T_{ps}}{T_s}$ , are directly applicable to any range of values of  $\frac{T_{ps}}{T_s}$ , e.g., when higher sampling rates are preferable. A very important insight which arises from the figures is that the sequence of RDFs  $R_n(D)$  is not convergent; hence, for example, one cannot approach the RDF for  $\frac{T_{ps}}{T_s} = 2.5$  by simply taking rational values of  $\frac{T_{ps}}{T_s}$  which approach 2.5. This verifies that the RDF for asynchronous sampling cannot be obtained by straightforward application of previous results, and indeed, the entire analysis carried out in the manuscript is necessary for the desired characterization.

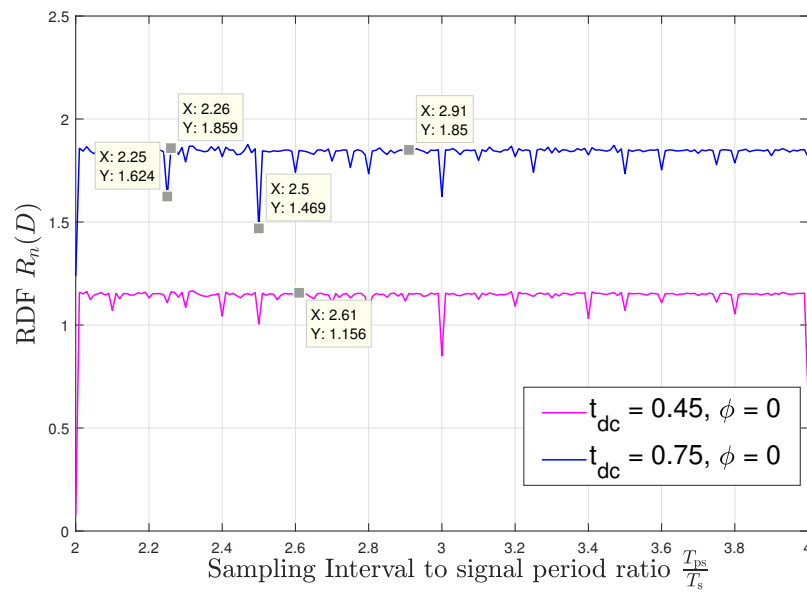


Figure 4.  $R_n(D)$  versus  $\frac{T_{ps}}{T_s}$ ; offset  $\phi = 0$ .

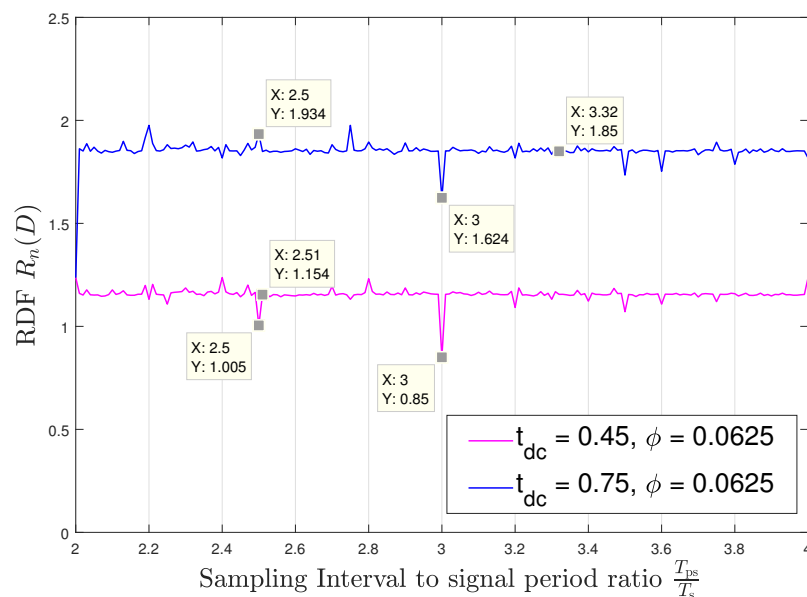


Figure 5.  $R_n(D)$  versus  $\frac{T_{ps}}{T_s}$ ; offset  $\phi = \frac{1}{16}$ .

We observe in Figures 4 and 5 that when  $\frac{T_{ps}}{T_s}$  has a fractional part with a relatively small integer denominator, the variations in the RDF are significant and depend on the sampling time offset. These variations can either degrade the ability to accurately represent the source, which are the observed peaks in Figures 4 and 5, or alternatively, allow to encode the signal to within the same distortion with smaller code rates, corresponding to the deeps in these figures. However, when  $\frac{T_{ps}}{T_s}$  approaches an irrational number, the period of the sampled variance function becomes very long, and consequently, the RDF is approximately constant and independent of the sampling time offset. As an example, consider  $\frac{T_{ps}}{T_s} = 2.5$  and  $t_{dc} = 75\%$ : For sampling time offset  $\phi = 0$  the RDF takes a value of 1.469 bits per source sample, as shown in Figure 4 while for the offset of  $\phi = \frac{1}{16}$  the RDF peaks to 1.934 bits per source sample as can be seen in Figure 5. On the other hand, when approaching asynchronous sampling, the RDF takes a nearly constant value of 1.85 bits per source sample for all the considered



values of  $\frac{T_{ps}}{T_s}$  and this value is invariant to the offset  $\phi$ . This follows since when the denominator of the fractional part of  $\frac{T_{ps}}{T_s}$  increases, then the DT period of the resulting sampled variance,  $p_n$ , increases and practically captures the entire set of values of the CT variance regardless of the sampling time offset. In a similar manner as with the study on capacity in [17], we conjecture that since asynchronous sampling captures the entire set of values of the CT variance, the respective RDF represents the RDF of the analog source, which does not depend on the specific sampling rate and offset. Figures 4 and 5 demonstrate how slight variations in the sampling rate can result in significant changes in the RDF. For instance, at  $\phi = 0$  we observe in Figure 4 that when the sampling rate switches from  $T_s = 2.25 \cdot T_{ps}$  to  $T_s = 2.26 \cdot T_{ps}$ , i.e., the sampling rate switches from being synchronous to being nearly asynchronous, then the RDF changes from 1.624 bits per source sample to 1.859 bits per source sample for  $t_{dc} = 75\%$ ; also, we observe in Figure 5 for  $t_{dc} = 45\%$ , that when the sampling rate switches from  $T_s = 2.5 \cdot T_{ps}$  to  $T_s = 2.51 \cdot T_{ps}$ , i.e., the sampling rate also switches from being synchronous to being nearly asynchronous, then the RDF changes from 1.005 bits per source sample to 1.154 bits per source sample.

Lastly, Figures 6 and 7 numerically evaluate the RDF versus the distortion constraint  $D \in [0.05, 0.19]$  for sampling time offsets of 0 and  $\frac{1}{16}$  respectively. At each  $\phi$ , the result is evaluated at three different values of synchronization mismatch  $\epsilon$ . For this setup, we fix  $t_{dc} = 75\%$ ,  $p = 2$  and  $\epsilon \in \{0.5, \frac{5\pi}{32}, 0.6\}$ . The only mismatch value that refers to the asynchronous sampling case is  $\epsilon = \frac{5\pi}{32}$  and its corresponding sampling interval is approximately 2.007  $\mu$ secs, which is a negligible variation from the sampling intervals corresponding to  $\epsilon \in \{0.5, 0.6\}$ , which are 2.000  $\mu$ secs and 1.923  $\mu$ secs, respectively. Observing both figures, we see that the RDF may vary significantly for very slight variation in the sampling rate. For instance, as shown in Figure 6 for  $\phi = 0$ , at  $D = 0.18$ , a slight change in the synchronization mismatch from  $\epsilon = \frac{5\pi}{32}$  (i.e.,  $T_s \approx 2.007 \mu$ secs) to  $\epsilon = 0.5$  (i.e.,  $T_s = 2.000 \mu$ secs) results to approximately 20% decrease in the RDF. For  $\phi = \frac{1}{16}$  the same change in the sampling synchronization mismatch at  $D = 0.18$  results in an increase in the RDF by roughly 4%. These results demonstrate the unique and counter-intuitive characteristics of the RDF of sampled WSCS signals which arise from our derivation. It is also interesting to examine how the RDF varies with the sampling time offset  $\phi$ . To that aim we plot in Figure 8 the RDF vs.  $\phi$  for the three sampling rates used in Figures 6 and 7 at  $D = 0.18$ . The points marked on the plot correspond to  $\phi = 0$  and  $\phi = \frac{1}{16}$  considered in Figures 6 and 7, respectively. We observe that the RDF is indeed periodic with  $\phi$ . These variations in the RDF occur as by changing  $\phi$  the number of high variance samples within a period of the variance of the DT process changes due to the duty cycle of the CT variance. Then, at  $\phi = 0$  the periodic variance of the DT process corresponding to  $T_s = 2.000 \mu$ secs has the smallest number of high variance values within a period, and when  $\phi = \frac{1}{16}$  the periodic variance of the DT process corresponding to asynchronous sampling has the smallest number of high variance values within a period. For the asynchronous sampling rate the sampling time offset does not matter as in any case (nearly) all values of the CT variance are reflected in the variance of the DT process.

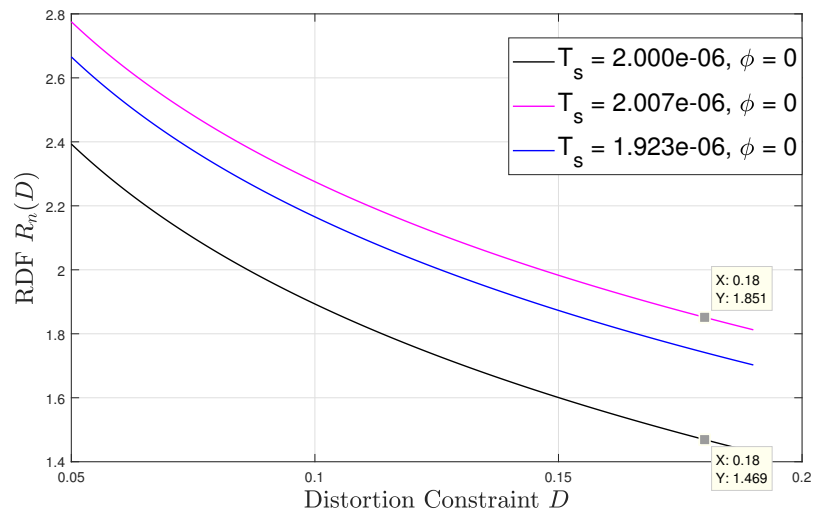


Figure 6.  $R_n(D)$  versus  $D$ ; offset  $\phi = 0$ .

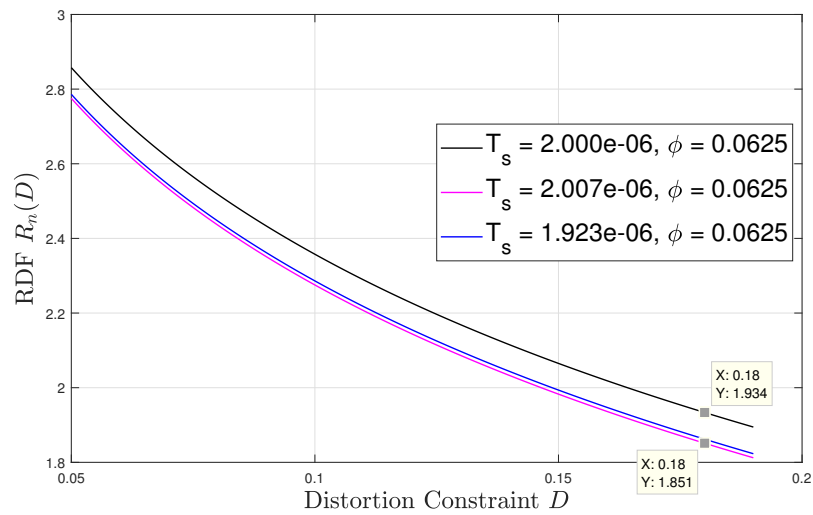


Figure 7.  $R_n(D)$  versus  $D$ ; offset  $\phi = \frac{1}{16}$ .

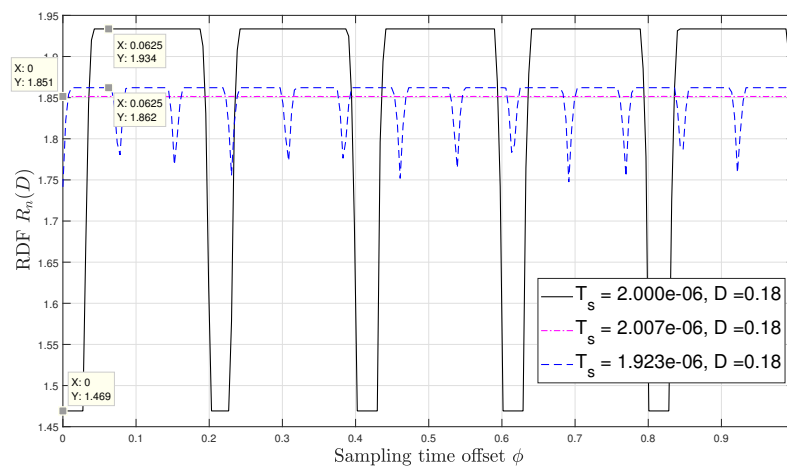


Figure 8.  $R_n(D)$  versus  $\phi$  at  $t_{dc} = 75\%$ .

## 6. Conclusions

In this work the RDF of a sampled CT WSCS Gaussian source process was characterized for scenarios in which the resulting DT process is memoryless and the distortion is relatively small. This characterization shows the relationship between the sampling rate and the minimal number of bits per source sample required for compression at a given distortion. For cases in which the sampling rate is synchronized with the period of the statistics of the source process, the resulting DT process is WSCS and standard information theoretic framework can be used for deriving its RDF. For asynchronous sampling, information stability does not hold, and hence we resort to the information spectrum framework to obtain a characterization. To that aim we derived a relationship between some relevant information spectrum quantities for uniformly convergent sequences of RVs. This relationship was further applied to characterize the RDF of an asynchronously sampled CT WSCS Gaussian source process as the limit superior of a sequence of RDFs, each corresponding to the synchronous sampling of the CT WSCS Gaussian process. The results were derived in the low distortion regime, i.e., under the condition that the distortion constraint  $D$  is less than the minimum variance of the source, and for sampling intervals which are larger than the correlation length of the CT process. Our numerical examples give rise to non-intuitive insights which follow from the derivations. In particular, the numerical evaluation demonstrates that the RDF for a sampled CT WSCS Gaussian source can change dramatically with minor variations in the sampling rate and the sampling time offset. In particular, when the sampling rate switches from being synchronous to being asynchronous and vice versa, the RDF may change considerably as the statistical model of the source switches between WSCS and WSACS. The resulting analysis enables determining the sampling system parameters in order to facilitate accurate and efficient source coding of acquired CT signals.

**Author Contributions:** Conceptualization, E.A., N.S. and R.D.; Methodology, E.A., N.S. and R.D.; Derivation, E.A., N.S. and R.D.; Writing—original draft preparation, E.A., N.S. and R.D.; writing—review and editing, E.A., N.S. and R.D.; Supervision, N.S. and R.D. All authors have read and approved the final published version.

**Funding:** This work was supported by the Israel Science Foundation under Grants 1685/16 and 0100101, and by the Israeli Ministry of Economy through the HERON 5G consortium.

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

## Appendix A. Proof of Lemma 1

**Proof.** To prove that the minimum achievable rate at a given maximum distortion for a code with arbitrary blocklength can be achieved by considering only codes whose blocklength is an integer multiple of  $r$ , we apply the following approach: We first show that every rate-distortion pair achievable when restricted to using source codes whose blocklength is an integer multiple of  $r$  is also achievable when using arbitrary blocklengths; We then prove that every achievable rate-distortion pair is also achievable when restricted to using codes whose blocklength is an integer multiple of  $r$ . Combining these two assertions proves that the rate-distortion function of the source  $\{S[i]\}_{i \in \mathcal{N}}$  can be obtained when restricting the blocklengths to be an integer multiple of  $r$ . Consequently, a reproduction signal  $\{\hat{S}[i]\}_{i \in \mathcal{N}}$  which achieves the minimal rate for a given  $D$  under the restriction to use only blocklengths which are an integer multiple of  $r$  is also the reproduction signal achieving the minimal rate without this restriction, and vice versa, thus proving the lemma.

To prove the first assertion, consider a rate-distortion pair  $(R, D)$  which is achievable when using codes whose blocklength is an integer multiple of  $r$ . It thus follows directly from Definition 6 that for every  $\eta > 0$ ,  $\exists b_0 \in \mathcal{N}$  such that for all  $b > b_0$  there exists a source code  $(R_{(b,r)}, b \cdot r)$  with rate  $R_{(b,r)} \leq R + \eta$  satisfying  $\bar{d}(S^{(b,r)}, \hat{S}^{(b,r)}) \leq D + \frac{\eta}{2}$ . We now show that we can construct a code with an arbitrary blocklength  $l = b \cdot r + j$  where  $0 < j < r$  (i.e., the blocklength  $l$  is not an integer multiple of  $r$ ) satisfying Definition 6 for all  $j \in \{1, \dots, r-1\}$  as follows: Apply the code  $(R_{(b,r)}, b \cdot r)$  to the first  $b \cdot r$  samples of  $S[i]$  and then concatenate each codeword by  $j$  zeros to obtain a source code having

codewords of length  $b \cdot r + j$ . The average distortion (i.e., see (2)) of the resulting  $(R_{(b \cdot r + j)}, b \cdot r + j)$  code is given by:

$$\begin{aligned} \bar{d}(\mathbf{S}^{(b \cdot r + j)}, \hat{\mathbf{S}}^{(b \cdot r + j)}) &= \frac{1}{b \cdot r + j} \left( \sum_{i=1}^{b \cdot r} \mathbb{E} \{ (S[i] - \hat{S}[i])^2 \} + \sum_{i=b \cdot r + 1}^{b \cdot r + j} \mathbb{E} \{ (S[i])^2 \} \right) \\ &= \frac{1}{b \cdot r + j} \left( b \cdot r \cdot \bar{d}(\mathbf{S}^{(b \cdot r)}, \hat{\mathbf{S}}^{(b \cdot r)}) + \sum_{i=1}^j \sigma_S^2[i] \right) \\ &= \frac{b \cdot r}{b \cdot r + j} \cdot \bar{d}(\mathbf{S}^{(b \cdot r)}, \hat{\mathbf{S}}^{(b \cdot r)}) + \frac{1}{b \cdot r + j} \sum_{i=1}^j \sigma_S^2[i]. \end{aligned} \tag{A.1}$$

Thus,  $\exists b_1 > b_0$  such that  $\frac{1}{b_1 \cdot r + j} \sum_{i=1}^j \sigma_S^2[i] < \frac{\eta}{2}$  and hence, for all  $b > b_1$

$$\begin{aligned} \bar{d}(\mathbf{S}^{(b \cdot r + j)}, \hat{\mathbf{S}}^{(b \cdot r + j)}) &= \frac{b \cdot r}{b \cdot r + j} \cdot \bar{d}(\mathbf{S}^{(b \cdot r)}, \hat{\mathbf{S}}^{(b \cdot r)}) + \frac{1}{b \cdot r + j} \sum_{i=1}^j \sigma_S^2[i] \\ &\leq \frac{b \cdot r}{b \cdot r + j} \cdot \bar{d}(\mathbf{S}^{(b \cdot r)}, \hat{\mathbf{S}}^{(b \cdot r)}) + \frac{\eta}{2} \\ &\leq \bar{d}(\mathbf{S}^{(b \cdot r)}, \hat{\mathbf{S}}^{(b \cdot r)}) + \frac{\eta}{2} \leq D + \eta. \end{aligned} \tag{A.2}$$

The rate  $R_{(b \cdot r + j)}$  satisfies:

$$R_{(b \cdot r + j)} = \frac{1}{b \cdot r + j} \cdot \log_2 M = R_{(b \cdot r)} \cdot \frac{b \cdot r}{b \cdot r + j} \leq (R + \eta) \cdot \frac{b \cdot r}{b \cdot r + j} \leq R + \eta. \tag{A.3}$$

Consequently, any rate-distortion pair achievable with codes whose blocklength is an integer multiple of  $r$  can be achieved by codes with arbitrary blocklengths.

Next, we prove that any achievable rate-distortion pair  $(R, D)$  can be achieved by codes whose blocklength is an integer multiple of  $r$ . To that aim, we fix  $\eta > 0$ . By Definition 6, it holds that there exists a code of blocklength  $l$  satisfying (3) and (4). To show that  $(R, D)$  is achievable using codes whose blocklength is an integer multiple of  $r$ , we assume here that  $l$  is not an integer multiple of  $r$ , hence, there exist some positive integers  $b$  and  $j$  such that  $j < r$  and  $l = b \cdot r + j$ . We denote this code by  $(R_{(b \cdot r + j)}, b \cdot r + j)$ . It follows from Definition 6 that  $R_{(b \cdot r + j)} \leq R + \eta$  and  $\bar{d}(\mathbf{S}^{(b \cdot r + j)}, \hat{\mathbf{S}}^{(b \cdot r + j)}) \leq D + \frac{\eta}{2}$ . Next, we construct a code  $(R_{(b+1) \cdot r}, (b+1) \cdot r)$  with codewords whose length is  $(b+1) \cdot r$ , i.e., an integer multiple of  $r$ , by adding  $r - j$  zeros at the end of each codeword of the code  $(R_{(b \cdot r + j)}, b \cdot r + j)$ . The average distortion can now be computed as follows:

$$\begin{aligned} \bar{d}(\mathbf{S}^{((b+1) \cdot r)}, \hat{\mathbf{S}}^{((b+1) \cdot r)}) &= \frac{1}{(b+1) \cdot r} \left( \sum_{i=1}^{b \cdot r + j} \mathbb{E} \{ (S[i] - \hat{S}[i])^2 \} + \sum_{i=b \cdot r + j + 1}^{(b+1) \cdot r} \mathbb{E} \{ (S[i])^2 \} \right) \\ &= \frac{1}{(b+1) \cdot r} \left( (b \cdot r + j) \cdot \bar{d}(\mathbf{S}^{(b \cdot r + j)}, \hat{\mathbf{S}}^{(b \cdot r + j)}) + \sum_{i=b \cdot r + j + 1}^{(b+1) \cdot r} \sigma_S^2[i] \right) \\ &= \frac{b \cdot r + j}{(b+1) \cdot r} \cdot \bar{d}(\mathbf{S}^{(b \cdot r + j)}, \hat{\mathbf{S}}^{(b \cdot r + j)}) + \frac{\sum_{i=b \cdot r + j + 1}^{(b+1) \cdot r} \sigma_S^2[i]}{(b+1) \cdot r}, \end{aligned} \tag{A.4}$$

Thus, since the variance is finite and bounded,  $\exists b_1 > b_0$  such that  $\frac{\sum_{i=b_1 \cdot r + j + 1}^{(b_1+1) \cdot r} \sigma_S^2[i]}{(b_1+1) \cdot r} < \frac{\eta}{2}$  for all  $b > b_1$ . Hence, for all  $b > b_1$

$$\begin{aligned} \bar{d}(\mathbf{S}^{((b+1) \cdot r)}, \hat{\mathbf{S}}^{((b+1) \cdot r)}) &\leq \frac{b \cdot r + j}{(b+1) \cdot r} \cdot \bar{d}(\mathbf{S}^{(b \cdot r + j)}, \hat{\mathbf{S}}^{(b \cdot r + j)}) + \frac{\eta}{2} \\ &\leq \bar{d}(\mathbf{S}^{(b \cdot r + j)}, \hat{\mathbf{S}}^{(b \cdot r + j)}) + \frac{\eta}{2} \leq D + \eta. \end{aligned} \tag{A.5}$$

The rate  $R_{(b+1) \cdot r}$  can be expressed as follows:

$$R_{(b+1) \cdot r} = \frac{1}{(b+1) \cdot r} \cdot \log_2 M = R_{(b \cdot r + j)} \cdot \frac{b \cdot r + j}{(b+1) \cdot r} \leq (R + \eta) \cdot \frac{b \cdot r + j}{(b+1) \cdot r} < R + \eta. \tag{A.6}$$

It follows that  $R_{(b+1) \cdot r} \leq R + \eta$  for any arbitrary  $\eta$  by selecting a sufficiently large  $b$ . This proves that every rate-distortion pair achievable with arbitrary blocklengths (e.g.,  $l = b \cdot r + j, j < r$ ) is also achievable when considering source codes whose blocklength is an integer multiple of  $r$  (i.e.,  $l = b \cdot r$ ). This concludes the proof.  $\square$

**Appendix B. Proof of Theorem 1**

Recall that  $\alpha \in \mathcal{R}^{++}$ . To prove the theorem, we fix a rate-distortion pair  $(R, D)$  that is achievable for the source  $\{S[i]\}_{i \in \mathcal{N}}$ . By Definition 6 this implies that for all  $\eta > 0$  there exists  $l_0(\eta) \in \mathcal{N}$  such that for all  $l > l_0(\eta)$  there exists a source code  $\mathcal{C}_l$  with rate  $R_{(l)} \leq R + \eta$  and MSE distortion  $D_{(l)} = \mathbb{E}\{\frac{1}{l} \|\mathbf{S}^{(l)} - \hat{\mathbf{S}}^{(l)}\|^2\} \leq D + \eta$ , where  $\|\cdot\|$  denotes the norm of a vector. Next, we use the code  $\mathcal{C}_l$  to define the source code  $\mathcal{C}_l^{(\alpha)}$ , which operates in the following manner: The encoder first scales its input block by  $1/\alpha$ . Then, the block is encoded using the source code  $\mathcal{C}_l$ . Finally, the selected codeword is scaled by  $\alpha$ . Since the  $\mathcal{C}_l^{(\alpha)}$  has the same number of codewords and the same blocklength as  $\mathcal{C}_l$ , it follows that its rate, denote  $R_{(l)}^{(\alpha)}$ , satisfied  $R_{(l)}^{(\alpha)} = R_{(l)} \leq R + \eta$ . Furthermore, by the construction of  $\mathcal{C}_l^{(\alpha)}$ , it holds that its reproduction vector when applied to  $\alpha \cdot \mathbf{S}^{(l)}$  is equal to the output of  $\mathcal{C}_l$  applied to  $\mathbf{S}^{(l)}$  scaled by  $\alpha$ , i.e.,  $\alpha \cdot \hat{\mathbf{S}}^{(l)}$ . Consequently, the MSE of  $\mathcal{C}_l^{(\alpha)}$  when applied to the source  $\{\alpha \cdot S[i]\}_{i \in \mathcal{N}}$ , denoted  $D_{(l)}^{(\alpha)}$ , satisfies  $D_{(l)}^{(\alpha)} = \mathbb{E}\{\frac{1}{l} \|\alpha \cdot \mathbf{S}^{(l)} - \alpha \cdot \hat{\mathbf{S}}^{(l)}\|^2\} = \alpha^2 \cdot D_{(l)} \leq \alpha^2 \cdot D + \alpha^2 \eta$ .

It thus follows that for all  $\tilde{\eta} > 0$  there exists  $\tilde{l}_0(\tilde{\eta}) = l_0(\min(\tilde{\eta}, \alpha^2 \tilde{\eta}))$  such that for all  $l > \tilde{l}_0(\tilde{\eta})$  there exists a code  $\mathcal{C}_l^{(\alpha)}$  with rate  $R_{(l)}^{(\alpha)} \leq R + \tilde{\eta}$  which achieves an MSE distortion of  $D_{(l)}^{(\alpha)} \leq \alpha^2 \cdot D + \tilde{\eta}$  when applied to the compression of  $\{\alpha \cdot S[i]\}_{i \in \mathcal{N}}$ . Hence,  $(R, \alpha^2 \cdot D)$  is achievable for compression of  $\{\alpha \cdot S[i]\}_{i \in \mathcal{N}}$  by Definition 6, proving the theorem.

**Appendix C. Proof of Theorem 3**

In this appendix, we prove (17b) by applying a similar approach as used for proving (17a) in ([17] Appendix A). We first note that Definition 9 can also be written as follows:

$$p\text{-}\limsup_{k \rightarrow \infty} Z_k \stackrel{(a)}{=} \inf \left\{ \beta \in \mathcal{R} \mid \limsup_{k \rightarrow \infty} \Pr(Z_k > \beta) = 0 \right\} \stackrel{(b)}{=} \inf \left\{ \beta \in \mathcal{R} \mid \liminf_{k \rightarrow \infty} F_k(\beta) = 1 \right\}. \tag{C.1}$$

For the equality (a), we note that the set of probabilities  $\{\Pr(Z_k > \beta)\}_{k \in \mathcal{N}}$  is non-negative and bounded in  $[0, 1]$ ; hence, for any  $\beta \in \mathcal{R}$  for which  $\limsup_{k \rightarrow \infty} \Pr(Z_k > \beta) = 0$ , it also holds from ([31] Theorem 3.17) that the limit of any subsequence of  $\{\Pr(Z_k > \beta)\}_{k \in \mathcal{N}}$  is also 0, since

non-negativity of the probability implies  $\liminf_{k \rightarrow \infty} \Pr(Z_k > \beta) \geq 0$ . Then, combined with the relationship  $\liminf_{k \rightarrow \infty} \Pr(Z_k > \beta) \leq \limsup_{k \rightarrow \infty} \Pr(Z_k > \beta)$ , we conclude:

$$0 \leq \liminf_{k \rightarrow \infty} \Pr(Z_k > \beta) \leq \limsup_{k \rightarrow \infty} \Pr(Z_k > \beta) = 0$$

$$\implies \liminf_{k \rightarrow \infty} \Pr(Z_k > \beta) = \limsup_{k \rightarrow \infty} \Pr(Z_k > \beta) \stackrel{(a)}{=} \lim_{k \rightarrow \infty} \Pr(Z_k > \beta) = 0,$$

where (a) follows from ([31] Example 3.18(c)). This implies  $\lim_{k \rightarrow \infty} \Pr(Z_k > \beta)$  exists and is equal to 0.

In the opposite direction, if  $\lim_{k \rightarrow \infty} \Pr(Z_k > \beta) = 0$  then it follows from ([31] Example 3.18(c)) that  $\limsup_{k \rightarrow \infty} \Pr(Z_k > \beta) = 0$ . Next, we note that since  $F_k(\beta)$  is bounded in  $[0, 1]$  then  $\liminf_{k \rightarrow \infty} F_k(\beta)$  is finite  $\forall \beta \in \mathcal{R}$ , even if  $\lim_{k \rightarrow \infty} F_k(\beta)$  does not exist. Equality (b) follows since  $\limsup_{k \rightarrow \infty} \Pr(Z_k > \beta) = \limsup_{k \rightarrow \infty} (1 - \Pr(Z_k \leq \beta))$  which according to ([32] Theorem 7.3.7) is equal to  $1 + \limsup_{k \rightarrow \infty} (-\Pr(Z_k \leq \beta))$ . By ([33] Chapter 1, page 29), this quantity is also equal to  $1 - \liminf_{k \rightarrow \infty} \Pr(Z_k \leq \beta) = 1 - \liminf_{k \rightarrow \infty} F_k(\beta)$ .

Next, we state the following lemma:

**Lemma C.1.** Given assumption AS2, for all  $\beta \in \mathcal{R}$  it holds that

$$\liminf_{k \rightarrow \infty} F_k(\beta) = \lim_{n \rightarrow \infty} \liminf_{k \rightarrow \infty} \tilde{F}_{k,n}(\beta). \tag{C.2}$$

**Proof.** To prove the lemma we first show that  $\liminf_{k \rightarrow \infty} F_k(\beta) \leq \lim_{n \rightarrow \infty} \liminf_{k \rightarrow \infty} \tilde{F}_{k,n}(\beta)$ , and then we show  $\liminf_{k \rightarrow \infty} F_k(\beta) \geq \lim_{n \rightarrow \infty} \liminf_{k \rightarrow \infty} \tilde{F}_{k,n}(\beta)$ . Recall that by AS2, for all  $\beta \in \mathcal{R}$  and  $k \in \mathcal{N}$ ,  $\tilde{F}_{k,n}(\beta)$  converges as  $n \rightarrow \infty$  to  $F_k(\beta)$ , uniformly over  $k$  and  $\beta$ , i.e., for all  $\eta > 0$  there exists  $n_0(\eta) \in \mathcal{N}$ ,  $k_0(n_0(\eta), \eta) \in \mathcal{N}$  such that for every  $n > n_0(\eta)$ ,  $\beta \in \mathcal{R}$  and  $k > k_0(n_0(\eta), \eta)$ , it holds that  $|\tilde{F}_{k,n}(\beta) - F_k(\beta)| < \eta$ . Consequently, for every subsequence  $0 < k_1 < k_2 < \dots$  such that  $\lim_{l \rightarrow \infty} \tilde{F}_{k_l,n}(\beta)$  exists for any  $n > n_0(\eta)$ , it follows from ([31] Theorem 7.11) that, as the convergence over  $k$  is uniform, the limits over  $n$  and  $l$  are interchangeable:

$$\lim_{n \rightarrow \infty} \lim_{l \rightarrow \infty} \tilde{F}_{k_l,n}(\beta) = \lim_{l \rightarrow \infty} \lim_{n \rightarrow \infty} \tilde{F}_{k_l,n}(\beta) = \lim_{l \rightarrow \infty} F_{k_l}(\beta). \tag{C.3}$$

The existence of such a convergent subsequence is guaranteed by the Bolzano-Weierstrass theorem ([31] Theorem 2.42) as  $\tilde{F}_{k,n}(\beta) \in [0, 1]$ .

From the properties of the limit inferior ([31] Theorem 3.17) it follows that there exists a subsequence of  $\{F_k(\beta)\}_{k \in \mathcal{N}}$ , denoted  $\{F_{k_m}(\beta)\}_{m \in \mathcal{N}'}$  such that  $\lim_{m \rightarrow \infty} F_{k_m}(\beta) = \liminf_{k \rightarrow \infty} F_k(\beta)$ . Consequently,

$$\liminf_{k \rightarrow \infty} F_k(\beta) = \lim_{m \rightarrow \infty} F_{k_m}(\beta) \stackrel{(a)}{=} \lim_{n \rightarrow \infty} \lim_{m \rightarrow \infty} \tilde{F}_{k_m,n}(\beta)$$

$$\stackrel{(b)}{\geq} \lim_{n \rightarrow \infty} \liminf_{k \rightarrow \infty} \tilde{F}_{k,n}(\beta), \tag{C.4}$$

where (a) follows from (C.3), and (b) follows from the definition of the limit inferior ([31] Definition 3.16). Similarly, by ([31] Theorem 3.17), for any  $n \in \mathcal{N}$  there exists a subsequence of

$\{\tilde{F}_{k,n}(\beta)\}_{k \in \mathcal{N}}$  which we denote by  $\{\tilde{F}_{k_l,n}(\beta)\}_{l \in \mathcal{N}}$  where  $\{k_l\}_{l \in \mathcal{N}}$  satisfy  $0 < k_1 < k_2 < \dots$ , such that  $\lim_{l \rightarrow \infty} \tilde{F}_{k_l,n}(\beta) = \liminf_{k \rightarrow \infty} \tilde{F}_{k,n}(\beta)$ . Therefore,

$$\begin{aligned} \lim_{n \rightarrow \infty} \liminf_{k \rightarrow \infty} \tilde{F}_{k,n}(\beta) &= \lim_{n \rightarrow \infty} \lim_{l \rightarrow \infty} \tilde{F}_{k_l,n}(\beta) \\ &\stackrel{(a)}{=} \lim_{l \rightarrow \infty} F_{k_l}(\beta) \stackrel{(b)}{\geq} \liminf_{k \rightarrow \infty} F_k(\beta), \end{aligned} \tag{C.5}$$

where (a) follows from (C.3), and (b) follows from the definition of the limit inferior ([31] Definition 3.16). Therefore,  $\liminf_{k \rightarrow \infty} F_k(\beta) \leq \lim_{n \rightarrow \infty} \liminf_{k \rightarrow \infty} \tilde{F}_{k,n}(\beta)$ . Combining (C.4) and (C.5) proves (C.2) in the statement of the lemma.  $\square$

**Lemma C.2.** *Given assumptions AS1–AS2, the sequence of RVs  $\{\tilde{Z}_{k,n}\}_{k,n \in \mathcal{N}}$  satisfies*

$$\lim_{n \rightarrow \infty} \left( \text{p-} \limsup_{k \rightarrow \infty} \tilde{Z}_{k,n} \right) = \inf \left\{ \beta \in \mathcal{R} \mid \lim_{n \rightarrow \infty} \liminf_{k \rightarrow \infty} \tilde{F}_{k,n}(\beta) = 1 \right\}. \tag{C.6}$$

**Proof.** Since by assumption AS1, for every  $n \in \mathcal{N}$ , every convergent subsequence of  $\{\tilde{Z}_{k,n}\}_{k \in \mathcal{N}}$  converges in distribution as  $k \rightarrow \infty$  to a deterministic scalar, it follows that every convergent subsequence of  $\tilde{F}_{k,n}(\beta)$  converges as  $k \rightarrow \infty$  to a step function, which is the CDF of the corresponding sublimit of  $\tilde{Z}_{k,n}$ . In particular, the limit  $\liminf_{k \rightarrow \infty} \tilde{F}_{k,n}(\beta)$  is a step function representing the CDF of the deterministic scalar  $\zeta_n$ , i.e.,

$$\liminf_{k \rightarrow \infty} \tilde{F}_{k,n}(\beta) = \begin{cases} 0 & \beta < \zeta_n \\ 1 & \beta \geq \zeta_n. \end{cases} \tag{C.7}$$

Since, by Lemma C.1, AS2 implies that the limit  $\lim_{n \rightarrow \infty} \liminf_{k \rightarrow \infty} \tilde{F}_{k,n}(\beta)$  exists (convergence to a discontinuous function is in the sense of ([31] Ex. 7.3)), then  $\lim_{n \rightarrow \infty} \zeta_n$  exists. Hence, we obtain that

$$\lim_{n \rightarrow \infty} \liminf_{k \rightarrow \infty} \tilde{F}_{k,n}(\beta) = \begin{cases} 0 & \beta < \lim_{n \rightarrow \infty} \zeta_n \\ 1 & \beta \geq \lim_{n \rightarrow \infty} \zeta_n, \end{cases} \tag{C.8}$$

and from the right-hand side of (C.6) we have that

$$\inf \left\{ \beta \in \mathcal{R} \mid \lim_{n \rightarrow \infty} \liminf_{k \rightarrow \infty} \tilde{F}_{k,n}(\beta) = 1 \right\} = \lim_{n \rightarrow \infty} \zeta_n. \tag{C.9}$$

Next, from (C.1) and (C.7) we note that

$$\text{p-} \limsup_{k \rightarrow \infty} \tilde{Z}_{k,n} = \inf \left\{ \beta \in \mathcal{R} \mid \liminf_{k \rightarrow \infty} \tilde{F}_{k,n}(\beta) = 1 \right\} = \zeta_n.$$

Consequently, the left-hand side of (C.6) is equal to  $\lim_{n \rightarrow \infty} \zeta_n$ . Combining with (C.9) we arrive at the equality (C.6) in the statement of the lemma.  $\square$

Substituting (C.2) into (C.1) results in

$$\text{p-} \limsup_{k \rightarrow \infty} Z_k = \inf \left\{ \beta \in \mathcal{R} \mid \lim_{n \rightarrow \infty} \liminf_{k \rightarrow \infty} \tilde{F}_{k,n}(\beta) = 1 \right\} \stackrel{(a)}{=} \lim_{n \rightarrow \infty} \left( \text{p-} \limsup_{k \rightarrow \infty} \tilde{Z}_{k,n} \right), \tag{C.10}$$

where (a) follows from (C.6). Equation (C.10) concludes the proof for (17b).

**Appendix D. Proof of Theorem 4**

In this appendix we detail the proof of Theorem 4. The outline of the proof is given as follows:

- We first show in Appendix D.1 that for any  $k \in \mathcal{N}$ , the PDF of the random vector  $\mathbf{S}_n^{(k)}$ , representing the first  $k$  samples of the CT WSCS source  $S_c(t)$  sampled at time instants  $T_s(n) = \frac{T_{ps}}{p+\epsilon_n}$ , converges in the limit as  $n \rightarrow \infty$  and for any  $k \in \mathcal{N}$  to the PDF of  $\mathbf{S}_\epsilon^{(k)}$ , which represents the first  $k$  samples of the CT WSCS source  $S_c(t)$ , sampled at time instants  $T_s = \frac{T_{ps}}{p+\epsilon}$ . We prove that this convergence is uniform in  $k \in \mathcal{N}$  and in the realization vector  $\mathbf{s}^{(k)} \in \mathcal{R}^k$ . This is stated in Lemma D.1.
- Next, in Appendix D.2 we apply Theorem 3 to relate the mutual information density rates for the random source vector  $\mathbf{S}_n^{(k)}$  and its reproduction  $\hat{\mathbf{S}}_n^{(k)}$  with that of the random source vector  $\mathbf{S}_\epsilon^{(k)}$  and its reproduction  $\hat{\mathbf{S}}_\epsilon^{(k)}$ . To that aim, let the functions  $F_{S_n, \hat{S}_n}$  and  $F_{S_\epsilon, \hat{S}_\epsilon}$  denote the joint distributions of an arbitrary dimensional source and reproduction vectors corresponding to the synchronously sampled and to the asynchronously sampled source process respectively. We define the following mutual information density rates:

$$\tilde{Z}'_{k,n} (F_{S_n, \hat{S}_n}) \triangleq \frac{1}{k} \log \frac{p_{\mathbf{S}_n^{(k)} | \hat{\mathbf{S}}_n^{(k)}} (\mathbf{S}_n^{(k)} | \hat{\mathbf{S}}_n^{(k)})}{p_{\mathbf{S}_n^{(k)}} (\mathbf{S}_n^{(k)})}, \tag{D.1a}$$

and

$$Z'_{k,\epsilon} (F_{S_\epsilon, \hat{S}_\epsilon}) \triangleq \frac{1}{k} \log \frac{p_{\mathbf{S}_\epsilon^{(k)} | \hat{\mathbf{S}}_\epsilon^{(k)}} (\mathbf{S}_\epsilon^{(k)} | \hat{\mathbf{S}}_\epsilon^{(k)})}{p_{\mathbf{S}_\epsilon^{(k)}} (\mathbf{S}_\epsilon^{(k)})}, \tag{D.1b}$$

$k, n \in \mathcal{N}$ . The RVs  $\tilde{Z}'_{k,n} (F_{S_n, \hat{S}_n})$  and  $Z'_{k,\epsilon} (F_{S_\epsilon, \hat{S}_\epsilon})$  in (D.1) denote the mutual information density rates ([14] Definition 3.2.1) between the DT source process and the corresponding reproduction process for the case of synchronous sampling and for the case of asynchronous sampling, respectively.

We then show that if the pairs of source process and optimal reproduction process  $\{S_n[i], \hat{S}_n[i]\}_{i \in \mathcal{N}}$  and  $\{S_\epsilon[i], \hat{S}_\epsilon[i]\}_{i \in \mathcal{N}}$  satisfy that  $p_{\mathbf{S}_n^{(k)}} (\mathbf{s}^{(k)}) \xrightarrow{n \rightarrow \infty} p_{\mathbf{S}_\epsilon^{(k)}} (\mathbf{s}^{(k)})$  uniformly with respect to  $\hat{\mathbf{s}}^{(k)} \in \mathcal{R}^k$  and  $k \in \mathcal{N}$ , and that  $p_{\mathbf{S}_n^{(k)} | \hat{\mathbf{S}}_n^{(k)}} (\mathbf{s}^{(k)} | \hat{\mathbf{s}}^{(k)}) \xrightarrow{n \rightarrow \infty} p_{\mathbf{S}_\epsilon^{(k)} | \hat{\mathbf{S}}_\epsilon^{(k)}} (\mathbf{s}^{(k)} | \hat{\mathbf{s}}^{(k)})$  uniformly in  $\left( (\hat{\mathbf{s}}^{(k)})^T, (\mathbf{s}^{(k)})^T \right)^T \in \mathcal{R}^{2k}$  and  $k \in \mathcal{N}$ , then  $\tilde{Z}'_{k,n} (F_{S_n, \hat{S}_n}) \xrightarrow[n \rightarrow \infty]{(dist.)} Z'_{k,\epsilon} (F_{S_\epsilon, \hat{S}_\epsilon})$  uniformly in  $k \in \mathcal{N}$ . In addition, Lemma D.3 proves that every subsequence of  $\left\{ \tilde{Z}'_{k,n} (F_{S_n, \hat{S}_n}) \right\}_{k \in \mathcal{N}}$  w.r.t.  $k$ , indexed as  $k_l$  converges in distribution, in the limit  $l \rightarrow \infty$  to a deterministic scalar.

- Lastly, in Appendix D.3 we combine the above results to show in Lemmas D.5 and D.6 that  $R_\epsilon(D) \leq \limsup_{n \rightarrow \infty} R_n(D)$  and  $R_\epsilon(D) \geq \limsup_{n \rightarrow \infty} R_n(D)$  respectively; implying that  $R_\epsilon(D) = \limsup_{n \rightarrow \infty} R_n(D)$ , which proves the theorem.

To facilitate our proof we will need uniform convergence in  $k \in \mathcal{N}$ , of  $p_{\mathbf{S}_n^{(k)}} (\mathbf{s}^{(k)})$ ,  $p_{\hat{\mathbf{S}}_n^{(k)}} (\hat{\mathbf{s}}^{(k)})$  and  $p_{\mathbf{S}_n^{(k)} | \hat{\mathbf{S}}_n^{(k)}} (\mathbf{s}^{(k)} | \hat{\mathbf{s}}^{(k)})$  to  $p_{\mathbf{S}_\epsilon^{(k)}} (\mathbf{s}^{(k)})$ ,  $p_{\hat{\mathbf{S}}_\epsilon^{(k)}} (\hat{\mathbf{s}}^{(k)})$  and  $p_{\mathbf{S}_\epsilon^{(k)} | \hat{\mathbf{S}}_\epsilon^{(k)}} (\mathbf{s}^{(k)} | \hat{\mathbf{s}}^{(k)})$ , respectively. To that aim, we will make the following scaling assumption w.l.o.g.:

**Assumption D.1.** The variance of the source and the allowed distortion are scaled by some factor  $\alpha^2$  such that

$$\alpha^2 \cdot \min \left\{ D, \left( \min_{0 \leq t \leq T_{ps}} \sigma_{S_c}^2(t) - D \right) \right\} > \frac{1}{2\pi}. \tag{D.2}$$



Note that this assumption has no effect on the generality of the RDF for multivariate stationary processes detailed in ([5] Section 10.3.3), ([34] Section IV). Moreover, by Theorem 1, for every  $\alpha > 0$  it holds that when any rate  $R$  achievable when compressing the original source  $S_c(t)$  with distortion not larger than  $D$  is achievable when compressing the scaled source  $\alpha \cdot S_c(t)$  with distortion not larger than  $\alpha^2 \cdot D$ . Note that if for the source  $S_c(t)$  the distortion satisfies  $D < \min_{0 \leq t \leq T_{ps}} \sigma_{S_c}^2(t)$ , then for the scaled source and distortion we have  $\alpha^2 \cdot D < \min_{0 \leq t \leq T_{ps}} \alpha^2 \cdot \sigma_{S_c}^2(t)$ .

Appendix D.1. Convergence in Distribution of  $S_n^{(k)}$  to  $S_\epsilon^{(k)}$  Uniformly with Respect to  $k \in \mathcal{N}$

In order to prove the uniform convergence in distribution,  $S_n^{(k)} \xrightarrow[n \rightarrow \infty]{(dist.)} S_\epsilon^{(k)}$ , uniformly with respect to  $k \in \mathcal{N}$ , we first prove, in Lemma D.1, that as  $n \rightarrow \infty$  the sequence of PDFs of  $S_n^{(k)}$ ,  $p_{S_n^{(k)}}(s^{(k)})$ , converges to the PDF of  $S_\epsilon^{(k)}$ ,  $p_{S_\epsilon^{(k)}}(s^{(k)})$ , uniformly in  $s^{(k)} \in \mathcal{R}^k$  and in  $k \in \mathcal{N}$ . Next, we show in Corollary D.1 that  $S_n^{(k)} \xrightarrow[n \rightarrow \infty]{(dist.)} S_\epsilon^{(k)}$  uniformly in  $k \in \mathcal{N}$ .

To that aim, let us define the set  $\mathcal{K} \triangleq \{1, 2, \dots, k\}$  and consider the  $k$ -dimensional zero-mean, memoryless random vectors  $S_n^{(k)}$  and  $S_\epsilon^{(k)}$  with their respective diagonal correlation matrices expressed below:

$$R_n^{(k)} \triangleq \mathbb{E}\{(S_n^{(k)})(S_n^{(k)})^T\} = \text{diag}(\sigma_{S_n}^2[1], \dots, \sigma_{S_n}^2[k]), \tag{D.3a}$$

$$R_\epsilon^{(k)} \triangleq \mathbb{E}\{(S_\epsilon^{(k)})(S_\epsilon^{(k)})^T\} = \text{diag}(\sigma_{S_\epsilon}^2[1], \dots, \sigma_{S_\epsilon}^2[k]). \tag{D.3b}$$

Since  $\epsilon_n \triangleq \frac{|n-\epsilon|}{n}$  it holds that  $\frac{n-\epsilon-1}{n} \leq \epsilon_n \leq \frac{n-\epsilon}{n}$ ; therefore

$$\lim_{n \rightarrow \infty} \epsilon_n = \epsilon. \tag{D.4}$$

Now we note that since  $\sigma_{S_c}^2(t)$  is uniformly continuous, then by the definition of a uniformly continuous function, for each  $i \in \mathcal{N}$ , the limit in (D.4) implies that

$$\lim_{n \rightarrow \infty} \sigma_{S_n}^2[i] \equiv \lim_{n \rightarrow \infty} \sigma_{S_c}^2\left(i \cdot \frac{T_{ps}}{p + \epsilon_n}\right) = \sigma_{S_c}^2\left(i \cdot \frac{T_{ps}}{p + \epsilon}\right) \equiv \sigma_{S_\epsilon}^2[i]. \tag{D.5}$$

From Assumption D.1, it follows that  $\sigma_{S_n}^2[i]$  satisfies  $\sigma_{S_n}^2[i] > \frac{1}{2\pi}$ ; Hence, we can state the following lemma:

**Lemma D.1.** The PDF of  $S_n^{(k)}$ ,  $p_{S_n^{(k)}}(s^{(k)})$ , converges as  $n \rightarrow \infty$  to the PDF of  $S_\epsilon^{(k)}$ ,  $p_{S_\epsilon^{(k)}}(s^{(k)})$ , uniformly in  $s^{(k)} \in \mathcal{R}^k$  and in  $k \in \mathcal{N}$ :

$$\lim_{n \rightarrow \infty} p_{S_n^{(k)}}(s^{(k)}) = p_{S_\epsilon^{(k)}}(s^{(k)}), \quad \forall s^{(k)} \in \mathcal{R}^k, \forall k \in \mathcal{N}.$$

**Proof.** The proof of the lemma directly follows from the steps in the proof of ([17] Lemma B.1), which was applied to random Gaussian processes with independent entries and variance larger than  $\frac{1}{2\pi}$ . □

Lemma D.1 gives rise to the following corollary:

**Corollary D.1.** For any  $k \in \mathcal{N}$  it holds that  $S_n^{(k)} \xrightarrow[n \rightarrow \infty]{(dist.)} S_\epsilon^{(k)}$ , and convergence is uniform over  $k$ .

**Proof.** The corollary holds due to ([35] Theorem 1): Since  $p_{S_n^{(k)}}(s^{(k)})$  converges to  $p_{S_\epsilon^{(k)}}(s^{(k)})$  then  $S_n^{(k)} \xrightarrow[n \rightarrow \infty]{(dist.)} S_\epsilon^{(k)}$ . In addition, since the convergence of the PDFs is uniform in  $k \in \mathcal{N}$ , the convergence of the CDFs is also uniform in  $k \in \mathcal{N}$ . □

Appendix D.2. Showing that  $\tilde{Z}'_{k,n} \left( F_{S_n, \hat{S}_n}^{\text{opt}} \right)$  and  $Z'_{k,\epsilon} \left( F_{S_\epsilon, \hat{S}_\epsilon} \right)$  Satisfy the Conditions of Theorem 3

Let  $F_{S_n, \hat{S}_n}^{\text{opt}}$  denote the joint distribution for the source process and the corresponding optimal reproduction process satisfying the distortion constraint  $D$ . We next prove that for  $F_{S_n, \hat{S}_n}^{\text{opt}} \xrightarrow[n \rightarrow \infty]{(dist.)} F_{S_\epsilon, \hat{S}_\epsilon}$ , then  $\tilde{Z}'_{k,n} \left( F_{S_n, \hat{S}_n}^{\text{opt}} \right)$  and  $Z'_{k,\epsilon} \left( F_{S_\epsilon, \hat{S}_\epsilon} \right)$  satisfy AS1–AS2. In particular, Lemma D.2 proves that  $\tilde{Z}'_{k,n} \left( F_{S_n, \hat{S}_n}^{\text{opt}} \right) \xrightarrow[n \rightarrow \infty]{(dist.)} Z'_{k,\epsilon} \left( F_{S_\epsilon, \hat{S}_\epsilon} \right)$  uniformly in  $k \in \mathcal{N}$  for the optimal zero-mean Gaussian reproduction vectors with independent entries. Lemma D.3 proves that for any fixed  $n$ ,  $\tilde{Z}'_{k,n} \left( F_{S_n, \hat{S}_n}^{\text{opt}} \right)$  converges in distribution to a deterministic scalar as  $k \rightarrow \infty$ .

**Lemma D.2.** Let  $\{\hat{\mathbf{s}}_n^{(k)}\}_{n \in \mathcal{N}}$  and  $\{\mathbf{w}_n^{(k)}\}_{n \in \mathcal{N}}$  be two sets of mutually independent sequences of  $k \times 1$  zero-mean Gaussian random vectors related via the backward channel (20), each having independent entries and let PDFs  $p_{\hat{\mathbf{s}}_n^{(k)}} \left( \hat{\mathbf{s}}^{(k)} \right)$  and  $p_{\mathbf{w}_n^{(k)}} \left( \mathbf{w}^{(k)} \right)$ , respectively, denote their PDFs. Consider two other zero-mean Gaussian random vectors  $\hat{\mathbf{S}}_\epsilon^{(k)}$  and  $\mathbf{W}_\epsilon^{(k)}$  each having independent entries with the PDFs  $p_{\hat{\mathbf{S}}_\epsilon^{(k)}} \left( \hat{\mathbf{s}}^{(k)} \right)$  and  $p_{\mathbf{W}_\epsilon^{(k)}} \left( \mathbf{w}^{(k)} \right)$ , respectively, such that  $\lim_{n \rightarrow \infty} p_{\hat{\mathbf{s}}_n^{(k)}} \left( \hat{\mathbf{s}}^{(k)} \right) = p_{\hat{\mathbf{S}}_\epsilon^{(k)}} \left( \hat{\mathbf{s}}^{(k)} \right)$  uniformly in  $\hat{\mathbf{s}}^{(k)} \in \mathcal{R}^k$  and uniformly with respect to  $k \in \mathcal{N}$ , and  $\lim_{n \rightarrow \infty} p_{\mathbf{w}_n^{(k)}} \left( \mathbf{w}^{(k)} \right) = p_{\mathbf{W}_\epsilon^{(k)}} \left( \mathbf{w}^{(k)} \right)$  uniformly in  $\mathbf{w}^{(k)} \in \mathcal{R}^k$  and uniformly with respect to  $k \in \mathcal{N}$ . Then, the RVs  $\tilde{Z}'_{k,n} \left( F_{S_n, \hat{S}_n}^{\text{opt}} \right)$  and  $Z'_{k,\epsilon} \left( F_{S_\epsilon, \hat{S}_\epsilon} \right)$ , defined via (D.1) satisfy  $\tilde{Z}'_{k,n} \left( F_{S_n, \hat{S}_n}^{\text{opt}} \right) \xrightarrow[n \rightarrow \infty]{(dist.)} Z'_{k,\epsilon} \left( F_{S_\epsilon, \hat{S}_\epsilon} \right)$  uniformly over  $k \in \mathcal{N}$ .

**Proof.** To begin the proof, for  $(\mathbf{s}^{(k)}, \hat{\mathbf{s}}^{(k)}) \in \mathcal{R}^{2k}$ , define

$$f_{k,n} \left( \mathbf{s}^{(k)}, \hat{\mathbf{s}}^{(k)} \right) \triangleq \frac{p_{S_n^{(k)} | \hat{S}_n^{(k)}} \left( \mathbf{s}^{(k)} | \hat{\mathbf{s}}^{(k)} \right)}{p_{S_n^{(k)}} \left( \mathbf{s}^{(k)} \right)}, \quad f_{k,\epsilon} \left( \mathbf{s}^{(k)}, \hat{\mathbf{s}}^{(k)} \right) \triangleq \frac{p_{S_\epsilon^{(k)} | \hat{S}_\epsilon^{(k)}} \left( \mathbf{s}^{(k)} | \hat{\mathbf{s}}^{(k)} \right)}{p_{S_\epsilon^{(k)}} \left( \mathbf{s}^{(k)} \right)}. \tag{D.6}$$

Now, we recall the backward channel relationship (20):

$$\mathbf{S}_n^{(k)} = \hat{\mathbf{S}}_n^{(k)} + \mathbf{W}_n^{(k)}, \tag{D.7}$$

where  $\hat{\mathbf{S}}_n^{(k)}$  and  $\mathbf{W}_n^{(k)}$  are mutually independent zero-mean, Gaussian random vectors with independent entries, corresponding to the optimal compression process and its respective distortion. From this relationship we obtain

$$\begin{aligned} p_{S_n^{(k)} | \hat{S}_n^{(k)}} \left( \mathbf{s}^{(k)} | \hat{\mathbf{s}}^{(k)} \right) &\stackrel{(a)}{=} p_{\hat{\mathbf{S}}_n^{(k)} + \mathbf{W}_n^{(k)} | \hat{S}_n^{(k)}} \left( \mathbf{s}^{(k)} | \hat{\mathbf{s}}^{(k)} \right) \\ &= p_{\mathbf{W}_n^{(k)} | \hat{S}_n^{(k)}} \left( \mathbf{s}^{(k)} - \hat{\mathbf{s}}^{(k)} | \hat{\mathbf{s}}^{(k)} \right) \\ &\stackrel{(b)}{=} p_{\mathbf{W}_n^{(k)}} \left( \mathbf{s}^{(k)} - \hat{\mathbf{s}}^{(k)} \right), \end{aligned} \tag{D.8}$$

where (a) follows since  $\mathbf{S}_n^{(k)} = \hat{\mathbf{S}}_n^{(k)} + \mathbf{W}_n^{(k)}$ , see (D.7), and (b) follows since  $\mathbf{W}_n^{(k)}$  and  $\hat{\mathbf{S}}_n^{(k)}$  are mutually independent. The joint PDF of  $\mathbf{S}_n^{(k)}$  and  $\hat{\mathbf{S}}_n^{(k)}$  can be expressed via the conditional PDF as:

$$p_{S_n^{(k)}, \hat{S}_n^{(k)}} \left( \mathbf{s}^{(k)}, \hat{\mathbf{s}}^{(k)} \right) = p_{S_n^{(k)} | \hat{S}_n^{(k)}} \left( \mathbf{s}^{(k)} | \hat{\mathbf{s}}^{(k)} \right) \cdot p_{\hat{S}_n^{(k)}} \left( \hat{\mathbf{s}}^{(k)} \right) \stackrel{(a)}{=} p_{\mathbf{W}_n^{(k)}} \left( \mathbf{s}^{(k)} - \hat{\mathbf{s}}^{(k)} \right) \cdot p_{\hat{S}_n^{(k)}} \left( \hat{\mathbf{s}}^{(k)} \right), \tag{D.9}$$

where (a) follows from (D.8). Since  $\hat{\mathbf{S}}_n^{(k)}$  and  $\mathbf{W}_n^{(k)}$  are Gaussian and mutually independent and since the product of two multivariate Gaussian PDFs is also a multivariate Gaussian PDF ([36]

Section 3), it follows from (D.9) that  $\mathbf{S}_n^{(k)}$  and  $\hat{\mathbf{S}}_n^{(k)}$  are jointly Gaussian. Following the mutual independence of  $\mathbf{W}_n^{(k)}$  and  $\hat{\mathbf{S}}_n^{(k)}$ , the right hand side (RHS) of (D.9) is also equivalent to the joint PDF of  $\left[ \left( \mathbf{W}_n^{(k)} \right)^T, \left( \hat{\mathbf{S}}_n^{(k)} \right)^T \right]^T$  denoted by  $p_{\mathbf{W}_n^{(k)}, \hat{\mathbf{S}}_n^{(k)}} \left( \mathbf{s}^{(k)} - \hat{\mathbf{s}}^{(k)}, \hat{\mathbf{s}}^{(k)} \right)$ . Now, from (D.8), the assumption  $\lim_{n \rightarrow \infty} p_{\mathbf{W}_n^{(k)}} \left( \mathbf{w}^{(k)} \right) = p_{\mathbf{W}_\epsilon^{(k)}} \left( \mathbf{w}^{(k)} \right)$  implies that a limit exists for the conditional PDF  $p_{\mathbf{S}_n^{(k)} | \hat{\mathbf{S}}_n^{(k)}} \left( \mathbf{s}^{(k)} | \hat{\mathbf{s}}^{(k)} \right)$ , this we denote by  $p_{\mathbf{S}_\epsilon^{(k)} | \hat{\mathbf{S}}_\epsilon^{(k)}} \left( \mathbf{s}^{(k)} | \hat{\mathbf{s}}^{(k)} \right)$ . Combining this with the assumption  $\lim_{n \rightarrow \infty} p_{\hat{\mathbf{S}}_n^{(k)}} \left( \hat{\mathbf{s}}^{(k)} \right) = p_{\hat{\mathbf{S}}_\epsilon^{(k)}} \left( \hat{\mathbf{s}}^{(k)} \right)$ , we have that,

$$\begin{aligned} \lim_{n \rightarrow \infty} p_{\mathbf{S}_n^{(k)}, \hat{\mathbf{S}}_n^{(k)}} \left( \mathbf{s}^{(k)}, \hat{\mathbf{s}}^{(k)} \right) &= \lim_{n \rightarrow \infty} \left( p_{\mathbf{S}_n^{(k)} | \hat{\mathbf{S}}_n^{(k)}} \left( \mathbf{s}^{(k)} | \hat{\mathbf{s}}^{(k)} \right) \cdot p_{\hat{\mathbf{S}}_n^{(k)}} \left( \hat{\mathbf{s}}^{(k)} \right) \right) \\ &\stackrel{(a)}{=} \lim_{n \rightarrow \infty} \left( p_{\mathbf{W}_n^{(k)}} \left( \mathbf{s}^{(k)} - \hat{\mathbf{s}}^{(k)} \right) \cdot p_{\hat{\mathbf{S}}_n^{(k)}} \left( \hat{\mathbf{s}}^{(k)} \right) \right) \\ &\stackrel{(b)}{=} \lim_{n \rightarrow \infty} \left( p_{\mathbf{W}_n^{(k)}} \left( \mathbf{s}^{(k)} - \hat{\mathbf{s}}^{(k)} \right) \right) \cdot \lim_{n \rightarrow \infty} \left( p_{\hat{\mathbf{S}}_n^{(k)}} \left( \hat{\mathbf{s}}^{(k)} \right) \right) \\ &= p_{\mathbf{S}_\epsilon^{(k)} | \hat{\mathbf{S}}_\epsilon^{(k)}} \left( \mathbf{s}^{(k)} | \hat{\mathbf{s}}^{(k)} \right) \cdot p_{\hat{\mathbf{S}}_\epsilon^{(k)}} \left( \hat{\mathbf{s}}^{(k)} \right) \\ &= p_{\mathbf{S}_\epsilon^{(k)}, \hat{\mathbf{S}}_\epsilon^{(k)}} \left( \mathbf{s}^{(k)}, \hat{\mathbf{s}}^{(k)} \right), \end{aligned} \tag{D.10}$$

where (a) follows from (D.8), and (b) follows since the limit for each sequence in the product exists ([31] Theorem 3.3); Convergence is uniform in  $\left( \left( \hat{\mathbf{s}}^{(k)} \right)^T, \left( \mathbf{s}^{(k)} \right)^T \right)^T \in \mathcal{R}^{2k}$  and  $k \in \mathcal{N}$ , as each sequence converges uniformly in  $k \in \mathcal{N}$  ([31] Page 165). Observe that the joint PDF for the zero-mean Gaussian random vectors  $\left[ \mathbf{S}_n^{(k)}, \hat{\mathbf{S}}_n^{(k)} \right]$  is given by the general expression:

$$p_{\mathbf{S}_n^{(k)}, \hat{\mathbf{S}}_n^{(k)}} \left( \mathbf{s}^{(k)}, \hat{\mathbf{s}}^{(k)} \right) = \left( \text{Det} \left( 2\pi \tilde{\mathbf{C}}_n^{(2k)} \right) \right)^{-\frac{1}{2}} \exp \left( -\frac{1}{2} \left[ \left( \hat{\mathbf{s}}^{(k)} \right)^T, \left( \mathbf{s}^{(k)} \right)^T \right] \left( \tilde{\mathbf{C}}_n^{(2k)} \right)^{-1} \left[ \left( \hat{\mathbf{s}}^{(k)} \right)^T, \left( \mathbf{s}^{(k)} \right)^T \right]^T \right), \tag{D.11}$$

where  $\tilde{\mathbf{C}}_n^{(2k)}$  denotes the joint covariance matrix of  $\left[ \left( \hat{\mathbf{S}}_n^{(k)} \right)^T, \left( \mathbf{S}_n^{(k)} \right)^T \right]^T$ . From (D.11) we note that  $p_{\mathbf{S}_n^{(k)}, \hat{\mathbf{S}}_n^{(k)}} \left( \mathbf{s}^{(k)}, \hat{\mathbf{s}}^{(k)} \right)$  is a continuous mapping of  $\tilde{\mathbf{C}}_n^{(2k)}$  with respect to the index  $n$ , see ([17] Lemma B.1). Hence the convergence in (D.10) of  $p_{\mathbf{S}_n^{(k)}, \hat{\mathbf{S}}_n^{(k)}} \left( \mathbf{s}^{(k)}, \hat{\mathbf{s}}^{(k)} \right)$  as  $n \rightarrow \infty$  directly implies the convergence of  $\tilde{\mathbf{C}}_n^{(2k)}$  as  $n \rightarrow \infty$  to a limit which we denote by  $\tilde{\mathbf{C}}_\epsilon^{(2k)}$ . It therefore follows that the limit function  $p_{\mathbf{S}_\epsilon^{(k)}, \hat{\mathbf{S}}_\epsilon^{(k)}} \left( \mathbf{s}^{(k)}, \hat{\mathbf{s}}^{(k)} \right)$  corresponds to the PDF of a Gaussian vector with the covariance matrix  $\tilde{\mathbf{C}}_\epsilon^{(2k)}$ .

The joint PDF for the zero-mean Gaussian random vectors  $\left[ \mathbf{W}_n^{(k)}, \hat{\mathbf{S}}_n^{(k)} \right]$  can be obtained using their mutual independence as:

$$\begin{aligned} &p_{\mathbf{W}_n^{(k)}, \hat{\mathbf{S}}_n^{(k)}} \left( \mathbf{s}^{(k)} - \hat{\mathbf{s}}^{(k)}, \hat{\mathbf{s}}^{(k)} \right) \\ &= \left( \text{Det} \left( 2\pi \Sigma_n^{(2k)} \right) \right)^{\frac{1}{2}} \exp \left( -\frac{1}{2} \left[ \left( \mathbf{s}^{(k)} - \hat{\mathbf{s}}^{(k)} \right)^T, \left( \hat{\mathbf{s}}^{(k)} \right)^T \right] \left( \Sigma_n^{(2k)} \right)^{-1} \left[ \left( \mathbf{s}^{(k)} - \hat{\mathbf{s}}^{(k)} \right)^T, \left( \hat{\mathbf{s}}^{(k)} \right)^T \right]^T \right), \end{aligned} \tag{D.12}$$

where  $\Sigma_n^{(2k)}$  denotes the joint covariance matrix of  $\left[ \left( \mathbf{W}_n^{(k)} \right)^T, \left( \hat{\mathbf{S}}_n^{(k)} \right)^T \right]^T$ . Since the vectors  $\mathbf{W}_n^{(k)}$  and  $\hat{\mathbf{S}}_n^{(k)}$  are zero-mean, mutually independent and, by the relationship (20), each vector has independent

entries, it follows that  $\Sigma_n^{(2k)}$  is a diagonal matrix with each diagonal element taking the value of the corresponding temporal variance at the respective index  $i \in \{1, 2, \dots, k\}$ . i.e.,

$$\begin{aligned} \Sigma_n^{(2k)} &\triangleq \mathbb{E} \left\{ \left( (\mathbf{W}_n^{(k)})^T, (\hat{\mathbf{S}}_n^{(k)})^T \right) \cdot \left( (\mathbf{W}_n^{(k)})^T, (\hat{\mathbf{S}}_n^{(k)})^T \right) \right\} \\ &= \text{diag}(\mathbb{E} \{ (W_n[1])^2 \}, \mathbb{E} \{ (W_n[2])^2 \}, \dots, \mathbb{E} \{ (W_n[k])^2 \}, \sigma_{\hat{S}_n}^2[1], \sigma_{\hat{S}_n}^2[2] \dots, \sigma_{\hat{S}_n}^2[k]). \end{aligned} \quad (\text{D.13})$$

The convergence of  $p_{\mathbf{W}_n^{(k)}, \hat{\mathbf{S}}_n^{(k)}}(\mathbf{s}^{(k)} - \hat{\mathbf{s}}^{(k)}, \hat{\mathbf{s}}^{(k)})$ , from (D.10), implies a convergence of the diagonal elements in (D.13) as  $n \rightarrow \infty$ . Hence  $\Sigma_n^{(2k)}$  converges as  $n \rightarrow \infty$  to a diagonal joint covariance matrix which we denote by  $\Sigma_\epsilon^{(2k)}$ . This further implies that the limiting vectors  $\mathbf{W}_\epsilon^{(k)}$  and  $\hat{\mathbf{S}}_\epsilon^{(k)}$  are zero-mean, mutually independent and each vector has independent entries in  $i \in [1, 2, \dots, k]$ .

Relationship (D.10) implies that the joint limit distribution satisfies  $p_{\mathbf{S}_\epsilon^{(k)}, \hat{\mathbf{S}}_\epsilon^{(k)}}(\mathbf{s}^{(k)}, \hat{\mathbf{s}}^{(k)}) = p_{\hat{\mathbf{S}}_\epsilon^{(k)}}(\hat{\mathbf{s}}^{(k)}) p_{\mathbf{W}_\epsilon^{(k)}}(\mathbf{s}^{(k)} - \hat{\mathbf{s}}^{(k)})$ . Consequently, we can define an asymptotic backward channel that satisfies (D.10) via the expression:

$$\mathbf{S}_\epsilon^{(k)}[i] = \hat{\mathbf{S}}_\epsilon^{(k)}[i] + \mathbf{W}_\epsilon^{(k)}[i]. \quad (\text{D.14})$$

Next, by convergence of the joint PDF  $p_{\mathbf{W}_n^{(k)}}(\mathbf{s}^{(k)} - \hat{\mathbf{s}}^{(k)}) \cdot p_{\hat{\mathbf{S}}_n^{(k)}}(\hat{\mathbf{s}}^{(k)})$  uniformly in  $k \in \mathcal{N}$  and in  $\left( (\mathbf{s}^{(k)})^T, (\hat{\mathbf{s}}^{(k)})^T \right)^T \in \mathcal{R}^{2k}$ , it follows from ([35] Theorem 1) that  $\left[ (\hat{\mathbf{S}}_n^{(k)})^T, (\mathbf{W}_n^{(k)})^T \right]^T \xrightarrow[n \rightarrow \infty]{(dist.)} \left[ (\hat{\mathbf{S}}_\epsilon^{(k)})^T, (\mathbf{W}_\epsilon^{(k)})^T \right]^T$  and the convergence is uniform in  $k \in \mathcal{N}$  and in  $\left( (\mathbf{s}^{(k)})^T, (\hat{\mathbf{s}}^{(k)})^T \right)^T \in \mathcal{R}^{2k}$ . Then, by the continuous mapping theorem (CMT) ([37] Theorem 7.7), we have

$$\left[ (\mathbf{S}_n^{(k)})^T, (\hat{\mathbf{S}}_n^{(k)})^T \right]^T = \left[ (\hat{\mathbf{S}}_n^{(k)} + \mathbf{W}_n^{(k)})^T, (\hat{\mathbf{S}}_n^{(k)})^T \right]^T \xrightarrow[n \rightarrow \infty]{(dist.)} \left[ (\hat{\mathbf{S}}_\epsilon^{(k)} + \mathbf{W}_\epsilon^{(k)})^T, (\hat{\mathbf{S}}_\epsilon^{(k)})^T \right]^T = \left[ (\mathbf{S}_\epsilon^{(k)})^T, (\hat{\mathbf{S}}_\epsilon^{(k)})^T \right]^T.$$

Now, using the extended CMT ([37] Theorem 7.24), we will show that  $f_{k,n}(\mathbf{S}_n^{(k)}, \hat{\mathbf{S}}_n^{(k)}) \xrightarrow[n \rightarrow \infty]{(dist.)} f_{k,\epsilon}(\mathbf{S}_\epsilon^{(k)}, \hat{\mathbf{S}}_\epsilon^{(k)})$  for each  $k \in \mathcal{N}$ , following the same approach as in the proof of ([17] Lemma B.2). Then, since  $\tilde{Z}'_{k,n}(F_{\mathbf{S}_n, \hat{\mathbf{S}}_n}^{\text{opt}}) = \frac{1}{k} \log f_{k,n}(\mathbf{S}_n^{(k)}, \hat{\mathbf{S}}_n^{(k)})$  and  $Z'_k(F_{\mathbf{S}_\epsilon, \hat{\mathbf{S}}_\epsilon}) = \frac{1}{k} \log f_{k,\epsilon}(\mathbf{S}_\epsilon^{(k)}, \hat{\mathbf{S}}_\epsilon^{(k)})$ , we conclude that  $\tilde{Z}'_{k,n}(F_{\mathbf{S}_n, \hat{\mathbf{S}}_n}^{\text{opt}}) \xrightarrow[n \rightarrow \infty]{(dist.)} Z'_k(F_{\mathbf{S}_\epsilon, \hat{\mathbf{S}}_\epsilon})$ , where it also follows from the proof of ([17] Lemma B.2) that the convergence is uniform in  $k \in \mathcal{N}$ . Specifically, to prove that  $f_{k,n}(\mathbf{S}_n^{(k)}, \hat{\mathbf{S}}_n^{(k)}) \xrightarrow[n \rightarrow \infty]{(dist.)} f_{k,\epsilon}(\mathbf{S}_\epsilon^{(k)}, \hat{\mathbf{S}}_\epsilon^{(k)})$ , we will show that the following two properties hold:

- P1 The distribution of  $\left[ (\mathbf{S}_\epsilon^{(k)})^T, (\hat{\mathbf{S}}_\epsilon^{(k)})^T \right]^T$  is separable (as defined in ([37] Pg. 101)).
- P2 For any convergent sequence  $\left( (\mathbf{s}_n^{(k)})^T, (\hat{\mathbf{s}}_n^{(k)})^T \right)^T \in \mathcal{R}^{2k}$  such that  $\lim_{n \rightarrow \infty} (\mathbf{s}_n^{(k)}, \hat{\mathbf{s}}_n^{(k)}) = (\mathbf{s}_\epsilon^{(k)}, \hat{\mathbf{s}}_\epsilon^{(k)})$ , then  $\lim_{n \rightarrow \infty} f_{k,n}(\mathbf{s}_n^{(k)}, \hat{\mathbf{s}}_n^{(k)}) = f_{k,\epsilon}(\mathbf{s}_\epsilon^{(k)}, \hat{\mathbf{s}}_\epsilon^{(k)})$ .

To prove property P1, we show that  $U^{(k)} \triangleq \left[ (\mathbf{S}_\epsilon^{(k)})^T, (\hat{\mathbf{S}}_\epsilon^{(k)})^T \right]^T$  (here we misuse the dimension notation as  $U^{(k)}$  denotes a  $2k$ -dimensional vector) is separable ([37] Pg. 101), i.e., we show that  $\forall \eta > 0$ , there exists  $\beta > 0$  such that  $\Pr(\|U^{(k)}\|^2 > \beta) < \eta$ . To that aim, recall first that by Markov's inequality ([29] Pg. 114), it follows that  $\Pr(\|U^{(k)}\|^2 > \beta) < \frac{1}{\beta} \mathbb{E} \left\{ \|U^{(k)}\|^2 \right\}$ . For the

asynchronously sampled source process, we note that  $\sigma_{S_\epsilon}^2[i] \triangleq \mathbb{E} \left\{ (S_\epsilon[i])^2 \right\} \in [0, \max_{0 \leq t \leq T_{ps}} \sigma_{S_\epsilon}^2(t)]$ . By the independence of  $W_\epsilon^{(k)}$  and  $\hat{S}_\epsilon^{(k)}$ , and by the fact that their mean is zero, we have, from (D.14) that  $\mathbb{E} \left\{ (S_\epsilon[i])^2 \right\} = \mathbb{E} \left\{ (\hat{S}_\epsilon[i])^2 \right\} + \mathbb{E} \left\{ (W_\epsilon[i])^2 \right\} \leq \max_{0 \leq t \leq T_{ps}} \sigma_{S_\epsilon}^2(t)$ ; Hence  $\mathbb{E} \left\{ (\hat{S}_\epsilon[i])^2 \right\} \leq \max_{0 \leq t \leq T_{ps}} \sigma_{S_\epsilon}^2(t)$ , and  $\mathbb{E} \left\{ (W_\epsilon[i])^2 \right\} \leq \max_{0 \leq t \leq T_{ps}} \sigma_{S_\epsilon}^2(t)$ . This further implies that  $\mathbb{E} \left\{ \left\| U^{(k)} \right\|^2 \right\} = \mathbb{E} \left\{ \left\| \left[ (S_\epsilon^{(k)})^T, (\hat{S}_\epsilon^{(k)})^T \right]^T \right\|^2 \right\} \leq 2 \cdot k \cdot \max_{0 \leq t \leq T_{ps}} \sigma_{S_\epsilon}^2(t)$ ; therefore for each  $\beta > \frac{1}{\eta} \mathbb{E} \left\{ \left\| U^{(k)} \right\|^2 \right\}$  we have that  $\Pr \left( \left\| U^{(k)} \right\|^2 > \beta \right) < \eta$ , and thus  $U^{(k)}$  is separable.

By the assumption in this lemma it follows that  $\forall \eta > 0$  there exists  $n_0(\eta) > 0$  such that for all  $n > n_0(\eta)$  we have that  $\forall w^{(k)} \in \mathcal{R}^k, |p_{W_n^{(k)}}(w^{(k)}) - p_{W_\epsilon^{(k)}}(w^{(k)})| < \eta$ , for all sufficiently large  $k \in \mathcal{N}$ . Consequently, for all  $\left( (s^{(k)})^T, (\hat{s}^{(k)})^T \right)^T \in \mathcal{R}^{2k}, n > n_0(\eta)$  and a sufficiently large  $k \in \mathcal{N}$ , it follows from (D.8) that

$$|p_{S_n^{(k)}|\hat{S}_n^{(k)}}(s^{(k)}|\hat{s}^{(k)}) - p_{S_\epsilon^{(k)}|\hat{S}_\epsilon^{(k)}}(s^{(k)}|\hat{s}^{(k)})| = |p_{W_n^{(k)}}(s^{(k)} - \hat{s}^{(k)}) - p_{W_\epsilon^{(k)}}(s^{(k)} - \hat{s}^{(k)})| < \eta. \tag{D.15}$$

Following the continuity of  $p_{S_n^{(k)}|\hat{S}_n^{(k)}}(s^{(k)}|\hat{s}^{(k)})$  and of  $p_{S_n^{(k)}}(s^{(k)}), f_{k,n}(s^{(k)}, \hat{s}^{(k)})$  is also continuous ([31] Theorem 4.9); hence, when  $\lim_{n \rightarrow \infty} (s_n^{(k)}, \hat{s}_n^{(k)}) = (s^{(k)}, \hat{s}^{(k)})$ , then  $\lim_{n \rightarrow \infty} f_{k,n}(s_n^{(k)}, \hat{s}_n^{(k)}) = f_{k,\epsilon}(s^{(k)}, \hat{s}^{(k)})$ . This satisfies condition P2 for the extended CMT; Therefore, by the extended CMT, we have that  $f_{k,n}(s_n^{(k)}, \hat{s}_n^{(k)}) \xrightarrow[n \rightarrow \infty]{(dist.)} f_{k,\epsilon}(s^{(k)}, \hat{s}^{(k)})$ . Since the RVs  $\tilde{Z}'_{k,n}(F_{S_n, \hat{S}_n}^{opt})$  and  $Z'_{k,\epsilon}(F_{S_\epsilon, \hat{S}_\epsilon})$ , defined in (D.1), are also continuous mappings of  $f_{k,n}(s_n^{(k)}, \hat{s}_n^{(k)})$  and of  $f_{k,\epsilon}(s^{(k)}, \hat{s}^{(k)})$ , respectively, it follows from the CMT ([37] Theorem 7.7) that  $\tilde{Z}'_{k,n}(F_{S_n, \hat{S}_n}^{opt}) \xrightarrow[n \rightarrow \infty]{(dist.)} Z'_{k,\epsilon}(F_{S_\epsilon, \hat{S}_\epsilon})$ .

Finally, to prove that the convergence  $\tilde{Z}'_{k,n}(F_{S_n, \hat{S}_n}^{opt}) \xrightarrow[n \rightarrow \infty]{(dist.)} Z'_{k,\epsilon}(F_{S_\epsilon, \hat{S}_\epsilon})$  is uniform in  $k \in \mathcal{N}$ , we note that as  $\hat{S}_n^{(k)}$  and  $\hat{S}_\epsilon^{(k)}$  have independent entries, and the backward channels (21) and (D.14) are memoryless. Hence, it follows from the proof of ([17] Lemma B.2), that the characteristic function of the RV  $k \cdot \tilde{Z}'_{k,n}(F_{S_n, \hat{S}_n}^{opt})$  which is denoted by  $\Phi_{k \cdot \tilde{Z}_{k,n}}(\alpha) \triangleq \mathbb{E} \left\{ e^{j \cdot \alpha \cdot k \cdot \tilde{Z}_{k,n}} \right\}$  converges to the characteristic function of  $k \cdot Z'_{k,\epsilon}(F_{S_\epsilon, \hat{S}_\epsilon})$ , denoted by  $\Phi_{k \cdot Z_{k,\epsilon}}(\alpha)$ , uniformly over  $k \in \mathcal{N}$ . Thus, for all sufficiently small  $\eta > 0, \exists k_0 \in \mathcal{N}, n_0(\eta, k_0) \in \mathcal{N}$  such that  $\forall n > n_0(\eta, k_0)$ , and  $\forall k > k_0$

$$|\Phi_{k \cdot \tilde{Z}_{k,n}}(\alpha) - \Phi_{k \cdot Z_{k,\epsilon}}(\alpha)| < \eta, \quad \forall \alpha \in \mathcal{R}. \tag{D.16}$$

Hence, following Lévy's convergence theorem ([38] Theorem 18.1) we conclude that  $k \cdot \tilde{Z}'_{k,n}(F_{S_n, \hat{S}_n}^{opt}) \xrightarrow[n \rightarrow \infty]{(dist.)} k \cdot Z'_{k,\epsilon}(F_{S_\epsilon, \hat{S}_\epsilon})$  and that this convergence is uniform for sufficiently large  $k$ . Finally, since the CDFs of  $k \cdot \tilde{Z}'_{k,n}(F_{S_n, \hat{S}_n}^{opt})$  and  $k \cdot Z'_{k,\epsilon}(F_{S_\epsilon, \hat{S}_\epsilon})$  obtained at  $\alpha \in \mathcal{R}$  are equivalent to the CDFs of  $\tilde{Z}'_{k,n}(F_{S_n, \hat{S}_n}^{opt})$  and  $Z'_{k,\epsilon}(F_{S_\epsilon, \hat{S}_\epsilon})$  obtained at  $\frac{\alpha}{k} \in \mathcal{R}$  respectively, we can conclude that  $\tilde{Z}'_{k,n}(F_{S_n, \hat{S}_n}^{opt}) \xrightarrow[n \rightarrow \infty]{(dist.)} Z'_{k,\epsilon}(F_{S_\epsilon, \hat{S}_\epsilon})$ , uniformly in  $k \in \mathcal{N}$ .  $\square$

The following convergence lemma D.3 corresponds to ([17] Lemma B.3),

**Lemma D.3.** *Let  $n \in \mathcal{N}$  be given. Every subsequence of  $\left\{ \tilde{Z}'_{k_l, n}(F_{S_n, \hat{S}_n}^{opt}) \right\}_{k_l \in \mathcal{N}}$ , indexed by  $k_l$ , converges in distribution, in the limit as  $l \rightarrow \infty$ , to a finite deterministic scalar.*

**Proof.** Recall that the RVs  $\tilde{Z}'_{k,n} \left( F_{\hat{S}_n, S_n}^{\text{opt}} \right)$  represent the mutual information density rate between  $k$  samples of the source process  $S_n[i]$  and the corresponding samples of its reproduction process  $\hat{S}_n[i]$ , where these processes are jointly distributed via the Gaussian distribution measure  $F_{\hat{S}_n, S_n}^{\text{opt}}$ . Further, recall that the relationship between the source signal and the reproduction process which achieves the RDF can be described via the backward channel in (21) for a Gaussian source. The channel (21) is a memoryless additive WSCS Gaussian noise channel with period  $p_n$ , thus, by [21], it can be equivalently represented as a  $p_n \times 1$  multivariate memoryless additive stationary Gaussian noise channel, which is an information stable channel ([39] Section 1.5). For such channels in which the source and its reproduction obey the RDF-achieving joint distribution  $F_{\hat{S}_n, S_n}^{\text{opt}}$ , the mutual information density rate converges as  $k$  increases, almost surely, to the finite and deterministic mutual information rate ([14] Theorem 5.9.1). Since almost sure convergence implies convergence in distribution ([37] Lemma 7.21), this proves the lemma.  $\square$

Appendix D.3. Showing that  $R_\epsilon(D) = \limsup_{n \rightarrow \infty} R_n(D)$

This section completes the proof to Theorem 4. We note from (14) that the RDF for the source process  $S_n[i]$  (for fixed length coding and MSE distortion measure) is given by:

$$R_n(D) = \inf_{F_{\hat{S}_n, S_n}: \bar{d}_S(F_{\hat{S}_n, S_n}) \leq D} \left\{ \text{p-} \limsup_{k \rightarrow \infty} \tilde{Z}'_{k,n} \left( F_{\hat{S}_n, S_n}^{\text{opt}} \right) \right\}, \tag{D.17}$$

where  $\bar{d}_S \left( F_{\hat{S}_n, S_n} \right) = \limsup_{k \rightarrow \infty} \frac{1}{k} \mathbb{E} \left\{ \left\| \mathbf{S}_n^{(k)} - \hat{\mathbf{S}}_n^{(k)} \right\|^2 \right\}$ .

We now state the following lemma characterizing the asymptotic statistics of the optimal reconstruction  $\hat{\mathbf{S}}_n^{(k)}$  process and the respective noise process  $\mathbf{W}_n^{(k)}$  used in the backward channel relationship (21):

**Lemma D.4.** Consider the RDF-achieving distribution with distortion  $D$  for compression of a vector Gaussian source process  $\mathbf{S}_n^{(k)}$  characterized by the backward channel (21). Then, there exists a subsequence in the index  $n \in \mathcal{N}$  denoted  $n_1 < n_2 < \dots$ , such that for the RDF-achieving distribution, the sequences of reproduction vectors  $\{\hat{\mathbf{S}}_{n_l}^{(k)}\}_{l \in \mathcal{N}}$  and backward channel noise vectors  $\{\mathbf{W}_{n_l}^{(k)}\}_{l \in \mathcal{N}}$  satisfy that  $\lim_{l \rightarrow \infty} p_{\hat{\mathbf{S}}_{n_l}^{(k)}} \left( \hat{\mathbf{s}}^{(k)} \right) = p_{\hat{\mathbf{S}}_\epsilon^{(k)}} \left( \hat{\mathbf{s}}^{(k)} \right)$  uniformly in  $\hat{\mathbf{s}}^{(k)} \in \mathcal{R}^k$  and uniformly with respect to  $k \in \mathcal{N}$ , as well as  $\lim_{l \rightarrow \infty} p_{\mathbf{W}_{n_l}^{(k)}} \left( \mathbf{w}^{(k)} \right) = p_{\mathbf{W}_\epsilon^{(k)}} \left( \mathbf{w}^{(k)} \right)$  uniformly in  $\mathbf{w}^{(k)} \in \mathcal{R}^k$  and uniformly with respect to  $k \in \mathcal{N}$ , where  $p_{\hat{\mathbf{S}}_\epsilon^{(k)}} \left( \hat{\mathbf{s}}^{(k)} \right)$  and  $p_{\mathbf{W}_\epsilon^{(k)}} \left( \mathbf{w}^{(k)} \right)$  are Gaussian PDFs.

**Proof.** Recall from the analysis of the RDF for WSCS processes that for each  $n \in \mathcal{N}$ , the marginal distributions of the RDF-achieving reproduction process  $\hat{S}_n[i]$  and the backward channel noise  $W_n[i]$  is Gaussian, memoryless, zero-mean, and with variances  $\sigma_{\hat{S}_n}^2 [i] \triangleq \mathbb{E} \left\{ \left( \hat{S}_n[i] \right)^2 \right\}$  and

$$\mathbb{E} \left\{ \left( W_n[i] \right)^2 \right\} = \sigma_{S_n}^2 [i] - \sigma_{\hat{S}_n}^2 [i], \tag{D.18}$$

respectively. Consequently, the sequences of reproduction vectors  $\{\hat{\mathbf{S}}_n^{(k)}\}_{n \in \mathcal{N}}$  and backward channel noise vectors  $\{\mathbf{W}_n^{(k)}\}_{n \in \mathcal{N}}$  are zero-mean Gaussian with independent entries for each  $k \in \mathcal{N}$ . Since  $\sigma_{\hat{S}_n}^2 [i] \leq \max_{t \in \mathcal{R}} \sigma_{S_c}^2 (t)$ , then, from (D.18), it follows that  $\sigma_{\hat{S}_n}^2 [i]$  is also bounded in the interval  $[0, \max_{t \in \mathcal{R}} \sigma_{S_c}^2 (t)]$  for all  $n \in \mathcal{N}$ . Therefore, by Bolzano-Weierstrass theorem ([31] Theorem 2.42),  $\sigma_{\hat{S}_n}^2 [i]$  has a convergent subsequence, and we let  $n_1 < n_2 < \dots$  denote the indexes of this convergent subsequence and let the limit of the subsequence be denoted by  $\sigma_{\hat{S}_\epsilon}^2 [i]$ . From the CMT, as applied in the proof of ([17] Lemma B.1), the convergence  $\sigma_{\hat{S}_{n_l}}^2 [i] \xrightarrow{l \rightarrow \infty} \sigma_{\hat{S}_\epsilon}^2 [i]$  for each  $i \in \mathcal{N}$  implies that the subsequence of PDFs

$p_{\hat{S}_{n_l}^{(k)}}(\hat{s}^{(k)})$  corresponding to the memoryless Gaussian random vectors  $\{\hat{S}_{n_l}^{(k)}\}_{l \in \mathcal{N}}$  converges as  $l \rightarrow \infty$  to a Gaussian PDF which we denote by  $p_{\hat{S}_\epsilon^{(k)}}(\hat{s}^{(k)})$ , and the convergence of  $p_{\hat{S}_{n_l}^{(k)}}(\hat{s}^{(k)})$  is uniform in  $s^{(k)}$  for any fixed  $k \in \mathcal{N}$ . By Remark 2, it holds that  $W_n[i]$  is a memoryless stationary process with variance  $\mathbb{E}\{(W_n[i])^2\} = D$  and by Equation (D.18),  $\sigma_{\hat{S}_n}^2[i] = \sigma_{S_n}^2[i] - D$ . Hence by Assumption D.1 and by the proof of ([17] Lemma B.1), it follows that for a fixed  $\eta > 0$  and  $k_0 \in \mathcal{N}$ ,  $\exists n_0(\eta, k_0)$  such that for all  $n > n_0(\eta, k_0)$  and for all sufficiently large  $k$ , it holds that  $|p_{\hat{S}_{n_l}^{(k)}}(\hat{s}^{(k)}) - p_{\hat{S}_\epsilon^{(k)}}(\hat{s}^{(k)})| < \eta$  for every  $\hat{s}^{(k)} \in \mathcal{R}^k$ . Since  $n_0(\eta, k_0)$  does not depend on  $k$  (only on the fixed  $k_0$ ), this implies that the convergence is uniform with respect to  $k \in \mathcal{N}$ .

The fact that  $W_n[i]$  is a zero-mean stationary Gaussian process with variance  $D$  for each  $n \in \mathcal{N}$ , implies that the sequence of PDFs  $p_{W_n^{(k)}}(w^{(k)})$  converges as  $n \rightarrow \infty$  to a Gaussian PDF which we denote by  $p_{W^{(k)}}(w^{(k)})$ , hence its subsequence with indices  $n_1 < n_2 < \dots$  also converges to  $p_{W^{(k)}}(w^{(k)})$ . Since  $D > \frac{1}{2\pi}$  by Assumption D.1 combined with the proof of ([17] Lemma B.1) it follows that this convergence is uniform in  $w^{(k)}$  and in  $k \in \mathcal{N}$  to  $p_{W^{(k)}}(w^{(k)})$ .

Following the proof of Corollary D.1, it holds that the subsequences of the memoryless Gaussian random vectors  $\{\hat{S}_{n_l}^{(k)}\}$  and  $\{W_{n_l}^{(k)}\}$  converge in distribution as  $l \rightarrow \infty$  to a Gaussian distribution, and the convergence is uniform in  $k \in \mathcal{N}$  for any fixed  $k \in \mathcal{N}$ . Hence, as shown in Lemma D.2 the joint distribution  $[(S_{n_l}^{(k)})^T, (\hat{S}_{n_l}^{(k)})^T]^T \xrightarrow[n \rightarrow \infty]{(dist.)} [(S_\epsilon^{(k)})^T, (\hat{S}_\epsilon^{(k)})^T]^T$ , and the limit distribution is jointly Gaussian.  $\square$

**Lemma D.5.** *The RDF of  $\{S_\epsilon[i]\}$  satisfies  $R_\epsilon(D) \leq \limsup_{n \rightarrow \infty} R_n(D)$ , and the rate  $\limsup_{n \rightarrow \infty} R_n(D)$  is achievable for the source  $\{S_\epsilon[i]\}$  with distortion  $D$  when the reproduction process which obeys a Gaussian distribution.*

**Proof.** According to Lemma D.4, we note that the sequence of joint distributions  $\{F_{S_n, \hat{S}_n}^{opt}\}_{n \in \mathcal{N}}$  has a convergent subsequence, i.e., there exists a set of indexes  $n_1 < n_2 < \dots$  such that the sequence of distributions with independent entries  $\{F_{S_{n_l}, \hat{S}_{n_l}}^{opt}\}_{l \in \mathcal{N}}$  converges in the limit  $l \rightarrow \infty$  to a joint Gaussian distribution  $F'_{S_\epsilon, \hat{S}_\epsilon}$  and the convergence is uniform in  $k \in \mathcal{N}$ . Hence, this satisfies the condition of Lemma D.2; This implies that  $\tilde{Z}'_{k, n_l}(F_{S_{n_l}, \hat{S}_{n_l}}^{opt}) \xrightarrow[l \rightarrow \infty]{(dist.)} Z'_k(F'_{S_\epsilon, \hat{S}_\epsilon})$  uniformly in  $k \in \mathcal{N}$ . Moreover, by Lemma D.3 every subsequence of  $\{\tilde{Z}'_{k, n_l}(F_{S_{n_l}, \hat{S}_{n_l}}^{opt})\}_{l \in \mathcal{N}}$  converges in distribution to a finite deterministic scalar as  $k \rightarrow \infty$ . Therefore, by Theorem 3 it holds that

$$\begin{aligned} \lim_{l \rightarrow \infty} \left( \text{p-lim sup}_{k \rightarrow \infty} \tilde{Z}'_{k, n_l} \left( F_{S_{n_l}, \hat{S}_{n_l}}^{opt} \right) \right) &= \text{p-lim sup}_{k \rightarrow \infty} Z'_{k, \epsilon} \left( F'_{S_\epsilon, \hat{S}_\epsilon} \right) \\ &\geq \inf_{F_{S_\epsilon, \hat{S}_\epsilon}} \left\{ \text{p-lim sup}_{k \rightarrow \infty} Z'_{k, \epsilon} \left( F_{S_\epsilon, \hat{S}_\epsilon} \right) \right\} = R_\epsilon(D). \end{aligned} \tag{D.19}$$

From (14) we have that  $R_n(D) = \text{p-lim sup}_{k \rightarrow \infty} \tilde{Z}'_{k, n} \left( F_{S_n, \hat{S}_n}^{opt} \right)$ , then from (D.19), it follows that

$$R_\epsilon(D) \leq \lim_{l \rightarrow \infty} R_{n_l}(D) \stackrel{(a)}{\leq} \limsup_{n \rightarrow \infty} R_n(D), \tag{D.20}$$

where (a) follows since, by ([31] Definition 3.16), the limit of every subsequence is not greater than the limit superior. Noting that  $F'_{S_\epsilon, \hat{S}_\epsilon}$  is Gaussian by Lemma D.4 concludes the proof.  $\square$

**Lemma D.6.** *The RDF of  $\{S_\epsilon[i]\}$  satisfies  $R_\epsilon(D) \geq \limsup_{n \rightarrow \infty} R_n(D)$ .*

**Proof.** To prove this lemma, we first show that for a joint distribution  $F_{S_\epsilon, \hat{S}_\epsilon}$  which achieves a rate-distortion pair  $(R_\epsilon, D)$  it holds that  $R_\epsilon \geq \mathbb{E}\{Z'_{k,\epsilon}(F'_{S_\epsilon, \hat{S}_\epsilon})\}$ : Recall that  $(R_\epsilon, D)$  is an achievable rate-distortion pair for the source  $\{S_\epsilon[i]\}$ , namely, there exists a sequence of codes  $\{C_l\}$  whose rate-distortion approach  $(R_\epsilon, D)$  when applied to  $\{S_\epsilon[i]\}$ , This implies that for any  $\eta > 0$  there exists  $l_0(\eta)$  such that  $\forall l > l_0(\eta)$  it holds that  $C_l$  has a code rate  $R_l = \frac{1}{l} \log_2 M_l$  satisfying  $R_l \leq R_\epsilon + \eta$  by (3). Recalling Definition 5, the source code maps  $S_\epsilon^{(l)}$  into a discrete index  $J_l \in \{1, 2, \dots, M_l\}$ , which is in turn mapped into  $\hat{S}_\epsilon^{(l)}$ , i.e.,  $S_\epsilon^{(l)} \mapsto J_l \mapsto \hat{S}_\epsilon^{(l)}$  form a Markov chain. Since  $J_l$  is a discrete random variable taking values in  $\{1, 2, \dots, M_l\}$ , it holds that

$$\begin{aligned} \log_2 M_l &\geq H(J_l) \\ &\stackrel{(a)}{\geq} I(S_\epsilon^{(l)}; J_l) \\ &\stackrel{(b)}{\geq} I(S_\epsilon^{(l)}; \hat{S}_\epsilon^{(l)}), \end{aligned} \tag{D.21}$$

where (a) follows since  $I(S_\epsilon^{(l)}; J_l) = H(J_l) - H(J_l|S_\epsilon^{(l)})$  which is not larger than  $H(J_l)$  as  $J_l$  takes discrete values; while (b) follows from the data processing inequality ([5] Chapter 2.8). Now, (D.21) implies that for each  $l > l_0(\eta)$ , the reproduction obtained using the code  $C_l$  satisfies  $\frac{1}{l} I(S_\epsilon^{(l)}; \hat{S}_\epsilon^{(l)}) \leq \frac{1}{l} \log M_l \leq R_\epsilon + \eta$ . Since for every arbitrarily small  $\eta \rightarrow 0$ , this inequality holds for all  $l > l_0(\eta)$ , i.e., for all sufficiently large  $l$ , it follows that  $R_\epsilon \geq \limsup_{k \rightarrow \infty} \frac{1}{k} I(S_\epsilon^{(k)}; \hat{S}_\epsilon^{(k)})$ . Hence, replacing the blocklength symbol from  $l$  to  $k$ , as  $\frac{1}{k} I(S_\epsilon^{(k)}; \hat{S}_\epsilon^{(k)}) = \mathbb{E}\{Z'_{k,\epsilon}(F'_{S_\epsilon, \hat{S}_\epsilon})\}$  ([5] Equation (2.3)), we conclude that

$$R_\epsilon(D) \geq \limsup_{k \rightarrow \infty} \mathbb{E}\{Z'_{k,\epsilon}(F'_{S_\epsilon, \hat{S}_\epsilon})\}. \tag{D.22}$$

Next, we consider  $\limsup_{k \rightarrow \infty} \mathbb{E}\{Z'_{k_1,\epsilon}(F'_{S_\epsilon, \hat{S}_\epsilon})\}$ : Let  $Z'_{k_1,\epsilon}(F'_{S_\epsilon, \hat{S}_\epsilon})$  be a subsequence of  $\mathbb{E}\{Z'_{k,\epsilon}(F'_{S_\epsilon, \hat{S}_\epsilon})\}$  with the indexes  $k_1 < k_2 < \dots$  such that its limit equals the limit superior. i.e.,  $\lim_{l \rightarrow \infty} \mathbb{E}\{Z'_{k_l,\epsilon}(F'_{S_\epsilon, \hat{S}_\epsilon})\} = \limsup_{k \rightarrow \infty} \mathbb{E}\{Z'_{k,\epsilon}(F'_{S_\epsilon, \hat{S}_\epsilon})\}$ . Since by Lemma D.2, the sequence of non-negative RVs  $\{Z'_{k_l,n}(F_{S_n, \hat{S}_n}^{\text{opt}})\}_{n \in \mathcal{N}}$  convergences in distribution to  $Z'_{k_l,\epsilon}(F'_{S_\epsilon, \hat{S}_\epsilon})$  as  $n \rightarrow \infty$  uniformly in  $k \in \mathcal{N}$ , it follows from ([40] Theorem 3.5) that  $\mathbb{E}\{Z'_{k_l,\epsilon}(F'_{S_\epsilon, \hat{S}_\epsilon})\} = \lim_{n \rightarrow \infty} \mathbb{E}\{Z'_{k_l,n}(F_{S_n, \hat{S}_n}^{\text{opt}})\}$ . Moreover, we define a family of distributions  $\mathcal{F}(D)$  such that  $\mathcal{F}(D) = \{F_{S, \hat{S}} : D(F_{S, \hat{S}}) \leq D\}$ . Consequently, Equation (D.22) can now be written as:

$$\begin{aligned} R_\epsilon(D) &\geq \limsup_{k \rightarrow \infty} \mathbb{E}\{Z'_{k,\epsilon}(F'_{S_\epsilon, \hat{S}_\epsilon})\} = \lim_{l \rightarrow \infty} \lim_{n \rightarrow \infty} \mathbb{E}\{Z'_{k_l,n}(F_{S_n, \hat{S}_n}^{\text{opt}})\} \\ &\stackrel{(a)}{=} \lim_{n \rightarrow \infty} \lim_{l \rightarrow \infty} \mathbb{E}\{Z'_{k_l,n}(F_{S_n, \hat{S}_n}^{\text{opt}})\} \\ &\stackrel{(b)}{=} \limsup_{n \rightarrow \infty} \lim_{l \rightarrow \infty} \mathbb{E}\{Z'_{k_l,n}(F_{S_n, \hat{S}_n}^{\text{opt}})\} \\ &\geq \limsup_{n \rightarrow \infty} \lim_{l \rightarrow \infty} \inf_{F_{S, \hat{S}} \in \mathcal{F}(D)} \mathbb{E}\{Z'_{k_l,n}(F_{S, \hat{S}})\} \\ &\stackrel{(c)}{=} \limsup_{n \rightarrow \infty} \lim_{l \rightarrow \infty} \inf_{F_{S, \hat{S}} \in \mathcal{F}(D)} \frac{1}{k_l} I(\hat{S}_n^{(k_l)}; S_n^{(k_l)}), \end{aligned} \tag{D.23}$$



where (a) follows since the convergence  $\tilde{Z}'_{k_l, n} \left( F_{S_n, \hat{S}_n}^{\text{opt}} \right) \xrightarrow[n \rightarrow \infty]{(dist.)} Z'_{k_l, \epsilon} \left( F'_{S_\epsilon, \hat{S}} \right)$  is uniform with respect to  $k_l$ , thus the limits are interchangeable ([31] Theorem 7.11); (b) follows since the limit of the subsequence  $\mathbb{E} \left\{ \tilde{Z}'_{k_l, n} \left( F_{S_n, \hat{S}_n}^{\text{opt}} \right) \right\}$  exists in the index  $n$ , and is therefore equivalent to the limit superior,  $\limsup_{n \rightarrow \infty} \mathbb{E} \left\{ \tilde{Z}'_{k_l, n} \left( F_{S_n, \hat{S}_n}^{\text{opt}} \right) \right\}$  ([31] Page 57); and (c) holds since mutual information is the expected value of the mutual information density rate ([5] Equation (2.30)). Finally, we recall that in the proof of Lemma D.3 it was established that the backward channel for the RDF at the distortion constraint  $D$ , defined in (21), is information stable, hence for such backward channels, we have from ([41] Theorem 1) that the minimum rate is defined as  $R_n(D) = \lim_{k \rightarrow \infty} \inf_{F_{S, \hat{S}} \in \mathcal{F}(D)} \frac{1}{k} I \left( \hat{S}_\epsilon^{(k)}; S_n^{(k)} \right)$  and the limit exists; Hence,  $\lim_{k \rightarrow \infty} \inf_{F_{S, \hat{S}} \in \mathcal{F}(D)} \frac{1}{k} I \left( \hat{S}_\epsilon^{(k)}; S_n^{(k)} \right) = \lim_{l \rightarrow \infty} \inf_{F_{S, \hat{S}} \in \mathcal{F}(D)} \frac{1}{k_l} I \left( \hat{S}_\epsilon^{(k_l)}; S_n^{(k_l)} \right)$  in the index  $k$ . Substituting this into Equation (D.23) yields the result:

$$R_\epsilon(D) \geq \limsup_{n \rightarrow \infty} R_n(D). \quad (\text{D.24})$$

This proves the lemma.  $\square$

Combining the Lemmas D.5 and D.6 proves that  $R_\epsilon(D) = \limsup_{n \rightarrow \infty} R_n(D)$  and the rate is achievable with Gaussian inputs, completing the proof of the theorem.

## References

- Gardner, W.; Brown, W.; Chen, C.K. Spectral correlation of modulated signals: Part II-digital modulation. *IEEE Trans. Commun.* **1987**, *35*, 595–601.
- Giannakis, G.B. Cyclostationary signal analysis. In *Digital Signal Processing Handbook*; CRC PRESS: Boca Raton, FL, USA, 1998; pp. 17–21.
- Gardner, W.A.; Napolitano, A.; Paura, L. Cyclostationarity: Half a century of research. *Signal Process.* **2006**, *86*, 639–697.
- Berger, T.; Gibson, J.D. Lossy source coding. *IEEE Trans. Inf. Theory* **1998**, *44*, 2693–2723.
- Cover, T.M.; Thomas, J.A. *Elements of Information Theory*; John Wiley & Sons: New York, NY, USA, 2006.
- Wolf, J.K.; Wyner, A.D.; Ziv, J. Source coding for multiple descriptions. *Bell Syst. Tech. J.* **1980**, *59*, 1417–1426.
- Wyner, A.D.; Ziv, J. The rate-distortion function for source coding with side information at the decoder. *IEEE Trans. Inf. Theory* **1976**, *22*, 1–10.
- Oohama, Y. Gaussian multiterminal source coding. *IEEE Trans. Inf. Theory* **1997**, *43*, 1912–1923.
- Pandya, A.; Kansal, A.; Pottie, G.; Srivastava, M. Lossy source coding of multiple Gaussian sources: m-helper problem. In Proceedings of the IEEE Information Theory Workshop, San Antonio, TX, USA, 24–29 October 2004; pp. 34–38.
- Gallager, R.G. *Information Theory and Reliable Communication*; Springer: Berlin, Germany, 1968; Volume 588.
- Harrison, M.T. The generalized asymptotic equipartition property: Necessary and sufficient conditions. *IEEE Trans. Inf. Theory* **2008**, *54*, 3211–3216.
- Kipnis, A.; Goldsmith, A.J.; Eldar, Y.C. The distortion rate function of cyclostationary Gaussian processes. *IEEE Trans. Inf. Theory* **2018**, *64*, 3810–3824.
- Napolitano, A. Cyclostationarity: New trends and applications. *Signal Process.* **2016**, *120*, 385–408.
- Han, T.S. *Information-Spectrum Methods in Information Theory*; Springer: Berlin, Germany, 2003; Volume 50.
- Verdú, S.; Han, T.S. A general formula for channel capacity. *IEEE Trans. Inf. Theory* **1994**, *40*, 1147–1157.
- Zeng, W.; Mitran, P.; Kavcic, A. On the information stability of channels with timing errors. In Proceedings of the IEEE International Symposium on Information Theory (ISIT), Seattle, WA, USA, 9–14 July 2006; pp. 1885–1889.
- Shlezinger, N.; Abakasanga, E.; Dabora, R.; Eldar, Y.C. The Capacity of Memoryless Channels with Sampled Cyclostationary Gaussian Noise. *IEEE Trans. Commun.* **2020**, *68*, 106–121.
- Shannon, C.E. Communication in the presence of noise. *Proc. IEEE* **1998**, *86*, 447–457.

19. Cherif, F. A various types of almost periodic functions on Banach spaces: Part I. *Int. Math. Forum* **2011**, *6*, 921–952.
20. Shlezinger, N.; Dabora, R. On the capacity of narrowband PLC channels. *IEEE Trans. Commun.* **2015**, *63*, 1191–1201.
21. Shlezinger, N.; Dabora, R. The capacity of discrete-time Gaussian MIMO channels with periodic characteristics. In Proceedings of the IEEE International Symposium on Information Theory (ISIT), Barcelona, Spain, 10–16 October 2016; pp. 1058–1062.
22. Shlezinger, N.; Zahavi, D.; Murin, Y.; Dabora, R. The secrecy capacity of Gaussian MIMO channels with finite memory. *IEEE Trans. Inf. Theory* **2017**, *63*, 1874–1897.
23. Heath, R.W.; Giannakis, G.B. Exploiting input cyclostationarity for blind channel identification in OFDM systems. *IEEE Trans. Signal Process.* **1999**, *47*, 848–856.
24. Shaked, R.; Shlezinger, N.; Dabora, R. Joint estimation of carrier frequency offset and channel impulse response for linear periodic channels. *IEEE Trans. Commun.* **2017**, *66*, 302–319.
25. Shlezinger, N.; Dabora, R. Frequency-shift filtering for OFDM signal recovery in narrowband power line communications. *IEEE Trans. Commun.* **2014**, *62*, 1283–1295.
26. El Gamal, A.; Kim, Y.H. *Network Information Theory*; Cambridge University Press: Cambridge, UK, 2011.
27. Wu, X.; Xie, L.L. On the optimal compressions in the compress-and-forward relay schemes. *IEEE Trans. Inf. Theory* **2013**, *59*, 2613–2628.
28. Zitikovic, G. Lecture Notes on the Theory of Probability Parts I and II. Available online: [https://web.math.utexas.edu/users/gordanz/lecture\\_notes\\_page.html](https://web.math.utexas.edu/users/gordanz/lecture_notes_page.html) (accessed on 12 March 2020).
29. Papoulis, A. *Probability, Random Variables, and Stochastic Processes*; McGraw-Hill: New York, NY, USA, 2002.
30. Zamir, R.; Kochman, Y.; Erez, U. Achieving the Gaussian rate–distortion function by prediction. *IEEE Trans. Inf. Theory* **2008**, *54*, 3354–3364.
31. Rudin, W. *Principles of Mathematical Analysis*; International series in pure and applied mathematics; McGraw-Hill: New York, NY, USA, 1976.
32. Dixmier, J. *General Topology*; Springer-Verlag: New York, NY, USA, 1984.
33. Stein, E.M.; Shakarchi, R. *Real Analysis: Measure Theory, Integration, and Hilbert Spaces*; Princeton University Press: Princeton, NJ, USA, 2009.
34. Kolmogorov, A. On the Shannon theory of information transmission in the case of continuous signals. *IRE Trans. Inf. Theory* **1956**, *2*, 102–108.
35. Scheffé, H. A useful convergence theorem for probability distributions. *Ann. Math. Stat.* **1947**, *18*, 434–438.
36. Bromiley, P. Products and convolutions of Gaussian probability density functions. *Tina-Vision Memo* **2003**, *3*, 1.
37. Kosorok, M.R. *Introduction to Empirical Processes and Semiparametric Inference.*; Springer: New York, NY, USA, 2008.
38. Williams, D. *Probability with Martingales*; Cambridge University Press: Cambridge, UK, 1991.
39. Dobrushin, R.L. A general formulation of the fundamental theorem of Shannon in the theory of information. *Uspekhi Matematicheskikh Nauk* **1959**, *14*, 3–104.
40. Billingsley, P. *Convergence of Probability Measures*; John Wiley & Sons: New York, NY, USA, 2013.
41. Venkataramanan, R.; Pradhan, S.S. Source coding With feed-forward: Rate-distortion theorems and error exponents for a general source. *IEEE Trans. Inf. Theory* **2007**, *53*, 2154–2179.

