The Distortion Rate Function of Cyclostationary Gaussian Processes

Alon Kipnis, *Student Member, IEEE*, Andrea J. Goldsmith, *Fellow, IEEE*, and Yonina C. Eldar, *Fellow, IEEE*

***Abstract*— A general expression for the quadratic distortion rate function (DRF) of cyclostationary Gaussian processes in terms of their spectral properties is derived. This expression can be seen as the result of orthogonalization over the different components in the polyphase decomposition of the process. We use this expression to derive, in a closed form, the DRF of several cyclostationary processes arising in practice. We first consider the DRF of a combined sampling and source coding problem. It is known that the optimal coding strategy for this problem involves source coding applied to a signal with the same structure as one resulting from pulse amplitude modulation (PAM). Since a PAM-modulated signal is cyclostationary, our DRF expression can be used to solve for the minimal distortion in the combined sampling and source coding problem. We also analyze in more detail the DRF of a source with the same structure as a PAM-modulated signal, and show that it is obtained by reverse waterfilling over an expression that depends on the energy of the pulse and the baseband process modulated to obtain the PAM signal. This result is then used to explore the effect of the symbol rate in PAM on the DRF of its output. In addition, we also study the DRF of sources with an amplitude-modulation structure, and show that the DRF of a narrow-band Gaussian stationary process modulated by either a deterministic or a random phase sine-wave equals the DRF of the baseband process.**

***Index Terms*— Source coding, rate-distortion, modulation, Gaussian processes.**

I. Introduction

THE distortion rate function (DRF) describes the average minimal distortion achievable in sending an information source over a rate-limited noiseless link. Sources with memory posses an inherent statistical dependency that can be exploited in the context of data compression. However, not many closed-form expressions for the DRF of such sources are known, and those are usually limited to the class of stationary processes. Two notable exceptions are the DRFs of the Wiener process, derived by Berger [1], and of auto-regressive Gaussian processes, derived by Gray [2]. Indeed, information sources are rarely stationary in practice, and source coding techniques that are based on stationary assumptions about the source may achieve poor performance if the source has time-varying statistics.

A *Cyclostationary process* (CSP) (also known as *periodically correlated*, *periodically stationary* or *block-stationary* process) is a processes whose statistics are invariant to time shifts by integer multiples of a given time constant, denoted as the *period* of the process. As described in the survey by Gardner *et al.* [3], CSPs have been used in many fields to model periodic time-variant phenomena. In particular, they arise naturally in synchronous communication where block coding and modulation by periodic signals are employed. Spectral properties of CSPs were first studied by Bennett in [4] who also coined the term *cyclostationary*. These spectral properties and others that will be exploited in our derivations are also reviewed in [3] and in the references therein.

A coding theorem with respect to CSPs was first considered by Nedoma [5] (who referred to such processes as *block-stationary*). This theorem implies that the optimal tradeoff between code rate and distortion in encoding a CSP is given by its Shannon's DRF, i.e., by an optimization over a class of joint probability distributions subject to a mutual information rate constraint. Since CSPs are a special class of *asymptotic mean stationary* (AMS) processes, this coding theorem also arises as special cases from Gray's work on AMS processes [6]. Nevertheless, it seems that the only existing mechanism for evaluating this DRF is by the Karhunen-Loève (KL) expansion [7] of the process. In this method it is required to solve for the eigenvalues of a Fredholm integral equation for each finite blocklength, and use a waterfilling expression over these eigenvalues. The DRF is then obtained in the limit as the size of the blocklength goes to infinity. This evaluation, however, does not exploit the special block periodicity of CSPs. Moreover, it does not provide intuition on the optimal source coding technique in terms of spectral properties of the process. In contrast, the DRF of a stationary Gaussian process is obtained by waterfilling over its power spectral density, which provides clear intuition about how the source code represents each frequency component of the process [8].

An interesting motivation for evaluating the DRF of CSPs using their spectral properties arises in the following example: suppose we are interested in finding the DRF of the process obtained by modulating a continuous-time Gaussian stationary process $U(\cdot)$ by a cosine wave with random phase Φ, namely

$$X_\Phi(t) = \sqrt{2}U(t)\cos(2\pi f_0 t + \Phi), \quad t \in \mathbb{R}, \tag{1}$$

Manuscript received May 20, 2015; revised April 28, 2017; accepted July 27, 2017. Date of publication August 21, 2017; date of current version April 19, 2018. This work was supported in part by NSF under Grant CCF-1320628, in part by the NSF Center for Science of Information under Grant CCF-0939370, and in part by BSF Transformative Science under Grant 2010505. This paper was presented at the 2014 International Symposium on Information Theory.

A. Kipnis and A. J. Goldsmith are with the Department of Electrical Engineering, Stanford University, Stanford, CA 94305 USA (e-mail: alonkipnis@gmail.com).

Y. C. Eldar is with the Department of Electrical Engineering, Technion–Israel Institute of Technology, Haifa 32000, Israel.

Communicated by E. Tuncel, Associate Editor for Source Coding.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIT.2017.2741978

where Φ is uniformly distributed over $[0, 2\pi)$. This process is commonly given as an example of a wide-sense stationary process in signal processing textbooks (e.g. [9, Example 8.18]). Note that due to the random phase, $X_\Phi(\cdot)$ is not Gaussian and in fact is non-ergodic, since the distribution of the phase cannot be learned from a single realization of $X_\Phi(\cdot)$. It seems that in the context of rate-distortion theory, the spectrum of $X_\Phi(\cdot)$ can only be used to derive an upper bound on its DRF given by the DRF of a Gaussian stationary process with the same second order statistics [10, Th. 4.6.5]. The theory of AMS processes [11] implies that the DRF of $X_\Phi(\cdot)$ is given by the DRF of each one of its ergodic components [6, Th. 11.3.1], corresponding to different realizations $\varphi \in [0, 2\pi)$ of the phase Φ. Each ergodic component satisfies a source coding theorem such that its DRF is evaluated by optimizing over probability distributions subject to a mutual information rate constraint. One may think that this decomposition provides a recipe to evaluate the DRF of $X_\Phi(\cdot)$ by averaging over the DRF of the process $X_\varphi(\cdot)$, obtained by fixing the phase in $X_\Phi(\cdot)$. However, while the process $X_\varphi(\cdot)$ is Gaussian, it is no longer stationary – but rather cyclostationary. Since $X_\varphi(\cdot)$ arises by modulating a stationary process whose spectrum is known, it raises the question as to whether a simple expression for its DRF can be given in terms of the spectrum of the original stationary process. Such an expression for the DRF of the CSP $X_\varphi(\cdot)$ would lead to the DRF of the non-ergodic, non-Gaussian, stationary process $X_\Phi(\cdot)$.

In this work we derive an expression for (Shannon's) DRF of Gaussian CSPs which uses their spectral properties, and therefore generalizes the waterfilling expression for the DRF of Gaussian stationary processes derived by Kolmogorov [12]. This expression is obtained by considering the polyphase components of the process, which can be seen as a set of stationary processes that comprise the CSP (e.g. [13]). We show that the DRF of a discrete-time CSP can be obtained in closed form by orthogonalizing over these components at each frequency band. For continuous-time CSPs, we obtain an expression that is based on increasingly fine discrete-time approximations of the continuous-time signal. The DRF evaluated for these approximations converges to the DRF of the continuous-time process under mild conditions on its covariance function.

The main results of this paper are divided into two parts. In the first part we derive a general expression for evaluating the DRF of a second order Gaussian CSP in terms of its spectral properties. This expression is given in the form of a reverse waterfilling solution over the eigenvalues of a spectral density matrix defined in terms of the *time-varying spectral density* of the source. For discrete-time Gaussian processes, the size of this matrix equals the discrete period of the source. We extend our result to Gaussian CSPs in continuous-time by taking increasingly finer discrete time approximations. The resulting expression is a function of the eigenvalues of an infinite matrix. In addition, we derive a lower bound on the DRF which can be obtained without evaluating the matrix eigenvalues. We show that this bound is tight when the polyphase components of the process are highly correlated.

In the second part of the paper we use our general DRF expression to study the distortion-rate performance in several specific cases. In particular:

- We derive a closed form expression for the DRF of a process with a pulse-amplitude modulation (PAM) signal structure. We show how this expression can be used to derive the minimal distortion in estimating a stationary Gaussian process from a rate-limited version of its sub-Nyquist samples.
- We study the effect of the symbol rate in PAM on the DRF of the modulated signal at the output of the modulator. In particular, we quantify the intuition that the complexity of the signal at the output of the PAM increases with the symbol rate.
- We evaluate in closed form the DRF of a Gaussian stationary process modulated by a deterministic cosine wave. We show that the DRF of the modulated process equals that of the baseband stationary Gaussian process provided the latter is narrowband. We further conclude that the stationary, non-Gaussian and non-egodic process given by (1) has a DRF identical that of the modulated process without the random phase. These two results imply that the DRF of the stationary non-Gaussian amplitude modulated process is strictly smaller than the DRF of a Gaussian stationary process with the same second order statistics.

We note that the DRF of a signal obtained by modulating a baseband signal does not shed light onto the performance in using this modulation technique on the baseband signal to transmit it through a channel. Indeed, in order to measure the modulation method performance, the distortion of the received signal should be measured with respect to the baseband signal as in indirect source coding [10, Ch. 3.5], rather than with respect to the modulated signal. Nevertheless, the DRF of the modulated signal may be used to characterize the empirical distribution of any optimal source code that attains it, as has been done in [14] and [15] for i.i.d. signals. This distribution can then be used to derive the distortion in estimating the baseband signal from an encoded version of its modulated version, in a similar way to the case of i.i.d signals considered in [16]. However, since properties of the optimal codes in these modulation techniques are not yet known, the derivation of this distortion is beyond the scope of this paper.

The rest of this paper is organized as follows: in Section II we review concepts and notation from the theory of CSPs and rate distortion theory. Our main results are given in Section III, where we derive an expression for the DRF of a Gaussian CSP and a lower bound on this DRF. In Section IV we explore applications of our main result in various special cases. Concluding remarks are provided in Section V.

II. Background and Problem Formulation

A. Cyclostationary Processes

Throughout the paper, we consider zero mean Gaussian processes in both discrete and continuous time. We use round brackets to denote a continuous time index and square brackets for a discrete time index, i.e.

$$X(\cdot) = \{X(t),\ t \in \mathbb{R}\},$$

and

$$X[\cdot] = \{X[n],\ n \in \mathbb{Z}\}.$$

Matrices and vectors are denoted by bold letters.

The statistics of a zero mean Gaussian process $X(\cdot)$ is specified in terms of its autocorrelation function[1]

$$R_X(t,\tau) \triangleq \mathbb{E}\left[X(t+\tau)X(t)\right].$$

If in addition the autocorrelation function is periodic in t with a fundamental period T_0,

$$R_X(t+T_0,\tau) = R_X(t,\tau),$$

then we say that $X(\cdot)$ is a *cyclostationary process* (CSP) or simply *cyclostationary*[2] [3], [17]. We also assume that $R_X(t,\tau)$ is bounded and Riemann integrable on $[0,T_0]\times\mathbb{R}$, and therefore

$$\sigma_X^2 = \lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}\mathbb{E}X(t)^2dt = \frac{1}{T_0}\int_0^{T_0}R_X(t,0)dt$$

is finite.

Suppose that $R_X(t,\tau)$ has a convergent Fourier series representation in t for almost any $\tau\in\mathbb{R}$. Then the statistics of $X(\cdot)$ is uniquely determined by the *cyclic autocorrelation* (CA) function:

$$\hat{R}_X^n(\tau) \triangleq \frac{1}{T_0}\int_{-T_0/2}^{T_0/2}R_X(t,\tau)e^{-2\pi i n t/T_0}dt, \quad n\in\mathbb{Z}. \tag{2}$$

The Fourier transform of $\hat{R}_X^n(\tau)$ with respect to τ is denoted as the *cyclic power spectral density* (CPSD) function:

$$\hat{S}_X^n(f) = \int_{-\infty}^{\infty}\hat{R}_X^n(\tau)e^{-2\pi i\tau f}d\tau, \quad -\infty\le f\le\infty. \tag{3}$$

If $\hat{S}_X^n(f)$ is identically zero for all $n\neq 0$, then $R_X(t,\tau) = R_X(0,\tau)$ for all $0\le t\le T$ and the process $X(\cdot)$ is stationary. In such a case $S_X(f)\triangleq\hat{S}_X^0(f)$ is the *power spectral density* (PSD) function of $X(\cdot)$. The *time-varying power spectral density* (TPSD) function [17, Sec. 3.3] of $X(\cdot)$ is defined by the Fourier transform of $R_X(t,\tau)$ with respect to τ, i.e.

$$S_X^t(f) \triangleq \int_{-\infty}^{\infty}R_X(t,\tau)e^{-2\pi i f\tau}d\tau. \tag{4}$$

The Fourier series representation implies that

$$S_X^t(f) = \sum_{n\in\mathbb{Z}}\hat{S}_X^n(f)e^{2\pi i n t/T_0}. \tag{5}$$

[1]In [17] and in other references, the symmetric auto-correlation function

$$\tilde{R}_X(t,\tau)\triangleq\mathbb{E}\left[X(t+\tau/2)X(t-\tau/2)\right] = R_X(t-\tau/2,\tau),$$

the corresponding CPSD $\hat{\tilde{S}}_X^n(f)$ and TPSD $\tilde{S}_X^t(f)$, are used. The conversion between $\hat{S}^n(f)$ and the symmetric CPSD is given by $\hat{\tilde{S}}_X^n(f) = \hat{S}_X^n(f - n/(2T_0))$.

[2]It is customary to distinguish between wide-sense cyclostationarity which relates only to the second order statistics of the process, and strict-sense cyclostationarity which relates to the finite order statistics of the process [18, Ch. 10.4]. Both definitions coincide in the Gaussian case.

Associated with every CSP $X(\cdot)$ with period T_0 is a set of stationary discrete time processes $X^t[\cdot]$, $0\le t\le T_0$, defined by

$$X^t[n] = X\left(T_0 n + t\right), \quad n\in\mathbb{Z}. \tag{6}$$

These processes are called the *polyphase components* (PC) of the CSP $X(\cdot)$. The cross-correlation function of $X^{t_1}[\cdot]$ and $X^{t_2}[\cdot]$ is given by

$$\begin{aligned} R_{X^{t_1}X^{t_2}}[n,k] &= \mathbb{E}\left[X[T_0(n+k)+t_1]X[T_0 n+t_2]\right] \\ &= R_X\left(T_0 n+t_2, T_0 k+t_1-t_2\right) \\ &= R_X\left(t_2, T_0 k+t_1-t_2\right). \end{aligned} \tag{7}$$

Since $R_{X^{t_1}X^{t_2}}[n,k]$ depends only on k, this implies that $X^{t_1}[\cdot]$ and $X^{t_2}[\cdot]$ are jointly stationary. The PSD of $X^t[\cdot]$ is given by

$$\begin{aligned} S_{X^t}\left(e^{2\pi i\phi}\right) &\triangleq \sum_{k\in\mathbb{Z}}R_{X^tX^t}[0,k]e^{-2\pi i\phi k} \\ &= \sum_{k\in\mathbb{Z}}R_X\left(t,T_0 k\right)e^{-2\pi i\phi k}, \quad -\frac{1}{2}\le\phi\le\frac{1}{2}. \end{aligned} \tag{8}$$

Exploiting the spectral properties of sampled processes, we can use (8) and (5) to connect the functions $S_{X^t}\left(e^{2\pi i\phi}\right)$ and the CPSD of $X(\cdot)$ as follows:

$$\begin{aligned} S_{X^t}\left(e^{2\pi i\phi}\right) &= \frac{1}{T_0}\sum_{k\in\mathbb{Z}}S_X^t\left(\frac{\phi-k}{T_0}\right) \\ &= \frac{1}{T_0}\sum_{k\in\mathbb{Z}}\sum_{n\in\mathbb{Z}}\hat{S}_X^n\left(\frac{\phi-k}{T_0}\right)e^{2\pi i n t/T_0}. \end{aligned}$$

More generally, for $t_1,t_2\in[0,T_0]$ we have

$$\begin{aligned} &S_{X^{t_1}X^{t_2}}\left(e^{2\pi i\phi}\right) \\ &= \sum_{k\in\mathbb{Z}}R_{X^{t_1}X^{t_2}}[0,k]e^{-2\pi i k\phi} \\ &= \frac{1}{T_0}\sum_{k\in\mathbb{Z}}S_X^{t_2}\left(\frac{\phi-k}{T_0}\right)e^{2\pi i(t_1-t_2)\frac{\phi-k}{T_0}} \\ &= \frac{1}{T_0}\sum_{k\in\mathbb{Z}}\sum_{m\in\mathbb{Z}}\hat{S}_X^m\left(\frac{\phi-k}{T_0}\right)e^{2\pi i\left(m\frac{t_2}{T_0}+\frac{t_1-t_2}{T_0}(\phi-k)\right)}. \end{aligned} \tag{9}$$

We now turn to briefly describe the discrete-time counterpart of the CA, CPSD, TPSD and the polyphase components defined in (2), (3), (4) and (6), respectively.

A discrete time zero mean Gaussian process $X[\cdot]$ is said to be cyclostationary with period $M\in\mathbb{N}$ if its covariance function

$$R_X[n,k] = \mathbb{E}\left[X[n+k]X[n]\right]$$

is periodic in k with period M. For $m=0,\ldots,M-1$, the m^{th} cyclic autocorrelation (CA) function of $X[\cdot]$ is defined as

$$\hat{R}_X^m[k] \triangleq \sum_{n=0}^{M-1}R_X[n,k]e^{-2\pi i n m/M}.$$

The m^{th} CPSD function is then given by

$$\hat{S}_X^m\left(e^{2\pi i\phi}\right)\triangleq\sum_{k\in\mathbb{Z}}\hat{R}_X^m[k]e^{-2\pi i\phi k},$$

and the discrete TPSD function is

$$S_X^n\left(e^{2\pi i\phi}\right) \triangleq \sum_{k\in\mathbb{Z}} R_X[n,k]e^{-2\pi i\phi k}.$$

Finally, we have the discrete time Fourier transform relation

$$S_X^n\left(e^{2\pi i\phi}\right) = \frac{1}{M}\sum_{m=0}^{M-1}\hat{S}_X^m\left(e^{2\pi i\phi}\right)e^{2\pi i\phi nm/M}.$$

The m-th stationary component $\bar{X}^m[\cdot]$, $0 \le m \le M-1$ of $X[\cdot]$ is defined by

$$X^m[n] \triangleq X[Mn+m], \quad n\in\mathbb{Z}. \tag{10}$$

For $0 \le m, r, n \le M-1$ and $k \in \mathbb{Z}$ we have

$$\begin{aligned} R_{X^mX^r}[n,k] &= \mathbb{E}\left[X^m[n+k]X^r[n]\right] \\ &= \mathbb{E}\left[X[Mn+Mk+m]X[Mn+r]\right] \\ &= R_X[Mn+r, Mk+m-r] \\ &= R_X[r, Mk+m-r]. \end{aligned} \tag{11}$$

Using properties of multi-rate signal processing:

$$\begin{aligned} S_{X^mX^r}\left(e^{2\pi i\phi}\right) &= \sum_{k\in\mathbb{Z}} R_X[r, Mk+m-r]e^{-2\pi ik\phi} \\ &= \frac{1}{M}\sum_{n=0}^{M-1} S_X^r\left(e^{2\pi i\frac{\phi-n}{M}}\right)e^{2\pi i(m-r)\frac{\phi-n}{M}}. \end{aligned} \tag{12}$$

The discrete-time counterpart of (9) is then

$$S_{X^mX^r}\left(e^{2\pi i\phi}\right) = \frac{1}{M}\sum_{k=0}^{M-1}\sum_{n=0}^{M-1}\hat{S}_X^n\left(e^{2\pi i\frac{\phi-k}{M}}\right)e^{2\pi i\frac{nr+(m-r)(\phi-k)}{M}}. \tag{13}$$

The functions $S_{X^mX^r}\left(e^{2\pi i\phi}\right)$, $0 \le m, r \le M-1$ define an $M\times M$ matrix $\mathbf{S}_X\left(e^{2\pi i\phi}\right)$ with $(m+1, r+1)^{\text{th}}$ entry $S_{X^mX^r}\left(e^{2\pi i\phi}\right)$. This matrix completely determines the statistics of $X[\cdot]$, and can be seen as the PSD matrix associated with the stationary vector valued process $\mathbf{X}^M[n]$ defined by the stationary components of $X[\cdot]$:

$$\mathbf{X}^M[n] \triangleq \left(X^0[n], \ldots, X^{M-1}[n]\right), \quad n\in\mathbb{Z}. \tag{14}$$

We denote the autocorrelation matrix of $\mathbf{X}^M[\cdot]$ as the PSD-PC matrix. Note that the $(r+1, m+1)^{\text{th}}$ entry of this matrix is given by (11).

B. Examples

We now present two important modulation models which result in CSPs.

Example 1 (Amplitude Modulation (AM)): Given a Gaussian stationary process $U(\cdot)$ with PSD $S_U(f)$, consider the process

$$X_{\mathsf{AM}}(t) = \sqrt{2}U(t)\cos(2\pi f_0 t + \varphi),$$

where $f_0 > 0$ and $\varphi \in [0, 2\pi)$ are deterministic constants. This process is cyclostationary with period $T_0 = f_0^{-1}$ and CPSD [19, eq. (41)]

$$\hat{S}_{\mathsf{AM}}^m(f) = \frac{1}{2}\begin{cases} S_U(f+f_0)+S_U(f-f_0), & m=0, \\ S_U(f\mp f_0)e^{\pm 2i\varphi}, & m\pm 2, \\ 0, & \textit{otherwise}. \end{cases}$$

This leads to the TPSD

$$\begin{aligned} S_X^{\,t}(f) &= \frac{1}{2}S_U(f+f_0)(1+e^{-2(2\pi i f_0 t+\varphi)}) \\ &\quad + \frac{1}{2}S_U(f-f_0)(1+e^{2(2\pi i f_0 t+\varphi)}). \end{aligned} \tag{15}$$

Example 2 (Pulse-Amplitude Modulation (PAM)): Consider a Gaussian stationary process $U(\cdot)$ modulated by a deterministic signal $p(t)$ as follows:

$$X_{\mathsf{PAM}}(t) = \sum_{n\in\mathbb{N}} U(nT_0)p(t-nT_0). \tag{16}$$

This process is cyclostationary with period T_0 and CPSD [19, eq. (49)]

$$\hat{S}_{\mathsf{PAM}}^n(f) = \frac{1}{T_0}P(f)P^*\left(f-\frac{n}{T_0}\right)S_U(f), \quad n\in\mathbb{Z}, \tag{17}$$

where $P(f)$ is the Fourier transform of $p(t)$ and $P^*(f)$ is its complex conjugate. If T_0 is small enough such that the support of $P(f)$ is contained within the interval $\left(-\frac{1}{2T_0}, \frac{1}{2T_0}\right)$, then $\hat{S}_{\mathsf{PAM}}^n(f) = 0$ for all $n \neq 0$, which implies that $X_{PAM(\cdot)}$ is stationary.

C. The Distortion-Rate Function

For a fixed $T > 0$, let X_T be the reduction of $X(\cdot)$ to the interval $[-T, T]$. Define the distortion between two waveforms $x(\cdot)$ and $y(\cdot)$ over the interval $[-T, T]$ by

$$d_T\left(x(\cdot), y(\cdot)\right) \triangleq \frac{1}{2T}\int_{-T}^{T}\left(x(t)-y(t)\right)^2 dt. \tag{18}$$

We expand X_T by a Karhunen-Loève (KL) expansion [7, Ch. 9.7] as

$$X_T(t) = \sum_{k=1}^{\infty} X_k f_k(t), \quad -T \le t \le T, \tag{19}$$

where $\{f_k\}$ is a set of orthogonal functions over $[-T, T]$ satisfying the Fredholm integral equation

$$\lambda_k f_k(t) = \frac{1}{2T}\int_{-T}^{T} K_X(t,s)f_k(s)ds, \quad t\in[-T,T], \tag{20}$$

with corresponding eigenvalues $\{\lambda_k\}$, and where

$$K_X(t,s) \triangleq \mathbb{E}X(t)X(s) = R_X(s, t-s).$$

Assuming a similar expansion as (19) to an arbitrary random waveform Y_T, we have

$$\begin{aligned} \mathbb{E}d_T(X_T, Y_T) &= \frac{1}{2T}\int_{-T}^{T}\mathbb{E}\left(X(t)-Y(t)\right)^2 dt \\ &= \sum_{n\in\mathbb{Z}}\mathbb{E}\left(X_n - Y_n\right)^2. \end{aligned}$$

The mutual information between $X(\cdot)$ and $Y(\cdot)$ on the interval $[-T, T]$ is defined by

$$I_T\left(X(\cdot), Y(\cdot)\right) \triangleq \frac{1}{2T}\lim_{N\to\infty} I\left(\mathbf{X}_{-N}^N; \mathbf{Y}_{-N}^N\right),$$

where $\mathbf{X}_{-N}^N = (X_{-N}, \ldots, X_N)$, $\mathbf{Y}_{-N}^N = (Y_{-N}, \ldots, Y_N)$ and the X_ns and Y_ns are the coefficients in the KL expansion of

$X(\cdot)$ and $Y(\cdot)$, respectively. Denote by $\mathcal{P}_T$ the set of joint probability distributions $P_{X,\hat{X}}$ over the waveforms $\left(X(\cdot), \hat{X}(\cdot)\right)$, such that the marginal of $X(\cdot)$ agrees with the original distribution, and the average distortion $\mathbb{E}d_T\left(X(\cdot), \hat{X}(\cdot)\right)$ does not exceed D. The rate-distortion function (RDF) of $X(\cdot)$ is defined by

$$R(D) = \lim_{T\to\infty} R_T(D),$$

where

$$R_T(D) = \inf I_T\left(X(\cdot); \hat{X}(\cdot)\right)$$

and the infimum is over the set $\mathcal{P}_T$. It is well known that $R(D)$ and $R_T(D)$ are non-increasing convex functions of D, and therefore continuous in D over any open interval [10]. We define their inverse functions as the distortion-rate functions $D(R)$ and $D_T(R)$, respectively. We note that by its definition, $D(R)$ is bounded from above by the average power of $X(\cdot)$ over a single period:

$$\sigma_X^2 \triangleq \lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} \mathbb{E}X^2(t)dt = \lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} R_X(t,0)dt$$
$$= \frac{1}{T_0}\int_0^{T_0} R_X(t,0)dt = \hat{R}_X^0(0).$$

For Gaussian processes, we have the following parametric representation for $R_T(D)$ or $D_T(R)$ [7, eq. (9.7.41)]:

$$D_T(\theta) = \sum_{k=1}^{\infty} \min\{\theta, \lambda_k\} \tag{21a}$$

$$R_T(\theta) = \frac{1}{2}\sum_{k=1}^{\infty} \log^+(\lambda_k/\theta), \tag{21b}$$

where $\log^+ x \triangleq \max\{\log x, 0\}$.

In the discrete-time case the DRF is defined in a similar way as in the continuous-time setting described above by replacing the continuous-time index in (18), (19) and (20), and by changing integration to summation. Since the KL transform preserves norm and mutual information, this definition of the DRF in the discrete-time case is consistent with standard expressions for the DRF of a discrete-time source with memory as in [10, Ch. 4.5.2]. Note that with these definitions, the continuous-time distortion is measured in MSE per time unit while the discrete-time distortion is measured in MSE per source symbol. Similarly, in continuous-time, R represents bitrate, i.e., the number of bits per time unit. In the discrete-time setting we use the notation $\bar{R}$ to denote bits per source symbol.

Since the distribution of a zero-mean Gaussian CSP with period T_0 is determined by its second moment $R_X(t,\tau)$, we observe that such processes are T_0-ergodic and therefore *block-ergodic* as defined in [20, Definition 1]. It follows that a source coding theorem that associates $D(R)$ with the optimal MSE performance attainable in encoding $X(\cdot)$ at rate R is obtained from the main result of [20]. Another way to deduce a source coding theorem for CSPs is by noting that they belong to the class of asymptotic mean stationary process (AMS) [11, Excercise 6.3.1]. A source coding theorem for this class of processes in discrete-time can be found in [21]. Its extension to continuous-time follows immediately as long as the *flow* defined by the process, i.e. the mapping from the time index set $[-T, T]$ to the probability space, is measurable. This last condition is implicit in our definition of a continuous-time stationary process in terms of its finite dimensional probability distributions [22].

D. Problem Formulation: Evaluation of the DRF

In the special case in which $X(\cdot)$ is stationary, it is possible to obtain $D(R)$ without explicitly solving the Fredholm equation (20) or evaluating the KL eigenvalues: in this case, the density of these eigenvalues converges to the PSD $S_X(f)$ of $X(\cdot)$. This leads to the celebrated *reverse waterfilling* expression for the DRF of a stationary Gaussian process, originally derived by Kolmogorov [12]:

$$D(R_\theta) = \int_{-\infty}^{\infty} \min\{S_X(f), \theta\}\, d\phi. \tag{22a}$$

$$R_\theta = \frac{1}{2}\int_{-\infty}^{\infty} \log^+[S_X(f)/\theta]\, df. \tag{22b}$$

The discrete-time version of (22) is given by

$$D(\bar{R}_\theta) = \int_{-\frac{1}{2}}^{\frac{1}{2}} \min\left\{S_X\left(e^{2\pi i\phi}\right), \theta\right\} d\phi. \tag{23a}$$

$$\bar{R}_\theta = \frac{1}{2}\int_{-\frac{1}{2}}^{\frac{1}{2}} \log^+\left[S_X\left(e^{2\pi i\phi}\right)/\theta\right] d\phi. \tag{23b}$$

Equations (22) and (23) define the distortion as a function of the rate through a joint dependency on the water level parameter θ.

We note that stationarity is not a necessary condition for the existence of a density function for the eigenvalues in the KL expansion. For example, such a density function is known for the Wiener process [1] which is a non-stationary process.

The main problem we consider in this paper is the evaluation of $D(R)$ for a general Gaussian CSP. In principle, this evaluation can be obtained by computing the KL eigenvalues in (20) for each T, using (21) to obtain $D_T(R)$ and finally taking the limit as T goes to infinity. For general CSPs, however, an easy way to describe the density of the KL eigenvalues is in general unknown. As a result, the evaluation of the DRF directly by the KL eigenvalues usually does not lead to a closed-form solution. In the next section we derive an alternative representation for the function $D(R)$ which is based on an approximation of the kernel $K_X(t,s)$ used in (20). This representation leads to a simple expression for the DRF and allows the derivation of the DRF of the PAM and AM process of Examples 1-2 in closed forms.

III. Main Results

We now derive our main results regarding the DRF of a Gaussian CSP which do not involve the solution of the Fredholm integral equation (20). We begin by obtaining an expression for this DRF in terms of the eigenvalues of the PC-PSD matrix of the CSP. Next, we derive a lower bound on this DRF that does not require eigenvalue decomposition.

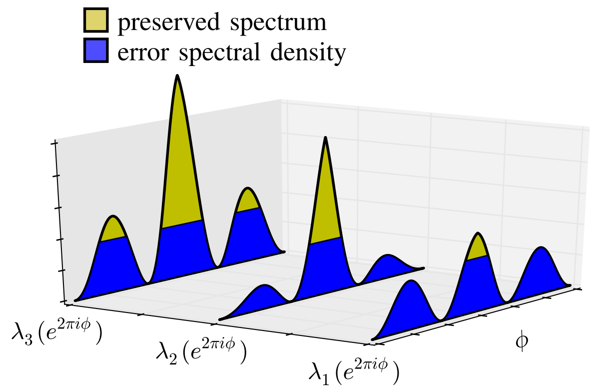

Fig. 1. Waterfilling interpretation of (25) for $M = 3$. The lossy compression error (blue) and the preserved spectrum (yellow) are associated with equations (25a) and (25b), respectively.

A. DRF in Terms of Spectral Properties

Our first observation is that in the discrete-time case, the DRF of a Gaussian CSP can be obtained by an expression for the DRF of a vector Gaussian stationary source. This expression is an extension of (23), which was derived in [23, eqs. (20) and (21)] and is given as follows:

$$D_{\mathbf{X}}(R_\theta) = \frac{1}{M}\sum_{m=1}^{M}\int_{-\frac{1}{2}}^{\frac{1}{2}} \min\left\{\lambda_m\left(e^{2\pi i\phi}\right), \theta\right\} d\phi \tag{24a}$$

$$R_\theta = \frac{1}{2M}\sum_{m=1}^{M}\int_{-\frac{1}{2}}^{\frac{1}{2}} \log^+\left[\lambda_m\left(e^{2\pi i\phi}\right)/\theta\right] d\phi, \tag{24b}$$

where $\lambda_1\left(e^{2\pi i\phi}\right), \ldots, \lambda_M\left(e^{2\pi i\phi}\right)$ are the eigenvalues of the PSD matrix $\mathbf{S_X}\left(e^{2\pi i\phi}\right)$ at frequency ϕ. In particular, we have the following result:

Theorem 1: Let $X[\cdot]$ be a discrete-time Gaussian cyclostationary process with period $M \in \mathbb{N}$. The distortion rate function of $X[\cdot]$ is given by

$$D(R_\theta) = \frac{1}{M}\sum_{m=1}^{M}\int_{-\frac{1}{2}}^{\frac{1}{2}} \min\left\{\lambda_m\left(e^{2\pi i\phi}\right), \theta\right\} d\phi \tag{25a}$$

$$\bar{R}_\theta = \frac{1}{2M}\sum_{m=1}^{M}\int_{-\frac{1}{2}}^{\frac{1}{2}} \log^+\left[\lambda_m\left(e^{2\pi i\phi}\right)/\theta\right] d\phi, \tag{25b}$$

where $\lambda_1\left(e^{2\pi i\phi}\right) \le \cdots \le \lambda_M\left(e^{2\pi i\phi}\right)$ are the eigenvalues of the PSD-PC matrix with $(m+1, r+1)^{th}$ entry given by

$$\mathbf{S}_{X^m X^r}\left(e^{2\pi i\phi}\right) = \frac{1}{M}\sum_{n=0}^{M-1} S_X^r\left(e^{2\pi i\frac{\phi-n}{M}}\right) e^{2\pi i(m-r)\frac{\phi-n}{M}}. \tag{26}$$

Proof: A full proof can be found in Appendix A. The idea is to use the polyphase decomposition (12) and the stationary vector valued process $\mathbf{X}^M[\cdot]$ defined in (14). The PSD matrix of the process is shown to coincide with the PSD-PC matrix of $X[\cdot]$. The proof shows that the DRF of $X[\cdot]$ is evaluated to the DRF of $\mathbf{X}^M[\cdot]$. The result then follows by applying (24) to $\mathbf{X}^M[\cdot]$. □

Equation (25) has the waterfilling interpretation illustrated in Fig. 1: the DRF is obtained by setting a single water-level over all eigenvalues of (26). These eigenvalues can be seen as the PSD of M independent processes obtained by the orthogonalization of the PC of $X[\cdot]$. The overall area below the water-level is the spectral density of the noise term in the test channel that attains Shannon's DRF, while the area above this level is associated with the reconstructed signal in this channel [10]. Compared to the limit in the discrete-time version of the KL expansion (21), expression (25) exploits the cyclostationary structure of the process by using its spectral properties. These spectral properties capture information on the entire time-horizon and not only over a finite blocklength as in the KL expansion.

The following theorem explains how to extend the above evaluation to the continuous-time case.

Theorem 2: Let $X(\cdot)$ be a Gaussian cyclostationary process with period T_0 and correlation function $R_X(t,\tau)$ Lipschitz continuous in its second argument. For a given $M \in \mathbb{N}$, denote

$$D_M(R_\theta) = \frac{1}{M}\sum_{m=1}^{M}\int_{-\frac{1}{2}}^{\frac{1}{2}} \min\left\{\lambda_m\left(e^{2\pi i\phi}\right), \theta_M\right\} d\phi \tag{27a}$$

$$R_\theta = \frac{1}{2T_0}\sum_{m=1}^{M}\int_{-\frac{1}{2}}^{\frac{1}{2}} \log^+\left[\lambda_m\left(e^{2\pi i\phi}\right)/\theta_M\right] d\phi, \tag{27b}$$

where $\lambda_1\left(e^{2\pi i\phi}\right) \le \cdots \le \lambda_M\left(e^{2\pi i\phi}\right)$ are the eigenvalues of the matrix $\mathbf{S}_X\left(e^{2\pi i\phi}\right)$ with its $(m+1, r+1)^{th}$ entry given by

$$\frac{1}{T_0}\sum_{k\in\mathbb{Z}} S_X^{rT_0/M}\left(\frac{\phi-k}{T_0}\right) e^{2\pi i(m-r)\frac{\phi-k}{M}} = \frac{1}{T_0}\sum_{k\in\mathbb{Z}}\sum_{n\in\mathbb{Z}} \hat{S}_X^n\left(\frac{\phi-k}{T_0}\right) e^{2\pi i\frac{nr+(m-r)(\phi-k)}{M}}. \tag{28}$$

Then the limit of D_M in M exists and the distortion-rate function of $X(\cdot)$ is given by

$$D(R) = \lim_{M\to\infty} D_M(\theta_M(R)). \tag{29}$$

1) Proof Sketch: The proof idea is to use a discrete-time CSP that approximates $X(\cdot)$. This approximation becomes tighter as M increases, so that the limit in (29) converges to the DRF of the continuous-time process. The proof details are given in Appendix B.

B. Discussion

The expression (27) is obtained by taking the limit in (25) over the time-period of a discrete-time CSP, where the code rate R is appropriately adjusted to bits per time unit. Although (27) only provides the DRF in terms of a limit, this limit is associated with the intra-cycle time resolution and not with the time horizon as in (21). This fact allows us to express the DRF in terms of the spectral properties of the process, that captures the "memory" of the process across the entire time horizon.

We note that limits of the form (29) have been obtained in closed-form using Szegő's Toeplitz distribution theorem [24, Sec. 5.2] when the underlying process is stationary and the matrix considered is Toeplitz [2], [10] or block Toeplitz [25], [26]. Unfortunately, the matrix in (28) is not

Toeplitz or block Toeplitz so Szegő's theorem is not applicable. In the following section we provide a few examples where the limit in (29) can be obtained in closed form which leads to a closed form expression for the DRF.

Equation (27) may be viewed as the extension to CSPs of the waterfilling expression (22) derived for stationary processes. While the latter can be understood as the limiting result of coding along orthogonal frequency bands [8], (27) implies that the DRF for CSPs is the result of two orthogonalization procedures: (1) over the PC inside a cycle, which is associated with the eigenvalue decomposition of the PSD-PC, and (2) over different frequency bands of the stationary processes resulting from the first orthogonalization. As in the case of stationary processes, the above procedure provides an intuitive way to attain the DRF: encode the mth orthogonal component resulting from the first orthogonalization using a separate bitstream whose rate is determined by

$$R_m \triangleq \frac{1}{T_0}\int_{-\frac{1}{2}}^{\frac{1}{2}} \log^+\left[\lambda_m\left(e^{2\pi i\phi}\right)/\theta_M\right]d\phi.$$

Within each component, the encoding is achieved by coding along orthogonal frequency bands [8] according to Pinsker's waterfilling expression (23) with bitrate R_m.

The main issue with evaluating the DRF using the limit in (29) is that it involves the eigenvalue decomposition of a matrix with growing dimension. Thus, we next derive a lower bound on the DRF that does not require the evaluation of these eigenvalues.

C. A Lower Bound

We now use the same decomposition of the process into its stationary PCs that led to (27) to derive a lower bound on the DRF. The basis for this bound is the following proposition, which holds for any source distribution and distortion measure (although we will consider here only the quadratic Gaussian case).

Proposition 1: Let $\mathbf{X}[\cdot]$ *be a vector valued process of dimension* M*. The distortion-rate function of* $\mathbf{X}[\cdot]$ *satisfies*

$$D_{\mathbf{X}}(R) \geq \frac{1}{M}\sum_{m=0}^{M-1} D_{X_m}(R). \quad (30)$$

Proof: Any rate R code for the process $\mathbf{X}[\cdot]$ can be seen as a rate R code for describing each of the coordinates $X_m[\cdot]$, $m = 0, \ldots, M-1$. At each coordinate, this code cannot achieve lower distortion than the optimal rate R code for that coordinate. □

Proposition 1 applied to Gaussian CSPs leads to the following result:

Proposition 2: Let $X[\cdot]$ *be a discrete-time Gaussian CSP with period* $M \in \mathbb{N}$*. The distortion rate function of* $X[\cdot]$ *satisfies*

$$D(\bar{R}) \geq \frac{1}{M}\sum_{m=0}^{M-1}\int_{-\frac{1}{2}}^{\frac{1}{2}} \min\left\{S_{X^m}\left(e^{2\pi i\phi}\right), \theta_m\right\}d\phi, \quad (31)$$

where for each $m = 0, \ldots, M-1$, θ_m *satisfies*

$$\bar{R}(\theta_m) = \frac{1}{2}\int_{-\frac{1}{2}}^{\frac{1}{2}} \log^+\left[S_{X^m}\left(e^{2\pi i\phi}\right)/\theta_m\right]d\phi, \quad (32)$$

and

$$S_{X^m}\left(e^{2\pi i\phi}\right) \triangleq S_{X^m X^m}\left(e^{2\pi i\phi}\right) = \frac{1}{M}\sum_{n=0}^{M-1} S_X^m\left(e^{2\pi i\frac{\phi-n}{M}}\right)$$

is the PSD of the m^{th} *PC of* $X[\cdot]$.

1) Proof: The claim is a direct application of Proposition 1 to our case of a discrete-time CSP: the summands on the RHS of (31) are the individual DRF of the PCs $X^m[\cdot]$, $m = 0, \ldots, M-1$, of $X[\cdot]$ obtained by (23).

Proposition 2 may be extended to the continuous-time case by approximating the outer integral in (33) by finite sums. This yields the following result:

Proposition 3: Let $X(\cdot)$ *be a continuous-time Gaussian cyclostationary process with period* $T_0 > 0$ *and correlation function* $R_X(t,\tau)$ *that is Lipschitz continuous in its second argument. The distortion rate function of* $X(\cdot)$ *satisfies*

$$D(R) \geq \frac{1}{T_0}\int_0^{T_0}\int_{-\frac{1}{2}}^{\frac{1}{2}} \min\left\{\sum_{n\in\mathbb{Z}} S_X^t\left(\frac{\phi-n}{T_0}\right), \theta_t\right\}d\phi dt, \quad (33)$$

where for each $0 \leq t \leq T_0$, θ_t *satisfies*

$$R(\theta_t) = \frac{1}{2T_0}\int_{-\frac{1}{2}}^{\frac{1}{2}} \log^+\left[\sum_{n\in\mathbb{Z}} S_X^t\left(\frac{\phi-n}{T_0}\right)/\theta_t\right]d\phi. \quad (34)$$

2) Proof: See Appendix C.

The bound (31) is obtained by averaging the minimal distortion at rate R in describing each one of the PCs of $X(\cdot)$. For each such component $X^t[\cdot]$ there is an associated water level θ_t obtained by solving (34) for θ_t. For $R = 0$, θ_t is always bigger than the essential supremum of

$$S_{X^t}\left(e^{2\pi i\phi}\right) = \frac{1}{T_0}\sum_{n\in\mathbb{Z}} S_X^t\left(\frac{\phi-n}{T_0}\right),$$

so the RHS of (31) equals the average over the total power of each one of the PCs of $X(\cdot)$ which are summed to $\sigma_X^2 = D_X(0)$. On the other hand, if $R \to \infty$, then $\theta_t \to 0$ for all $t \in [0, T_0]$, and again equality holds in (31). That is, the bound is tight in the two extremes of $R = 0$ and $R \to \infty$.

From a source coding point of view, the bound (33) can be understood as if a source code of rate R is applied to each of the PCs of $X(\cdot)$ individually. On the other hand, the DRF in (27) is obtained by applying a single rate R code to describe all PCs simultaneously. As a result, the bound becomes tight only when the PCs are highly correlated with one another, i.e., when a single PC determines the rest of them. A case where the latter holds is shown in the following example.

Example 3 (Equality in (33)*): Let* $X(\cdot)$ *be the PAM signal of Example 2 where the pulse* $p(t)$ *is given by*

$$p(t) = \begin{cases} 1 & 0 \leq t < T_0, \\ 0 & \textit{otherwise}. \end{cases}$$

The sample path of $X(\cdot)$ *has a staircase shape as illustrated in Fig. 2. This process is equivalent to the discrete time*

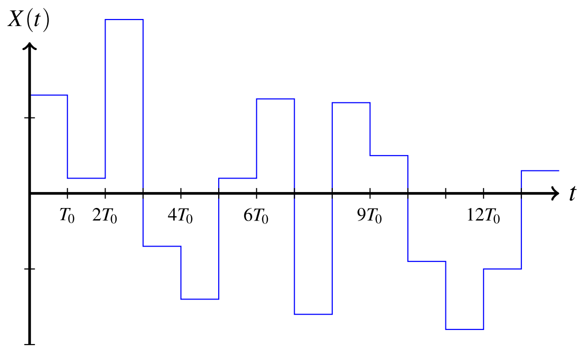

Fig. 2. An example of a continuous-time PAM process that attains equality in (33).

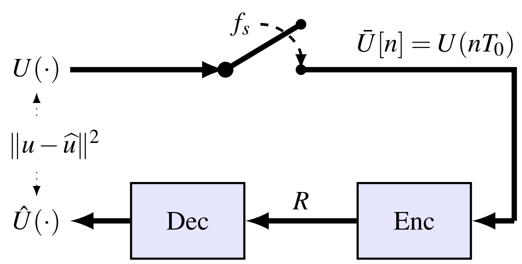

Fig. 3. Combined sampling and source coding system model.

process $\bar{U}[\cdot] \triangleq \{U(nT_0),\ n \in \mathbb{Z}\}$ *both in information rate and squared norm per period* T_0*, which is enough to conclude that* $D_X(R) = D_U(RT_0)$*. Indeed, the PCs in this case are maximally correlated, in the sense that a realization of* $X^0[\cdot] = \{X(nT_0),\ n \in \mathbb{Z}\}$ *determines the value of* $X^\Delta[\cdot] = \{X((n+\Delta)T_0),\ n \in \mathbb{Z}\}$ *for all* $0 \le \Delta < 1$*. In addition, for all* $0 \le t \le T_0$ *we have*

$$S_X^t\left(e^{2\pi i\phi}\right) = S_X^0\left(e^{2\pi i\phi}\right) = \sum_{n\in\mathbb{Z}} S_U\left(\frac{\phi-n}{T_0}\right),$$

where the latter is the PSD of the discrete time process $\bar{U}[\cdot]$*, so* (22) *implies that the RHS of* (33) *is the DRF of* $\bar{U}[\cdot]$*. We therefore conclude that the DRF of* $X(\cdot)$ *is given by the RHS of* (33)*.*

IV. Applications

In this section we apply the expression obtained in Theorem 2 to study the distortion-rate performance of a few CSPs that arise in practice.

A. Combined Sampling and Source Coding

We begin with the distortion-rate performance in the combined sampling and source coding problem considered in [27]. This problem is described by the system of Fig. 3: the source $U(\cdot)$ is a Gaussian stationary process with a known PSD $S_U(f)$. The source is uniformly sampled at rate $f_s = T_s^{-1}$, resulting in the discrete time process $\bar{U}[\cdot]$ defined by $\bar{U}[n] = U(n/f_s)$. The process $\bar{U}[\cdot]$ is then encoded at rate R bits per time unit. The goal is to estimate the source $U(\cdot)$ from its sampled and encoded version under a quadratic distortion. We denote by the function $D_{U|\bar{U}}(f_s, R)$ the minimal distortion attainable in this estimation, where the minimization is over all collections of encoders and decoders operating at bitrate R. Note that if $U(\cdot)$ is sampled above its Nyquist rate, then there is no loss of information in the sampling operation, and we get

$$D_{U|\bar{U}}(f_s, R) = D_U(R),$$

where $D_U(R)$ is found by (22). Therefore, the interesting case is that of sub-Nyquist sampling of $U(\cdot)$. In what follows we use Theorem 2 to derive $D_{U|\bar{U}}(f_s, R)$ in closed form.

Our first observation is that the combined sampling and source coding problem of Fig. 3 may be seen as an indirect source coding problem [28]; the distortion is measured with respect to the process $U(\cdot)$, but a different process, namely $\bar{U}[\cdot]$, is available to the encoder. Wolf and Ziv [29] have shown that the optimal source coding scheme under quadratic distortion for this class of problems is obtained as follows: the encoder first obtains the minimal mean square error (MMSE) estimate of the unseen source, and then an optimal source code is applied to describe this estimated sequence to the decoder. In the setting of Fig. 3, this implies that $D_{U|\bar{U}}(f_s, R)$ is attained by first obtaining the MMSE estimate

$$\widetilde{U}(t) = \mathbb{E}\left[U(t)|\bar{U}[\cdot]\right]$$

at the encoder, and then solving a standard source coding problem with respect to $\widetilde{U}(\cdot)$. Moreover, this scheme implies that the distortion decomposes into two parts:

$$D_{U|\bar{U}}(f_s, R) = \mathsf{mmse}(U|\bar{U}) + D_{\widetilde{U}}(R), \tag{35}$$

where $\mathsf{mmse}(U|\bar{U})$ is the MMSE in estimating $U(\cdot)$ from $\bar{U}[\cdot]$, and $D_{\widetilde{U}}(R)$ is the DRF of the process $\widetilde{U}(\cdot)$.

Standard linear estimation techniques [30] leads to

$$\widetilde{U}(t) = \sum_{n\in\mathbb{Z}} \bar{U}[n]w(t-nT_0) = \sum_{n\in\mathbb{Z}} U(nT_0)w(t-nT_0),$$

where the Fourier transform of $w(t)$ is given by

$$W(f) = \frac{S_U(f)}{\sum_{k\in\mathbb{Z}} S_U(f-k/T_0)}. \tag{36}$$

Moreover, the error in this estimation is

$$\mathsf{mmse}(U|\bar{U}) = \int_{-\infty}^{\infty} S_U(f)df - \int_{-\frac{1}{2T_0}}^{\frac{1}{2T_0}} \widetilde{S}_W(f)df, \tag{37}$$

where

$$\widetilde{S}_W(f) = \sum_{k\in\mathbb{Z}} |W(f-k/T_0)|^2 S_U(f-k/T_0). \tag{38}$$

We conclude from the above that $D_{\widetilde{U}}(R)$, and therefore $D_{U|\bar{U}}(f_s, R)$, is obtained by solving a source coding problem for an information source with a PAM signal structure, illustrated in Fig. 4. Since Example 2 implies that such a signal is cyclostationary with period $T_s = f_s^{-1}$, we can apply Theorem 2 in order to evaluate this DRF. By doing so, we obtain the following general result:

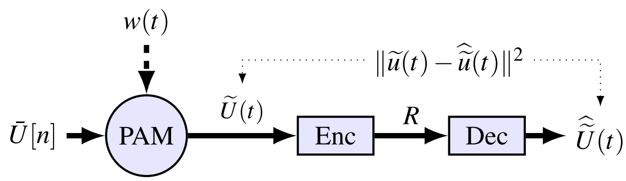

Fig. 4. Minimal distortion in Fig. 3 is obtained by a PAM modulator followed by an optimal source code (the Enc" in this figure) for the output of this modulator.

Proposition 4 (DRF of PAM-Modulated Signals): Let $X_{\text{PAM}}(\cdot)$ be defined by

$$X_{\text{PAM}}(t) = \sum_{n \in \mathbb{Z}} U(nT_0) p(t - nT_0), \quad t \in \mathbb{R}, \tag{39}$$

where $U(\cdot)$ is a Gaussian stationary process with[3] *PSD $S_U(f)$ and $p(t)$ is an analog deterministic signal with $\int_{-\infty}^{\infty} |p(t)|^2 dt < \infty$ and Fourier transform $P(f)$. Assume moreover, that the covariance function $R_{X_{\text{PAM}}}(t, \tau)$ of $X_{\text{PAM}}(\cdot)$ is Lipschitz continuous in its second argument. The distortion-rate function of $X_{\text{PAM}}(\cdot)$ is given by*

$$D(\theta) = \frac{1}{T_0} \int_{-\frac{1}{2T_0}}^{\frac{1}{2T_0}} \min\left\{\widetilde{S}(f), \theta\right\} df \tag{40a}$$

$$R(\theta) = \frac{1}{2} \int_{-\frac{1}{2T_0}}^{\frac{1}{2T_0}} \log^+ \left[\widetilde{S}(f)/\theta\right] df, \tag{40b}$$

where

$$\widetilde{S}(f) \triangleq \sum_{k \in \mathbb{Z}} |P(f - k/T_0)|^2 S_U(f - k/T_0). \tag{41}$$

Proof: See Appendix D. □

Proposition 4 applied to the process $\widetilde{U}(\cdot)$ implies that its DRF $D_{\widetilde{U}}(R)$ is given by waterfilling over the function

$$J(f) \triangleq \sum_{k \in \mathbb{Z}} |W(f - kf_s)|^2 S_U(f - kf_s).$$

As a result, we obtain from (35) and (40) the following expression for the minimal distortion in the combined sampling and source coding problem:

$$D_{U|\bar{U}}(f_s, R) = \mathsf{mmse}(U|\bar{U}) + \frac{1}{T_0} \int_{-\frac{1}{2T_0}}^{\frac{1}{2T_0}} \min\{J(f), \theta\} df, \tag{42a}$$

where

$$R(\theta) = \frac{1}{2} \int_{-\frac{1}{2T_0}}^{\frac{1}{2T_0}} \log^+ [J(f)/\theta] df. \tag{42b}$$

[3] Although we only use the value of $U(t)$ at $t \in \mathbb{Z}T_0$, it is convenient to treat $U(\cdot)$ as a continuous-time source so that the expressions emerging have only continuous-time spectrum.

B. Information Rates of Signals With PAM Structure

Proposition 4 provides a general closed form for the DRF of Gaussian processes with a PAM structure. In this subsection we use this expression to study the effect of the PAM of (39) on the complexity or the compressibility of the signal generated by this modulation, as a function of the symbol rate at its input. Intuitively, this complexity is increased with the symbol rate, since then the output contains more information on the modulated symbols per unit time. This increase, however, reaches saturation if these symbols become correlated. This phenomena can be quantified precisely by exploring the dependency of the DRF of the modulator output $X_{\text{PAM}}(\cdot)$ on the symbol rate.

For convenience, we assume that the origin of the random symbols in the PAM is a Gaussian stationary process $U(\cdot)$ sampled every T_0 time units. Therefore, the PAM output of (39) can be seen as a non-ideal reconstruction of $U(\cdot)$ from its uniform samples using pulses of shape $p(t)$, as illustrated in Fig. 6. Since the randomness in $X_{\text{PAM}}(\cdot)$ is only due to $U(\cdot)$, the process $X_{\text{PAM}}(\cdot)$ may be described with a lower bitrate than $U(\cdot)$ to attain the same distortion level, when this level is normalized to account for the effect of the modulator on the signal's energy. In other words, the DRF curve of $X_{\text{PAM}}(\cdot)$ is bounded from above by the DRF of $U(\cdot)$. In addition, we expect the DRF of $X_\Phi(\cdot)$ to increase with the symbol rate $1/T_0$, and saturate as this rate exceeds the Nyquist rate of $U(\cdot)$, provided the latter is bandlimited. Indeed, when $1/T_0$ is higher than the Nyquist rate of $U(\cdot)$, the support of $S_U(f)$ is contained within $(-1/2T_0, 1/2T_0)$. In this case, (40) implies that the DRF (and the RDF) of $X_{\text{PAM}}(\cdot)$ is obtained by waterfilling over the function

$$\widetilde{S}(f) = |P(f)|^2 S_U(f). \tag{43}$$

That is, the effect of the modulation in super-Nyquist sampling is identical to the effect of a linear filter with frequency response $P(f)$ applied to $U(\cdot)$. This filtering can be understood as a linear transformation of the coordinates represented by the frequency components [31, Ch. 22]. Assuming that $P(f)$ does not change the support of (43) (that is, this change in coordinates is invertible), the process $U(\cdot)$ may be recovered from $X_{\text{PAM}}(\cdot)$ with zero mean-square error. When the sampling frequency, i.e., the symbol rate $1/T_0$, is below the Nyquist rate of $U(\cdot)$, perfect recovery of $U(\cdot)$ is impossible in general. Intuitively, in this case $X_{\text{PAM}}(\cdot)$ is more compressible and therefore can be represented with fewer bits per unit time than $U(\cdot)$ for the same distortion level. A quantitative evaluation of this effect is given in Fig. 6, where the DRF of $X_{\text{PAM}}(\cdot)$ is compared to the DRF of $U(\cdot)$ for three sub-Nyquist symbol rates. Examples for the realization of $X_{\text{PAM}}(\cdot)$ and $U(\cdot)$ using sub- and super- Nyquist symbol rates are given in Fig. 5.

C. Amplitude Modulation With Random Phase

In this section we turn back to the two processes discussed in the introduction as our motivating examples and evaluate their DRFs using Theorem 2.

Consider the process $X_\Phi(\cdot)$ obtained by modulating a stationary Gaussian process $U(\cdot)$ by a cosine-wave of

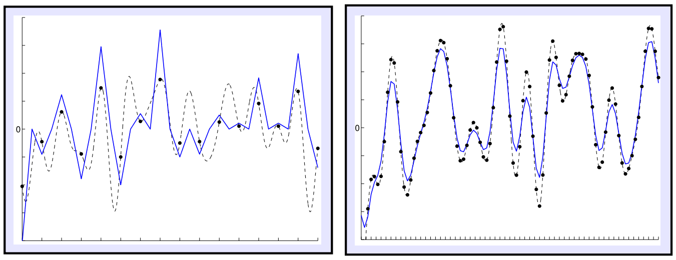

Fig. 5. Two realizations over time of the PAM process $X_{\mathrm{PAM}}(\cdot)$ (blue) and the baseband process $U(\cdot)$ (dashed) with the PSD and pulse shape given in Fig. 6, corresponding to sub-Nyquist (left) and super-Nyquist (right) symbol rates. Fig. 6 below shows that for the same distortion level, the PAM-modulated process on the left can be described with fewer bits per unit time than the PAM-modulated process on the right.

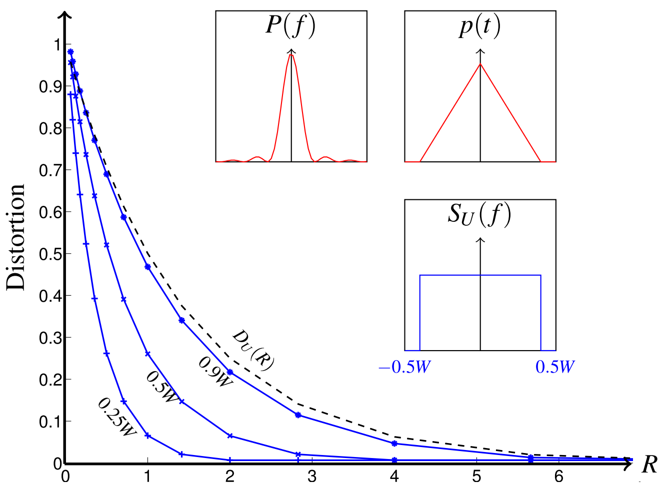

Fig. 6. The DRF of the PAM signal (16) for three values of symbol rate $1/T_0$ compared to the Nyquist rate W of $U(\cdot)$. The DRF of the baseband stationary signal $U(\cdot)$ (assuming that the modulation preserves the average power of the signal) is given by the dashed curve. The PSD of $U(\cdot)$ and the shape of the pulse $p(t)$ are given in the small frames. This figure shows that as the symbol rate decreases, the PAM process can be described with fewer bits per unit time for the same distortion level.

frequency f_0 and a random phase Φ uniform over $[0, 2\pi)$, as defined in (1). It is an elementary exercise [9, Example 8.18] to show that the process $X_\Phi(\cdot)$ is stationary with PSD

$$S_\Phi(f) = \frac{1}{2} S_U(f - f_0) + \frac{1}{2} S_U(f + f_0). \tag{44}$$

From [10, Th. 4.6.5], an upper bound on the DRF of $X_\Phi(t)$, denoted by $D_{X_\Phi}(R)$, is obtained by the DRF of a Gaussian process with the same PSD $S_\Phi(f)$ through the reverse-waterfilling (22). However, it seems that $D_{X_\Phi}(R)$ cannot be determined solely from the second order statistics of $X_\Phi(\cdot)$.

The main obstacle in deriving $D_{X_\Phi}(R)$ is the random phase of $X_\Phi[\cdot]$, which makes the process non-Gaussian and non-ergodic. This random phase may be handled using an asynchronous block code [6, Ch. 11.6], i.e. by adding a short prefix consisting of a source synchronization word to each block. Indeed, the following proposition follows directly from the proof of [6, Th. 11.6.1]:

Proposition 5: For any $\varphi \in [0, 2\pi)$ (deterministic), the DRF of $X_\Phi(\cdot)$ coincides with the DRF of the process

$$X_\varphi(t) = \sqrt{2} U(t) \cos(2\pi f_0 t + \varphi), \quad t \in \mathbb{R}. \tag{45}$$

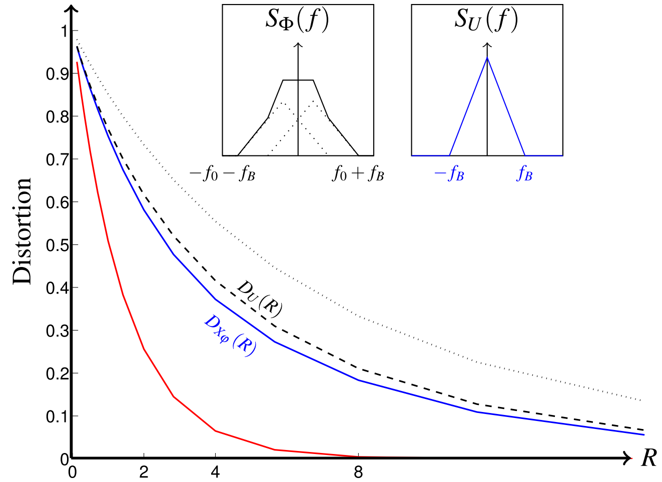

Fig. 7. The DRF of the processes $X_\varphi(\cdot)$ of (45) (blue), the baseband process $U(\cdot)$ (dashed), and the lower bound of Proposition 3. The PSD $S_U(f)$ is taken to be the pulse given in the small frame. Proposition 6 implies that $D_U(R)$ and $D_{X_\varphi}(R)$ coincide for $f_s > 2f_B$. Also shown is the DRF of the Gaussian stationary process with PSD $S_\Phi(f)$ (dotted), which provides an upper bound on $D_{X_\varphi}(R)$.

It was noted in Example 1 above that $X_\varphi(\cdot)$ is a CSP with the SCD function (15). It follows that $D_{X_\Phi}(R)$ is given by the DRF of the Gaussian CSP $X_\varphi(\cdot)$, generated by modulating the stationary Gaussian process $U(\cdot)$ using a deterministic cosine wave. Note that regardless of the carrier frequency f_0, the baseband process $U(\cdot)$ can always be recovered from $X_\varphi(\cdot)$, and that the $\sqrt{2}$ factor implies that the modulation preserves energy. These two facts are not enough to guarantee equality between the DRFs of the processes, since the modulation may lead to a 'change in coordinates' in the spectrum, in analogy with (43) and [31, Ch. 22]. In the following proposition we use Theorem 2 to show that this equality indeed holds as long as f_0 is larger than twice the bandwidth of $S_U(f)$.

Proposition 6: Let $U(\cdot)$ be a Gaussian stationary process bandlimited to $(-f_B, f_B)$. Let $f_0 > 2f_B$. The DRF of the process

$$X_\varphi(t) = \sqrt{2} U(t) \cos(2\pi f_0 t + \varphi), \quad t \in \mathbb{R},$$

equals the DRF of the stationary Gaussian process $U(\cdot)$.

Proof: See Appendix E. □

Proposition 7 asserts that the process $X_\varphi(\cdot)$ with AM signal structure suffers the same minimal distortion as the baseband process $U(\cdot)$ upon encoding each of them at rate R, provided the latter is narrowband. Fig. 7 shows that this equality does not necessarily hold when $U(\cdot)$ is not narrowband. Propositions 5 and 6 leads to the following conclusion:

Corollary 7: Let $U(\cdot)$ be a Gaussian stationary process bandlimited to $(-f_B, f_B)$. Assume that Φ is uniformly distributed over $(0, 2\pi)$ and $f_0 > 2f_B$. The distortion rate function of the stationary process

$$X_\Phi(t) = \sqrt{2} U(t) \cos(2\pi f_0 t + \Phi), \quad t \in \mathbb{R},$$

equals the DRF of the baseband process $U(\cdot)$.

It is interesting to note that the DRF of a Gaussian process with the same PSD as the stationary process $X_\Phi(\cdot)$ is strictly larger than the DRF of the baseband process $U(\cdot)$, and therefore provides an upper bound on $D_{X_\Phi}(R)$. This upper bound is illustrated in Fig. 7.

V. Conclusions

We derived an expression for the DRF of a class of Gaussian processes with periodically time-varying statistics, known as CSPs. This DRF is computed by reverse waterfilling over eigenvalues of a spectral density matrix associated with the polyphase components in the decomposition of the source. Unlike other general expressions for the DRF of Gaussian processes that use orthogonal basis expansions over increasing but finite time intervals, the expression we derive exploits the cyclostationarity of the process by orthogonalizing the polyphase components. Since these components are defined over the entire time horizon, the resulting expression can be given in terms of the spectrum of the process. In the continuous-time counterpart the solution is given in terms of a limit over functions of these eigenvalues.

While we leave open the question whether there exists a closed form solution to the above limit in general, we evaluated this limit in two special cases: a Gaussian signal with PAM signal structure, and a Gaussian signal with AM signal structure. We obtained the corresponding DRFs of these processes in terms of the power spectral density of the baseband stationary processes. The DRF result for a process with a PAM structure was then used to derive the DRF of a process under combined sampling and source coding.

In addition to an expression for the DRF of CSPs, we derived a lower bound on this DRF by averaging the minimal distortion attained in encoding each of the polyphase components over a single period. This bound is tight when high correlation among these components is present.

Appendix A

In this Appendix we prove Theorem 1. Consider the vector valued process $\mathbf{X}^M[\cdot]$ defined in (14). The rate-distortion function of $\mathbf{X}^M[\cdot]$ is given by (24):

$$D(\theta) = \frac{1}{M}\sum_{m=1}^{M}\int_{-\infty}^{\infty}\min\left\{\lambda_m\left(e^{2\pi i\phi}\right),\theta\right\}d\phi, \tag{46a}$$

$$R(\theta) = \frac{1}{2}\sum_{m=1}^{M}\int_{-\infty}^{\infty}\log^+\left[\lambda_m\left(e^{2\pi i\phi}\right)/\theta\right]d\phi, \tag{46b}$$

where $0 \le \lambda_1\left(e^{2\pi i\phi}\right) \le \cdots \le \lambda_M\left(e^{2\pi i\phi}\right)$ are the eigenvalues of the spectral density matrix $\mathbf{S}_{\mathbf{X}^M}\left(e^{2\pi i\phi}\right)$ obtained by taking the Fourier transform of covariance matrix $\mathbf{R}_{\mathbf{X}}[k] = \mathbb{E}\left[X^M[n+k](X^M[n])^T\right]$ entry-wise. The $(m,r)^{\text{th}}$ entry of $\mathbf{S}_{\mathbf{X}^M}\left(e^{2\pi i\phi}\right)$ is given by (12):

$$\begin{aligned}\left(\mathbf{S}_{\mathbf{X}^M}\left(e^{2\pi i\phi}\right)\right)_{m,r} &= S_X^{m,r}\left(e^{2\pi i\phi}\right)\\ &= \frac{1}{M}\sum_{k=0}^{M-1}S_X^r\left(e^{2\pi i\frac{\phi-k}{M}}\right)e^{2\pi i(m-r)\frac{\phi-k}{M}}.\end{aligned} \tag{47}$$

It is left to show that the DRF of $\mathbf{X}^M[\cdot]$ coincides with the DRF of $X[\cdot]$. By the source coding theorem for AMS processes [6, Th. 11.4.1] it is enough to show that the operational block coding distortion-rate function ([6, Ch. 11.2]) of both processes is identical. Indeed, any N block codebook for $\mathbf{X}^M[\cdot]$ is an MN block codebook for $X[\cdot]$ which achieves the same quadratic distortion averaged over the block. However, since $\mathbf{X}^M[\cdot]$ is stationary, by [6, Lemma 11.2.3] we know that any distortion above the DRF of $\mathbf{X}^M[\cdot]$ is attained for large enough N. This implies that the same is true for $X[\cdot]$.

Appendix B

In this Appendix we prove Theorem 2. Given a Gaussian cyclostationary process $X(\cdot)$ with period $T_0 > 0$, we define the discrete-time process $\bar{X}[\cdot]$ obtained by uniformly sampling $X(\cdot)$ at intervals T_0/M, i.e.

$$\bar{X}[n] = X(nT_0/M), \quad n \in \mathbb{Z}. \tag{48}$$

The autocorrelation function of $\bar{X}[\cdot]$ satisfies

$$\begin{aligned}R_{\bar{X}}[n+M,k] &= \mathbb{E}\left[\bar{X}[n+M+k]\bar{X}[n+M]\right]\\ &= \mathbb{E}\left[X(nT_0/M+T_0+kT_0/M)X(nT_0/M+T_0)\right]\\ &= R_X(nT_0/M+T_0, kT_0/M+T_0)\\ &= R_X(nT_0/M, kT_0/M)\\ &= R_{\bar{X}}[n,k],\end{aligned}$$

which means that $\bar{X}[\cdot]$ is a discrete-time Gaussian cyclostationary process with period M. The TPSD of $\bar{X}[\cdot]$ is given by

$$S_{\bar{X}}^m(e^{2\pi i\phi}) = \frac{M}{T_0}\sum_{k\in\mathbb{Z}}S_X^{mT_0/M}\left(\frac{M}{T_0}(\phi-k)\right).$$

This means that the PSD of the m^{th} PC of $\bar{X}[\cdot]$ is

$$\begin{aligned}S_{\bar{X}}^m\left(e^{2\pi i\phi}\right) &= \frac{1}{M}\sum_{n=0}^{M-1}S_{\bar{X}}^m\left(e^{2\pi i\frac{\phi-n}{M}}\right)\\ &= \frac{1}{T_0}\sum_{n=0}^{M-1}\sum_{k\in\mathbb{Z}}S_X^{mT_0/M}\left(\frac{\phi-Mk-n}{T_0}\right)\\ &= \frac{1}{T_0}\sum_{l\in\mathbb{Z}}S_X^{mT_0/M}\left(\frac{\phi-l}{T_0}\right).\end{aligned}$$

By applying Theorem 1 to $\bar{X}[\cdot]$, we obtain an expression for the DRF of $\bar{X}[\cdot]$ as a function of M:

$$D_M(\theta_M) = \frac{1}{M}\sum_{m=1}^{M}\int_{-\frac{1}{2}}^{\frac{1}{2}}\min\left\{\lambda_m\left(e^{2\pi i\phi}\right),\theta_M\right\}d\phi \tag{49a}$$

$$\bar{R}(\theta_M) = \frac{1}{2M}\sum_{m=1}^{M}\int_{-\frac{1}{2}}^{\frac{1}{2}}\log^+\left[\lambda_m\left(e^{2\pi i\phi}\right)/\theta_M\right]d\phi, \tag{49b}$$

where $\lambda_1\left(e^{2\pi i\phi}\right) \le \cdots \le \lambda_M\left(e^{2\pi i\phi}\right)$ are the eigenvalues of the matrix with $(m+1,r+1)^{\text{th}}$ entry

$$\begin{aligned}
&S_{\bar{X}^m\bar{X}^r}\left(e^{2\pi i\phi}\right)\\
&\quad = \frac{1}{M}\sum_{n=0}^{M-1} S_{\bar{X}}^{r}\left(e^{2\pi i\frac{\phi-n}{M}}\right)e^{2\pi i(m-r)\frac{\phi-n}{M}}\\
&\quad = \frac{1}{T_0}\sum_{n=0}^{M-1}\sum_{k\in\mathbb{Z}} S_X^{rT_0/M}\left(\frac{\phi-n-kM}{T_0}\right)e^{2\pi i(m-r)\frac{\phi-n}{M}},\\
&\quad = \frac{1}{T_0}\sum_{l\in\mathbb{Z}} S_X^{rT_0/M}\left(\frac{\phi-l}{T_0}\right)e^{2\pi i(m-r)\frac{\phi-l}{M}}.
\end{aligned} \tag{50}$$

In order to express the code-rate in bits per time unit, we multiply the number of bits per sample $\bar{R}$ by the sampling rate M/T_0. This shows that the DRF of $\bar{X}[\cdot]$, as measured in bits pertime unit R, is given by (27).

In order to complete the proof we rely on the following lemma:

Lemma 8: Let $X(\cdot)$ be as in Theorem 2 and let $\bar{X}[\cdot]$ be its uniformly sampled version at rate M/T_0 as in (48)*. Denote the DRF at rate R bits per time unit of the two processes by $D(R)$ and $\bar{D}(R)$, respectively. Then*

$$\lim_{M\to\infty}\bar{D}(R) = D(R).$$

The rest of the appendix is devoted to the proof of Lemma 8.

Throughout the next steps it is convenient to use the covariance kernels $K(t,s) = R_X(s,t-s)$ and $\bar{K}[n,k] = R_{\bar{X}}[n,k-n]$. For $M\in\mathbb{N}$, define

$$\widetilde{K}(t,s) = K\left(\lfloor tM/T_0\rfloor T_0/M, \lfloor sM/T_0\rfloor T_0/M\right).$$

For any fixed $T>0$, the kernel $\widetilde{K}(t,s)$ defines an Hermitian positive compact operator [32] on the space of square integrable functions over $[-T,T]$. The eigenvalues of this operator are given by the Fredholm integral equation

$$\tilde{\lambda}_l\tilde{f}_l(t) = \frac{1}{2T}\int_{-T}^{T}\widetilde{K}(t,s)\tilde{f}_l(s)ds,\quad -T\le t\le T, \tag{51}$$

where it can be shown that there are at most MT/T_0 nonzero eigenvalues $\{\tilde{\lambda}_l\}$ that satisfy (51). We define the function $\tilde{D}_T(R)$ by the following parametric expression:

$$\begin{aligned}
\tilde{D}_T(\theta) &= \sum_{l=1}^{\infty}\min\left\{\tilde{\lambda}_l,\theta\right\}\\
R(\theta) &= \frac{1}{2}\sum_{l=1}^{\infty}\log^{+}\left(\frac{\tilde{\lambda}_l}{\theta}\right)
\end{aligned} \tag{52}$$

(the eigenvalues in (52) are implicitly depend on T). Note that

$$\sum_{l=1}^{\infty}\tilde{\lambda}_l = \frac{1}{2T}\int_{-T}^{T}\widetilde{K}(t,t)dt = \frac{1}{2T}\sum_{n=-N}^{N}K(nT_0/M, nT_0/M), \tag{53}$$

where $N = MT/T_0$. Expression (53) converges to

$$\frac{1}{2T}\int_{-T}^{T}K(t,t)dt \le \sigma_X^2$$

as M goes to infinity due to our assumption that $R(t,\tau)$ is Riemann integrable and therefore so is $K(t,s)$. Since we are interested in the asymptotic of large M, we can assume that (53) is bounded. This implies that $\tilde{D}_T(R)$ is bounded.

We would like to claim that the eigenvalues $\{\tilde{\lambda}_l\}$ approximate the eigenvalues $\{\lambda_l\}$. We have the following lemma:

Lemma 9: Let $\{\lambda_l\}$ and $\{\tilde{\lambda}_l\}$ be the eigenvalues in the Fredholm integral equation of $K(t,s)$ and $\widetilde{K}(t,s)$, respectively. Assume that these eigenvalues are numbered in a descending order. Then

$$\left|\lambda_l - \tilde{\lambda}_l\right| \le 4CT_0/M,\quad l=1,2,\ldots. \tag{54}$$

A. Proof of Lemma 9

Approximations of the kind (54) can be obtained by Weyl's inequalities for singular values of operators defined by self-adjoint kernels [33]. In our case it suffices to use the following result [34, Corollary 1"]:

$$\left|\lambda_l - \tilde{\lambda}_l\right| \le 2\sup_{t,s\in[-T,T]}\left|K(t,s)-\widetilde{K}(t,s)\right|,\quad l=1,2,\ldots. \tag{55}$$

The assumption that $R_X(t,\tau)$ is Lipschitz continuous in τ implies that there exists a constant $C>0$ such that for any $t_1,t_2,s\in\mathbb{R}$,

$$\begin{aligned}
\left|K(t_1,s)-K(t_2,s)\right| &= \left|R_X(s,t_1-s)-R_X(s,t_2-s)\right|\\
&\le C\left|t_1-t_2\right|.
\end{aligned}$$

We therefore conclude that $K_X(t,s)$ is Lipschitz continuous in both of its arguments from symmetry. Lipschitz continuity of $K(t,s)$ implies that

$$\begin{aligned}
&\left|K(t_1,s_1)-K(t_2,s_2)\right|\\
&\qquad\le \left|K(t_1,s_1)-K(t_1,s_2)\right| + \left|K(t_1,s_2)-\widetilde{K}(t_2,s_2)\right|\\
&\qquad\le C\left|s_1-s_2\right| + C\left|t_1-t_2\right|.
\end{aligned}$$

As a result, (55) leads to

$$\begin{aligned}
&\left|\lambda_l-\tilde{\lambda}_l\right|\\
&\le 2\sup_{t,s}\left|K(t,s)-\widetilde{K}(t,s)\right|\\
&= 2\sup_{t,s\in[-T,T]}\left|K(t,s)-K\left(\lfloor tM/T_0\rfloor T_0/M, \lfloor sM/T_0\rfloor T_0/M\right)\right|\\
&\le 2C\left(\left|t-\lfloor tM/T_0\rfloor T_0/M\right| + \left|t-\lfloor sM/T_0\rfloor T_0/M\right|\right)\\
&\le 4CT_0/M,
\end{aligned}$$

which proves Lemma 9.

The significance of Lemma 9 is that the eigenvalues of the kernel $K(t,s)$ used in the expression for the DRF of $X(\cdot)$ can be approximated by the eigenvalues of $\widetilde{K}(t,s)$, where the error in each of these approximations converge, uniformly in T, to zero as M increases. Since only a finite number of eigenvalues participate in (21) and since both $D_T(R)$ and $\tilde{D}_T(R)$ are bounded continuous functions of their eigenvalues, we conclude that $\tilde{D}_T(R)$ converges to $D_T(R)$ uniformly in T. Now let $\epsilon>0$ and fix M_0 large enough such that for all $M>M_0$ and for all T

$$\left|D_T(R)-\tilde{D}_T(R)\right| \le \epsilon. \tag{56}$$

Recall that in addition to (23), the DRF of $\bar{X}[\cdot]$, denoted here as $\bar{D}(\bar{R})$, can also be obtained as the limit in N of the expression

$$\bar{D}_N(\theta) = \sum_{l=1}^{\infty} \min\left\{\bar{\lambda}_l, \theta\right\}$$
$$\bar{R}(\theta) = \frac{1}{2}\sum_{l=1}^{\infty} \log^+\left(\bar{\lambda}_l/\theta\right),$$

where $\bar{\lambda}_1, \bar{\lambda}_2, \dots$ are the eigenvalues in the KL expansion of $\bar{X}$ over $n = -N, \dots, N$:

$$\bar{\lambda}_l f_l[n] = \frac{1}{2N+1}\sum_{k=-N}^{N} K_{\bar{X}}[n,k] f_l[k], \quad l = 1, \dots, N, \tag{57}$$

(there are actually at most $2N+1$ distinct non-zero eigenvalues that satisfies (57)). Letting $T_N = T_0 M/N$ and $\tilde{f}_l(t) = f_l\left(\lfloor t/T_0 \rfloor M\right)$ (57) can also be written as

$$\bar{\lambda}_l f_l[n] = \int_{-T_N}^{T_N} \widetilde{K}_X(nT_0/M, s) f_l\left[\lfloor s/T_0 \rfloor M\right] ds, \ l = 1, 2, \dots,$$
$$\bar{\lambda}_l \tilde{f}_l(t) = \int_{-T_n}^{T_N} \widetilde{K}(t,s) \tilde{f}_l(s) ds, \ \ -T_N < t < T_N.$$

From the uniqueness of the KL expansion, we obtain that for any N, the eigenvalues of $\widetilde{K}(t,s)$ over $T_N = T_0 M/N$ are given by the eigenvalues of $\bar{K}[n,k]$ over $-N, \dots, N$. We conclude that

$$\bar{D}_N(\bar{R}) = \tilde{D}_{T_N}(R), \tag{58}$$

where $R = \bar{R} T_0/M$. Now take N large enough such that

$$\left|\bar{D}_N(R) - \bar{D}(R)\right| < \epsilon,$$

and

$$\left|D_{T_N}(R) - D(R)\right| < \epsilon.$$

For all $M \geq M_0$ we have

$$\begin{aligned}\left|D(R) - \bar{D}(R)\right| &= \Big|D(R) - D_{T_N}(R) + D_{T_N}(R) + \tilde{D}_{T_N}(R) \\ &\quad - \tilde{D}_{T_N}(R) + \bar{D}_N(R) - \bar{D}_N(R) - \bar{D}(R)\Big| \\ &\leq \left|D(R) - D_{T_N}(R)\right| \quad (59)\\ &\quad + \left|D_{T_N}(R) - \tilde{D}_{T_N}(R)\right| \quad (60)\\ &\quad + \left|\tilde{D}_{T_N}(R) - \bar{D}_N(R)\right| \quad (61)\\ &\quad + \left|\bar{D}_N(R) - \bar{D}(R)\right| \leq 3\epsilon, \quad (62)\end{aligned}$$

where the last transition is because: (59) and (62) are smaller than ϵ by the choice of N, (60) is smaller than ϵ from (56). and (61) equals zero from (58).

Appendix C

In this Appendix we prove Proposition 3. We use the process $\bar{X}[\cdot]$ defined in the proof of Theorem 2 as the uniform sampled version of $X(\cdot)$ at rate T_0/M. From Proposition 1 we conclude that the DRF of $\bar{X}[\cdot]$ satisfies

$$D_{\bar{X}}(\bar{R}) \geq \frac{1}{M}\sum_{m=0}^{M-1}\int_{-\frac{1}{2}}^{\frac{1}{2}} \min\left\{\frac{1}{T_0}\sum_{l\in\mathbb{Z}} S_X^{mT_0/M}\left(\frac{\phi-l}{T_0}\right), \theta_m\right\} d\phi, \tag{63}$$

where for all $m = 0, \dots, M-1$, θ_m is determined by

$$\bar{R} = \frac{1}{2}\int_{-\frac{1}{2}}^{\frac{1}{2}} \log^+\left[\frac{1}{T_0}\sum_{l\in\mathbb{Z}} S_X^{mT_0/M}\left(\frac{\phi-l}{T_0}\right)/\theta_m\right] d\phi.$$

Denote $t = mT_0/M$. As M approaches infinity, the RHS of (63) converges to an integral with respect to t over the interval $(0, T_0)$, which implies

$$\bar{D}(\bar{R}) \geq \frac{1}{T_0}\int_0^{T_0}\int_{-\frac{1}{2}}^{\frac{1}{2}} \min\left\{\sum_{l\in\mathbb{Z}} S_X^{\,t}\left(\frac{\phi-l}{T_0}\right), \theta_t\right\} d\phi, \tag{64}$$

and

$$\bar{R} = \frac{1}{2}\int_{-\frac{1}{2}}^{\frac{1}{2}} \log^+\left[\sum_{l\in\mathbb{Z}} S_X^{mT_0/M}\left(\frac{\phi-l}{T_0}\right)/\theta_m\right] d\phi, \tag{65}$$

where we denoted $\theta_t = T_0\theta_m$. In order to go from $\bar{R}$ to R we multiply (65) by M/T_0, so that (64) and (65) lead to (33). The fact that the function $\bar{D}(R)$ converges to $D(R)$ as M goes to infinity follows from the proof of Theorem. 2.

Appendix D

In this Appendix we provide a proof of Proposition 4. The entries of the matrix $\mathbf{S}\left(e^{2\pi i\phi}\right)$ in Theorem 2 are obtained by using the CPSD of the PAM process (17) in (28). For all $M \in \mathbb{N}$, this leads to

$$\begin{aligned}&\mathbf{S}_{m+1,r+1}\left(e^{2\pi i\phi}\right) \\ &= \frac{1}{T_0^2}\sum_{k\in\mathbb{Z}}\Bigg[P\left(\frac{\phi-k}{T_0}\right)S_U\left(\frac{\phi-k}{T_0}\right)e^{2\pi i(\phi-k)\frac{m-r}{M}} \\ &\qquad \times \sum_{n\in\mathbb{Z}} P^*\left(\frac{\phi-n-k}{T_0}\right)e^{2\pi i\frac{nr}{M}}\Bigg] \quad (66)\\ &= \frac{1}{T_0^2}\sum_{k\in\mathbb{Z}} P\left(\frac{\phi-k}{T_0}\right)S_U\left(\frac{\phi-k}{T_0}\right)e^{2\pi i(\phi-k)\frac{m}{M}} \\ &\qquad \times \sum_{l\in\mathbb{Z}} P^*\left(\frac{\phi-l}{T_0}\right)e^{-2\pi i(\phi-l)\frac{r}{M}}. \quad (67)\end{aligned}$$

The expression (67) consists of the product of a term depending only on m and a term depending only on r. We conclude that the matrix $\mathbf{S}\left(e^{2\pi i\phi}\right)$ can be written as the outer product of two M dimensional vectors, and thus it is of rank one. The single non-zero eigenvalue $\lambda_M\left(e^{2\pi i\phi}\right)$ of $\mathbf{S}\left(e^{2\pi i\phi}\right)$ is given by the trace of the matrix, which, by the orthogonality of the functions $e^{2\pi i\frac{nr}{M}}$ in (66), is evaluated as

$$\lambda_M\left(e^{2\pi i\phi}\right) = \frac{M}{T_0^2}\sum_{k\in\mathbb{Z}}\left|P\left(\frac{\phi-k}{T_0}\right)\right|^2 S_U\left(\frac{\phi-k}{T_0}\right). \tag{68}$$

We now use (68) in (27). In order to obtain (40), we change the integration variable from ϕ to $f = \phi/T_0$ and the water-level parameter θ to $T_0\theta/M$. Note that the final expression is independent of M, so that the limit in (29) is already given by this expression.

APPENDIX E

We now provide a proof of Proposition 6. Since $S_U(f)$ is compactly supported, the covariance function $R_U(\tau) = \mathbb{E}U(t+\tau)U(t)$ is an analytic function and therefore Lipsshitz continuous. Lipschitz continuity of $R_U(\tau)$ implies Lipschitz continuity of $R_X(t,\tau)$ in its second argument and therefore Theorem 2 applies: The DRF of the Gaussian CSP $X_\varphi(\cdot)$ with period $T_0 = f_0^{-1}$ is obtained by using Theorem 2 with the SCD (15). For all $M \in \mathbb{N}$ and $m, r = 0, \ldots, M-1$ we have,

$$\begin{aligned}\mathbf{S}_{m+1,r+1}&\left(e^{2\pi i\phi}\right)\\ &= \frac{f_0}{2}\sum_{k\in\mathbb{Z}} S_U\left(f_0(\phi-k-1)\right)\left(1+e^{4\pi i r/M}\right)e^{2\pi i(\phi-k)\frac{m-r}{M}}\\ &\quad+\frac{f_0}{2}\sum_{k\in\mathbb{Z}} S_U\left(f_0(\phi-k+1)\right)\left(1+e^{-4\pi i r/M}\right)e^{2\pi i(\phi-k)\frac{m-r}{M}}.\end{aligned}$$

Under the assumption that $f_0 > 2f_B$ we have that for all $\phi \in \left(-\frac{1}{2},\frac{1}{2}\right)$, $S_U\left(f_0(\phi-k\pm 1)\right) = 0$ for all $k \neq \pm 1$. This leads to

$$\begin{aligned}\mathbf{S}_{m+1,r+1}&\left(e^{2\pi i\phi}\right) \qquad (69)\\ &= S_U(f_0\phi)\frac{f_0\left(1+e^{4\pi i r/M}\right)}{2}e^{2\pi i(\phi+1)\frac{m-r}{M}}\\ &\quad+S_U(f_0\phi)\frac{f_0\left(1+e^{-4\pi i r/M}\right)}{2}e^{2\pi i(\phi-1)\frac{m-r}{M}}\\ &= 2f_0S_U(f_0\phi)e^{2\pi i\frac{m-r}{M}\phi}\cos\left(2\pi\frac{m}{M}\right)\cos\left(2\pi\frac{r}{M}\right). \qquad (70)\end{aligned}$$

From (70) we conclude that the matrix $\mathbf{S}\left(e^{2\pi i\phi}\right)$ can be written as

$$\mathbf{S}\left(e^{2\pi i\phi}\right) = 2f_0S_U(f_0\phi)\mathbf{S}_M\left(e^{2\pi i\phi}\right)\mathbf{S}_M^*\left(e^{2\pi i\phi}\right),$$

where $\mathbf{S}_M\left(e^{2\pi i\phi}\right) \in \mathbf{R}^{M\times 1}$ is given by

$$\left(1, e^{2\pi i\phi/M}\cos\left(\frac{2\pi}{M}\right), \ldots, e^{2\pi i\phi\frac{M-1}{M}}\cos\left(\frac{2\pi(M-1)}{M}\right)\right).$$

This means that $\mathbf{S}\left(e^{2\pi i\phi}\right)$ is a matrix of rank one, and its single non-zero eigenvalue is given by its trace:

$$\lambda_M\left(e^{2\pi i\phi}\right) = 2f_0S_U(f_0\phi)\sum_{m=0}^{M-1}\cos^2(2\pi m/M) = Mf_0S_U(f_0\phi).$$

We use this in (28):

$$\begin{aligned}D_M(R_\theta) &= \frac{f_0}{M}\int_{-\frac{1}{2}}^{\frac{1}{2}}\min\left\{Mf_0S_U(f_0\phi),\theta\right\}d\phi\\ &= f_0\int_{-\infty}^{\infty}\min\left\{S_U(f_0\phi),\theta/M\right\}d\phi\\ &= \int_{-\infty}^{\infty}\min\left\{S_U(f),\theta/M\right\}df, \qquad (71)\end{aligned}$$

where

$$\begin{aligned}R_\theta &= \frac{f_0}{2}\sum_{m=1}^{M}\int_{-\frac{1}{2}}^{\frac{1}{2}}\log^+\left[\lambda_m\left(e^{2\pi\phi}\right)\right]d\phi\\ &= \frac{f_0}{2}\int_{-\frac{1}{2}}^{\frac{1}{2}}\log^+\left[Mf_0S_U(f_0\phi)/\theta\right]d\phi\\ &= \frac{1}{2}\int_{-\infty}^{\infty}\log^+\left[S_U(f)/(\theta/M)\right]df. \qquad (72)\end{aligned}$$

From (71) and (72) we conclude that for every M, the parametric expression of D as a function of R is identical to the DRF of the stationary process $U(\cdot)$ given by (22).

ACKNOWLEDGMENT

The authors wish to thank Robert M. Gray for helpful remarks. They also wish to thank the anonymous reviewers and the AE for the depth and breadth of their reviews which have greatly improved the paper.

REFERENCES

[1] T. Berger, "Information rates of Wiener processes," *IEEE Trans. Inf. Theory*, vol. 16, no. 2, pp. 134–139, Mar. 1970.

[2] R. Gray, "Information rates of autoregressive processes," *IEEE Trans. Inf. Theory*, vol. 16, no. 4, pp. 412–421, Jul. 1970.

[3] W. A. Gardner, A. Napolitano, and L. Paura, "Cyclostationarity: Half a century of research," *Signal Process.*, vol. 86, no. 4, pp. 639–697, Apr. 2006.

[4] W. R. Bennett, "Statistics of regenerative digital transmission," *Bell Syst. Tech. J.*, vol. 37, no. 6, pp. 1501–1542, Nov. 1958.

[5] J. Nedoma, "On the ergodicity and R-ergodicity of stationary probability measures," *Zeitschrift Wahrscheinlichkeitstheorie*, vol. 2, pp. 90–97, 1963.

[6] R. M. Gray, *Entropy and Information Theory*. New York, NY, USA: Springer-Verlag, 1990.

[7] R. G. Gallager, *Information Theory and Reliable Communication*. Hoboken, NJ, USA: Wiley, 1968.

[8] T. Berger and J. D. Gibson, "Lossy source coding," *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2693–2723, Oct. 1998.

[9] S. Haykin, *An Introduction to Analog and Digital Communications*. Hoboken, NJ, USA: Wiley, 2009.

[10] T. Berger, *Rate Distortion Theory: A Mathematical Basis for Data Compression*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1971.

[11] R. M. Gray, *Probability, Random Processes, and Ergodic Properties*. New York, NY, USA: Springer, 2009.

[12] A. Kolmogorov, "On the Shannon theory of information transmission in the case of continuous signals," *IRE Trans. Inf. Theory*, vol. 2, no. 4, pp. 102–108, Dec. 1956.

[13] M. Vetterli, J. Kovačević, and V. K. Goyal, *Foundations of Signal Processing*. Cambridge, U.K.: Cambridge Univ. Press, 2014.

[14] T. Weissman and E. Ordentlich, "The empirical distribution of rate-constrained source codes," *IEEE Trans. Inf. Theory*, vol. 51, no. 11, pp. 3718–3733, Nov. 2005.

[15] A. Kanlis, "Compression and transmission of information at multiple resolutions," Ph.D. dissertation, Dept. Elect. Eng., Univ. Maryland, College Park, College Park, MD, USA, 1997.

[16] A. Kipnis, S. Rini, and A. J. Goldsmith. (2017). "Compress and estimate in multiterminal source coding." [Online]. Available: http://arxiv.org/abs/1602.02201

[17] W. A. Gardner, *Cyclostationarity in Communications and Signal Processing* (Electrical Engineering, Communications and Signal Processing). Piscataway, NJ, USA: IEEE Press, 1994.

[18] A. Papoulis and S. U. Pillai, *Probability, Random Variables and Stochastic Processes* (McGraw-Hill Electrical and Electronic Engineering Series). New York, NY, USA: McGraw-Hill, 2002.

[19] W. Gardner, "Spectral correlation of modulated signals: Part I—Analog modulation," *IEEE Trans. Commun.*, vol. 35, no. 6, pp. 584–594, Jun. 1987.

[20] T. Berger, "Rate distortion theory for sources with abstract alphabets and memory," *Inf. Control*, vol. 13, no. 3, pp. 254–273, 1968.

[21] R. M. Gray, *Entropy and Information Theory*, vol. 1. New York, NY, USA: Springer, 2011.

[22] I. Karatzas and S. Shreve, *Brownian Motion and Stochastic Calculus* (Graduate Texts in Mathematics). New York, NY, USA: Springer, 1991.

[23] V. Kafedziski, "Rate distortion of stationary and nonstationary vector Gaussian sources," in *Proc. IEEE/SP 13th Workshop Statist. Signal Process.*, Jul. 2005, pp. 1054–1059.

[24] U. Grenander and G. Szegő, *Toeplitz Forms and Their Applications* (California Monographs in Mathematical Sciences). Berkeley, CA, USA: Univ. California Press, 1958.

[25] P. Tilli, "Singular values and eigenvalues of non-Hermitian block Toeplitz matrices," *Linear Algebra Appl.*, vol. 272, nos. 1–3, pp. 59–69, Mar. 1996.

[26] Y. Chen, Y. C. Eldar, and A. J. Goldsmith, "Shannon meets Nyquist: Capacity of sampled Gaussian channels," *IEEE Trans. Inf. Theory*, vol. 59, no. 8, pp. 4889–4914, Aug. 2013.

[27] A. Kipnis, A. J. Goldsmith, Y. C. Eldar, and T. Weissman, "Distortion rate function of sub-Nyquist sampled Gaussian sources," *IEEE Trans. Inf. Theory*, vol. 62, no. 1, pp. 401–429, Jan. 2016.

[28] R. Dobrushin and B. Tsybakov, "Information transmission with additional noise," *IRE Trans. Inf. Theory*, vol. 8, no. 5, pp. 293–304, Sep. 1962.

[29] J. K. Wolf and J. Ziv, "Transmission of noisy information to a noisy receiver with minimum distortion," *IEEE Trans. Inf. Theory*, vol. 16, no. 4, pp. 406–411, Jul. 1970.

[30] M. B. Matthews, "On the linear minimum-mean-squared-error estimation of an undersampled wide-sense stationary random process," *IEEE Trans. Signal Process.*, vol. 48, no. 1, pp. 272–275, Jan. 2000.

[31] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, no. 3, pp. 379–423, Jul./Oct. 1948.

[32] H. König, *Eigenvalue Distribution of Compact Operators* (Operator Theory: Advances and Applications). Basel, Switzerland: Birkhäuser, 2013.

[33] H. Weyl, "Das asymptotische verteilungsgesetz der eigenwerte linearer partieller differentialgleichungen (mit einer anwendung auf die theorie der hohlraumstrahlung)," *Math. Ann.*, vol. 71, no. 4, pp. 441–479, 1912.

[34] H. Wielandt, "Error bounds for eigenvalues of symmetric integral equations," in *Proc. Symp. Appl. Math.*, vol. 6. 1956, pp. 261–282.

Alon Kipnis (S'14) received the B.Sc. degree in mathematics (*summa cum laude*) and the B.Sc. degree in electrical engineering (*summa cum laude*) in 2010, and the M.Sc. in mathematics in 2012, all from Ben-Gurion University of the Negev. He is currently a Ph.D. candidate in the department of electrical engineering at Stanford University. His research interests include information theory, signal processing and statistics.

Andrea J. Goldsmith (S'90–M'93–SM'99–F'05) Andrea Goldsmith is the Stephen Harris professor in the School of Engineering and a professor of Electrical Engineering at Stanford University. She was previously on the faculty of Electrical Engineering at Caltech. Her research interests are in information theory and communication theory, and their application to wireless communications and related fields. She co-founded and served as Chief Scientist of Plume WiFi (formerly Accelera), and previously co-founded and served as CTO of Quantenna Communications, Inc. (QNTA). She has also held industry positions at Maxim Technologies, Memorylink Corporation, and AT&T Bell Laboratories and currently serves on the technical advisory boards of several public and private companies. Dr. Goldsmith is a member of the National Academy of Engineering and the American Academy of Arts and Sciences. She is also a Fellow of the IEEE and of Stanford, and has received several awards for her work, including the IEEE ComSoc Edwin H. Armstrong Achievement Award as well as Technical Achievement Awards in Communications Theory and in Wireless Communications, the National Academy of Engineering Gilbreth Lecture Award, the IEEE ComSoc and Information Theory Society Joint Paper Award, the IEEE ComSoc Best Tutorial Paper Award, the Alfred P. Sloan Fellowship, the WICE Technical Achievement Award, and the Silicon Valley/San Jose Business Journal's Women of Influence Award. She is author of the book *Wireless Communications* and co-author of the books *MIMO Wireless Communications* and *Principles of Cognitive Radio*, all published by Cambridge University Press, as well as an inventor on 29 patents. She received the B.S., M.S. and Ph.D. degrees in Electrical Engineering from U.C. Berkeley.

Dr. Goldsmith has served on the Steering Committee for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS and as editor for the IEEE TRANSACTIONS ON INFORMATION THEORY, the Journal on *Foundations and Trends in Communications and Information Theory and in Networks*, the IEEE TRANSACTIONS ON COMMUNICATIONS, and the *IEEE Wireless Communications Magazine*. She participates actively in committees and conference organization for the IEEE Information Theory and Communications Societies and has served on the Board of Governors for both societies. She has also been a Distinguished Lecturer for both societies, served as President of the IEEE Information Theory Society in 2009, founded and chaired the student committee of the IEEE Information Theory society, and chaired the Emerging Technology Committee of the IEEE Communications Society. She currently chairs the IEEE TAB committee on diversity and inclusion, and the Women in Technology Leadership Roundtable working group on metrics. At Stanford she received the inaugural University Postdoc Mentoring Award, served as Chair of Stanford's Faculty Senate in 2009 and currently serves on its Faculty Senate, Budget Group, and University Advisory Board.

Yonina C. Eldar (S'98–M'02–SM'07–F'12) received the B.Sc. degree in Physics in 1995 and the B.Sc. degree in Electrical Engineering in 1996 both from Tel-Aviv University (TAU), Tel-Aviv, Israel, and the Ph.D. degree in Electrical Engineering and Computer Science in 2002 from the Massachusetts Institute of Technology (MIT), Cambridge.

She is currently a Professor in the Department of Electrical Engineering at the Technion - Israel Institute of Technology, Haifa, Israel, where she holds the Edwards Chair in Engineering. She is also a Research Affiliate with the Research Laboratory of Electronics at MIT, an Adjunct Professor at Duke University, and was a Visiting Professor at Stanford University, Stanford, CA. She is a member of the Israel Academy of Sciences and Humanities (elected 2017), an IEEE Fellow and a EURASIP Fellow. Her research interests are in the broad areas of statistical signal processing, sampling theory and compressed sensing, optimization methods, and their applications to biology and optics.

Dr. Eldar has received many awards for excellence in research and teaching, including the IEEE Signal Processing Society Technical Achievement Award (2013), the IEEE/AESS Fred Nathanson Memorial Radar Award (2014), and the IEEE Kiyo Tomiyasu Award (2016). She was a Horev Fellow of the Leaders in Science and Technology program at the Technion and an Alon Fellow. She received the Michael Bruno Memorial Award from the Rothschild Foundation, the Weizmann Prize for Exact Sciences, the Wolf Foundation Krill Prize for Excellence in Scientific Research, the Henry Taub Prize for Excellence in Research (twice), the Hershel Rich Innovation Award (three times), the Award for Women with Distinguished Contributions, the Andre and Bella Meyer Lectureship, the Career Development Chair at the Technion, the Muriel & David Jacknow Award for Excellence in Teaching, and the Technions Award for Excellence in Teaching (two times). She received several best paper awards and best demo awards together with her research students and colleagues including the SIAM outstanding Paper Prize, the UFFC Outstanding Paper Award, the Signal Processing Society Best Paper Award and the IET Circuits, Devices and Systems Premium Award, and was selected as one of the 50 most influential women in Israel.

She was a member of the Young Israel Academy of Science and Humanities and the Israel Committee for Higher Education. She is the Editor in Chief of *Foundations and Trends in Signal Processing*, a member of the IEEE Sensor Array and Multichannel Technical Committee and serves on several other IEEE committees. In the past, she was a Signal Processing Society Distinguished Lecturer, member of the IEEE Signal Processing Theory and Methods and Bio Imaging Signal Processing technical committees, and served as an associate editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING, the *EURASIP Journal of Signal Processing*, the *SIAM Journal on Matrix Analysis and Applications*, and the *SIAM Journal on Imaging Sciences*. She was Co-Chair and Technical Co-Chair of several international conferences and workshops.

She is author of the book *Sampling Theory: Beyond Bandlimited Systems* and co-author of the books *Compressed Sensing* and *Convex Optimization Methods in Signal Processing and Communications*, all published by Cambridge University Press.