On the $\epsilon$-Entropy and the Rate-Distortion Function of Certain Non-Gaussian Processes

JACOB BINIA, MOSHE ZAKAI, FELLOW, IEEE AND JACOB ZIV, FELLOW, IEEE

Abstract—Let $\xi = \{\xi(t), 0 \leq t \leq T\}$ be a process with covariance function $K(s,t)$ and $E_1^\infty \xi'(t) dt < \infty$. It is proved that for every $\epsilon > 0$ the $\epsilon$-entropy $H_\epsilon(\xi)$ satisfies

$$H_\epsilon(\xi) - H_{e,\epsilon}(\xi) \leq H_{\epsilon}(\xi) \leq H_\epsilon(\xi)e$$

where $\xi$ is a Gaussian process with the covariance $K(s,t)$ and $H_{e,\epsilon}(\xi)$ is the entropy of the measure induced by $\xi$ (in function space) with respect to that induced by $\xi$. It is also shown that if $H_{\epsilon}(\xi) < \infty$, then, as $\epsilon \to 0$

$$H_\epsilon(\xi) = H_{e,\epsilon}(\xi) - H_{e,\epsilon}(\xi) + o(1)$$

Furthermore, if there exists a Gaussian process $g = \{g(t); 0 \leq t \leq T\}$ such that $H_\epsilon(\xi) < \infty$, then the ratio between $H_\epsilon(\xi)$ and $H_\epsilon(g)$ goes to one as $\epsilon$ goes to zero. Similar results are given for the rate-distortion function, and some particular examples are worked out in detail. Some cases for which $H_{e,\epsilon}(\xi) = \infty$ are discussed, and asymptotic bounds on $H_\epsilon(\xi)$, expressed in terms of $H_{e,\epsilon}(\xi)$, are derived.

I. INTRODUCTION

THE $\epsilon$-ENTROPY (and its related normalized form, the rate-distortion function) provides an important mathematical tool for the analysis of communication sources and systems. Given a communication system, the $\epsilon$-entropy of the source and the channel capacity yield a lower bound on the minimum attainable distortion [1].

Let $\xi$ be a real-valued random variable and denote by $H_\epsilon(\xi)$ the $\epsilon$-entropy (rate-distortion function) of $\xi$, relative to a mean-square-error criterion. A well-known result of Shannon [2], [3] states that, if the probability distribution of $\xi$ possesses a density, then

$$h(\xi) + \frac{1}{2\pi\epsilon} \ln \frac{1}{2\pi\epsilon^2} \leq H_\epsilon(\xi) \leq h(\xi) + \frac{1}{2\pi\epsilon} \ln \frac{1}{2\pi\epsilon^2} + o(1)$$

where $\sigma^2$ denotes the variance of $\xi$ and $h(\xi) = -\int p_x(x) \ln p_x(x) dx$. Furthermore, it was shown by Gerrish and Schultheiss [4] that as $\epsilon \to 0$

$$H_\epsilon(\xi) = h(\xi) + \frac{1}{2\pi\epsilon} \ln \frac{1}{2\pi\epsilon^2} + o(1).$$

Results of the type (1) and (2) have been extended for $N$-dimensional random variables, provided that the random variables possess a probability density [1], [4], and [5].

The purpose of this paper is to derive results of the type (1) and (2) for random processes. A direct extension of (1) and (2) to random processes or to infinite-dimensional random variables is impossible, since the results are based on the existence of a probability density. Namely, the probability measure of $\xi$ is required to be absolutely continuous with respect to the Lebesgue measure. In this paper we replace this requirement by a requirement of absolute continuity with respect to a Gaussian measure. This enables us to prove the following bounds on the $\epsilon$-entropy of a random process $\xi = \{\xi(t), 0 \leq t \leq T\}$:

$$H_\epsilon(\xi) - H_{e,\epsilon}(\xi) \leq H_\epsilon(\xi) \leq H_{e,\epsilon}(\xi)$$

(3)

where $\xi$ is the Gaussian process with the same covariance as that of $\xi$ and $H_{e,\epsilon}(\xi)$ is the relative entropy of the measure induced by $\xi$ with respect to that induced by $\xi$ (cf., Section II). Furthermore, if $H_{e,\epsilon}(\xi) < \infty$, then we show that for $\epsilon \to 0$

$$H_\epsilon(\xi) = H_{e,\epsilon}(\xi) - H_{e,\epsilon}(\xi) + o(1).$$

(4)

In fact, for a finite-dimensional random variable $\xi$, the preceding lower bound on $H_\epsilon(\xi)$, as well as the asymptotic behavior (4) are stronger than previously known results (the upper bound on $H_\epsilon(\xi)$, for $N$-dimensional random variables, is already known; [1, sec. 4.6.2].

The results in (3) and (4) on $H_\epsilon(\xi)$ are expressed in terms of the $\epsilon$-entropy of a related Gaussian process $H_{e,\epsilon}(\xi)$ and the entropy of the measure of $\xi$ with respect to the measure of the related Gaussian process $\xi$. Results on the $\epsilon$-entropy of Gaussian processes are available in the literature [1], [6]-[10]. The entropy $H_{e,\epsilon}(\xi)$ of $\xi$ with respect to some Gaussian processes $g$ can be derived from known results on the Radon–Nikodym derivative of certain processes with respect to certain Gaussian processes [11]-[13]. It is shown that if there exists a Gaussian process $g$ for which $H_{e,\epsilon}(\xi) < \infty$, then $H_{e,\epsilon}(\xi) < \infty$. Therefore the bounds (3) also yield asymptotic results, for $\epsilon \to 0$. Several examples are given, and in one of these examples we show that if $d(\xi(t) = \xi(t) dt + \beta dw(t)$ (where $w(t)$ is a standard Brownian motion) then as $\epsilon \to 0$

$$H_\epsilon(\xi) \approx \frac{2T^2\beta^2}{\pi^2 \epsilon}$$

which is a generalization of previously known results [1], [10].

Section II is devoted to notation and certain preliminary results. The main results and examples are given in Section III, and their proofs are given in Section IV.
II. PRELIMINARIES AND NOTATION

Let $P_1$ and $P_2$ be two probability measures defined on a measurable space $(\Omega, \mathcal{B})$ and let $\{E_i\}$ be a finite $\mathcal{B}$-measurable partition of $\Omega$. The entropy $H_{P_2}(P_1)$ of $P_1$ with respect to $P_2$ is defined by [14, ch. 2]

$$H_{P_2}(P_1) = \sup \sum P_1(E_i) \ln \frac{P_1(E_i)}{P_2(E_i)}$$

where the supremum is taken over all partitions of $\Omega$. Let $P_1$ and $P_2$ be the distributions of the random variables $\xi$ and $\eta$ (taking values in the same space $(X, \mathcal{B}_X)$), respectively. In this case the notation for the entropy will be

$$H_{P_2}(\xi) = H_{P_2}(P_1) = H_{P_2}(P_1).$$

Let $\xi$ and $\eta$ be random variables with values in the measurable spaces $(X, \mathcal{B}_X)$ and $(Y, \mathcal{B}_Y)$, respectively. The mutual information $I(\xi, \eta)$ between these variables is defined as [14, ch. 2]

$$I(\xi, \eta) = \sup \sum P_{\xi,\eta}(E_i \times F_j) \ln \frac{P_{\xi,\eta}(E_i \times F_j)}{P_\xi(E_i) P_\eta(F_j)}$$

where the supremum is taken over all partitions $\{E_i\}$ of $X$ and $\{F_j\}$ of $Y$.

The $\varepsilon$-entropy $H_{\varepsilon}(\xi)$ of a random variable $\xi$ taking values in a metric space $(X, \mathcal{B}_X)$ with a metric $\rho(x,y)$, $x, y \in X$, is defined as

$$H_{\varepsilon}(\xi) = \inf E\beta^2(\xi, \eta)$$

where the infimum is taken over all joint distributions $P_{\xi,\eta}$ (defined on $(X \times X, \mathcal{B}_X \times \mathcal{B}_X)$), with a fixed marginal distribution $P_\xi$, such that

$$E\beta^2(\xi, \eta) \leq \varepsilon^2.$$  

The following lemma shows that the $\varepsilon$-entropy is independent of the representation of a process $\xi$ in different isometric spaces.

**Lemma 2.1:** Let $A$ be a linear, one-to-one and bi-measurable transformation from the metric space $(X, \mathcal{B}_X)$ onto the metric space $(\tilde{X}, \mathcal{B}_{\tilde{X}})$ and let $P_\xi$ be the distribution of $\xi$ in $(X, \mathcal{B}_X)$. We define the distribution $P_{\tilde{\xi}}$ of $\tilde{\xi}$ as $P_{\tilde{\xi}}(E) = P_\xi(E^{-1})$, where $E^{-1} = A^{-1}(E)$, $E \in \mathcal{B}_{\tilde{X}}$. If the metrics $\rho$ and $\tilde{\rho}$, defined in $X$ and $\tilde{X}$, respectively, satisfy

$$\tilde{\rho}(Ax, Ay) = \rho(x, y), \quad x, y \in X$$

then

$$H_{\varepsilon}(\tilde{\xi}) = H_{\varepsilon}(\xi).$$

The proof of Lemma 2.1 follows directly from the transformation properties, since such a transformation preserves the amount of information [15].

In particular, let $\xi = \{\xi(t), 0 \leq t \leq T\}$ be a stochastic process with $E \int_0^T \xi^2(t) dt < \infty$. Let $\{\phi(t)\}$ be a complete orthonormal system in $L^2[0,T]$, and consider the following transformation $A$ from $L^2[0,T]$ to $l^2$:

$$A: x(t) \rightarrow (x_1, x_2, \cdots), \quad x(t) \in L^2[0, T]$$

where

$$x_i = \int_0^T x(t) \phi_i(t) dt, \quad i = 1, 2, \cdots$$

Then we have

$$H_{A}(\xi) = H_{\varepsilon}(\xi_1, \xi_2, \cdots).$$

**Remarks:** a) From now on, the discussion is limited to the metrics of $L^2$. Although the results will be stated for processes in $L^2[0,T]$, they hold for all processes defined in spaces isometric to $L^2[0,T]$. b) The random function $\xi(t), 0 \leq t \leq T$ will be assumed to be measurable, separable, and satisfying $E \int_0^T \xi^2(t) dt < \infty$. Also, since the $\varepsilon$-entropy is independent of the process mean, we always assume $E\xi(t) = 0$.

**Lemma 2.2:** Let $\xi, \varepsilon, n = 1, 2, \cdots$ be a sequence of random processes such that

$$\lim_{n \rightarrow \infty} E \int_0^T [\xi(t) - \varepsilon_n(t)]^2 dt = 0.$$  

Then for every $\varepsilon > 0$

$$\lim_{n \rightarrow \infty} H_{\xi}(\varepsilon_n) = H_{\xi}(\xi).$$

**Proof:** We first show that if $\xi_1$ and $\xi_2$ are random processes such that

$$E \int_0^T [\xi_1(t) - \xi_2(t)]^2 dt \leq \theta$$

then

$$H_{\varepsilon_1}(\xi_2) \leq H_{\varepsilon_1}(\xi_1) \leq H_{\varepsilon_1}(\xi_2), \quad \varepsilon \geq \theta > 0.$$  

By the definition of the $\varepsilon$-entropy, there exists a process $\eta$ such that

$$E \int_0^T [\xi_1(t) - \eta(t)]^2 dt \leq \varepsilon^2$$

and

$$H_{\varepsilon_1}(\xi_1) \geq I(\xi_1, \eta) - \delta$$

where $\delta > 0$ is arbitrary. Note that we can choose a version of the preceding process $\eta$ such that $\eta$ and $\xi_2$ are conditionally independent, conditioned on $\xi_1$; therefore [14, sect. 3.4]

$$I(\xi_2, \eta) \leq I((\xi_2, \xi_1), \eta) = I(\xi_1, \eta) + E\xi(\xi_2, \eta | \xi_1) = I(\xi_1, \eta).$$

(11)

Furthermore, by (7) and (9)

$$E \int_0^T [\xi_2(t) - \eta(t)]^2 dt \leq (\varepsilon + \theta)^2.$$  

(12)

Therefore, by (10)-(12)

$$H_{\varepsilon_1}(\xi_2) \leq I(\xi_2, \eta) \leq I(\xi_1, \eta) \leq H_{\varepsilon_1}(\xi_1) \leq \delta.$$  

This completes the proof of the left side of (8); the right side of (8) follows by a similar argument.

Equation (6) follows from (5), (8) and the continuity of $H_{\varepsilon}(\xi)$ as a function of $\varepsilon$.

**Remark:** Let $\xi = \{\xi(t), 0 \leq t \leq T\}$ be a process with $E \int_0^T \xi^2(t) dt < \infty$. Suppose that $\xi$ is passed through a linear filter with transfer function $h(t)$, given by

$$h(t) = \begin{cases} \frac{1}{\delta}, & 0 \leq t \leq \delta \\ 0, & \text{otherwise} \end{cases}$$
Denote by $\xi_d$ the filter output in $[0,T]$. Then obviously $\xi_d$ is a quadratic-mean-continuous process and

$$
\lim_{\delta \to 0} \frac{1}{\delta} \int_{0}^{T} [\xi(t) - \xi_d(t)]^2 \, dt = 0.
$$

This result will allow us to assume throughout the proofs that $\xi$ is mean-continuous, and the results will remain valid without this assumption by Lemma 2.2.

Finally, the rate-distortion function $R_\epsilon(D)$ of a process $\xi = \{\xi(t), 0 \leq t < \infty\}$ is defined as ([1], [16])

$$
R_\epsilon(D) = \lim_{T \to \infty} \frac{1}{T} H_{\epsilon}(\xi_0^T)
$$

where $\xi_0^T = \{\xi(t), 0 \leq t \leq T\}$. Pinsker [17] proved, under some general conditions, that the rate-distortion function $R_\epsilon(D)$ of a stationary process $\xi$ is well defined and finite for $D > 0$.

The importance of the rate-distortion function in information theory stems from the following facts: 1) it provides general lower bounds on the performance of communication systems and 2) in some cases a coding theorem holds to the effect that the performance bounds derived from the rate-distortion function can be approached asymptotically by using appropriate delays [1].

III. MAIN RESULTS AND EXAMPLES

Let $\xi = \{\xi(t), 0 \leq t \leq T\}$ be a random process with a covariance function $K_\epsilon(s,t), 0 \leq s,t \leq T$. Let $\xi_g$ be a Gaussian process with the same covariance function. Let $P_\epsilon$ and $P_g$ be the measures induced by $\xi$ and $\xi_g$, respectively, in function space (namely the real-valued, Borel-measurable space over $[0,T]$), and denote by $H_{\epsilon}(\xi)$ the entropy of $P_\epsilon$ with respect to $P_g$. The following theorem extends the upper and lower bounds (Shannon, [2], [3]) on the $\epsilon$-entropy of a one-dimensional random variable to the case of a random process.

**Theorem 1:** For every $\epsilon > 0$

$$
H_{\epsilon}(\xi) - H_{\epsilon}(\xi_g) \leq H_\epsilon(\xi) \leq H_{\epsilon}(\xi_g) + o(1).
$$

If $H_{\epsilon}(\xi) < \infty$ then as $\epsilon \to 0$

$$
H_{\epsilon}(\xi) = H_{\epsilon}(\xi_g) + o(1).
$$

**Lemma 3.1:** Let $g = \{g(t), 0 \leq t \leq T\}$ be a Gaussian process. If $H_{\epsilon}(\xi) < \infty$ then as $\epsilon \to 0$

$$
H_{\epsilon}(\xi) = H_{\epsilon}(g) + o(1).
$$

Theorem 1 relates the $\epsilon$-entropy of a process $\xi$ to the $\epsilon$-entropy of $\xi_g$ and to $H_{\epsilon}(\xi)$. In the literature the absolute continuity of some non-Gaussian processes with respect to Gaussian measures are discussed. Furthermore, the $\epsilon$-entropy of these Gaussian processes is known for $\epsilon \to 0$ ([1], [6]–[10]). Therefore, in the following we try to express $H_{\epsilon}(\xi)$ in terms of $H_{\epsilon}(g)$ and $H_{\epsilon}(\xi)$, where $g$ is any given Gaussian process for which $H_{\epsilon}(\xi) < \infty$. We start with a lemma which enables us to replace $H_{\epsilon}(\xi)$ with $H_{\epsilon}(g)$ in the lower bound on $H_{\epsilon}(\xi)$.

**Lemma 3.1:** Let $g = \{g(t), 0 \leq t \leq T\}$ be a Gaussian process. If

$$
H_{\epsilon}(\xi) < \infty
$$

then $\xi_g$ and $g$ are equivalent Gaussian processes (i.e., the measures $P_\epsilon$ and $P_g$ are mutually absolutely continuous) and

$$
H_{\epsilon}(\xi) = H_{\epsilon}(g) + o(1).
$$

Note that since the entropy is always nonnegative [14], Lemma 3.1 implies

$$
H_\epsilon(\xi_g) - H_\epsilon(\xi) \leq H_\epsilon(\xi) - H_\epsilon(\xi_g) \leq H_\epsilon(\xi) - H_\epsilon(\xi_g).
$$

In the case that there exists a Gaussian process $g$ such that $H_\epsilon(\xi) < \infty$, we use Lemma 3.1 and the following known results about the behavior of the two equivalent Gaussian processes $\xi_g$ and $g$.

a) For equivalent Gaussian processes $\xi$ and $\xi_g$, it has been proved under some general conditions [8] that

$$
H_\epsilon(\xi_g) \approx H_\epsilon(g), \quad \epsilon \to 0.
$$

That is, $H_\epsilon(\xi_g)$ and $H_\epsilon(g)$ are “asymptotically equal,” in the sense that their ratio goes to one as $\epsilon$ goes to zero.

b) Suppose that $\xi_g$ and $g$ are also strongly equivalent. (Denote by $K(s,t), 0 \leq s,t \leq T$, the covariance function of a process $\xi$ and consider the integral operator $K$ defined on the space $L^2[0,T]$ by

$$
Kf(t) = \int_0^T K(s,t)f(s) \, ds, \quad f \in L^2[0,T], \quad t \in [0,T].
$$

Two Gaussian measures $P_1$ and $P_2$, with covariance functions $K_1(s,t)$ and $K_2(s,t)$, respectively, are defined to be strongly equivalent if the operator

$$
K = I - K_1 K_2^{-1/2} K_1 K_2^{-1/2}
$$

is of trace class, namely if $\sum |\lambda_i| < \infty$, where $\{\lambda_i\}$ are the eigenvalues of $K$. [18])

Then it has been shown that [10]

$$
H_\epsilon(\xi_g) = H_\epsilon(g) + o(1), \quad \epsilon \to 0
$$

which is tighter than the preceding asymptotic equality. Hence, the following theorem follows from Theorem 1, Lemma 3.1, and the asymptotic behavior of the $\epsilon$-entropies of equivalent and strongly equivalent Gaussian processes.

**Theorem 2:** If there exists a Gaussian process $g$ such that $H_\epsilon(\xi) < \infty$, then

$$
H_\epsilon(\xi_g) \approx H_\epsilon(g), \quad \epsilon \to 0.
$$

In addition, the Gaussian processes $\xi_g$ and $g$ are strongly equivalent, then

$$
H_\epsilon(\xi) = H_\epsilon(g) + o(1), \quad \epsilon \to 0.
$$

**Example 1:** Let $\xi$ be the solution of the stochastic equation

$$
\xi(t) = \xi(0) + \int_0^t \xi(s) \, ds + \beta w(t), \quad 0 \leq t \leq T
$$

or, more generally, let $\xi = \{\xi(t); 0 \leq t \leq T\}$ be a random process with $E \int_0^T \xi^2(t) \, dt < \infty$, and let $\xi$ be defined by

$$
\xi(t) = \xi(0) + \int_0^t \xi(s) \, ds + \beta w(t), \quad 0 \leq t \leq T
$$

where $w$ is a standard Brownian motion. We assume that $\xi(0) = 0$, or $\xi(0)$ is a random variable (independent of $\xi$ and
w) with density function such that $E\xi^2(0) < \infty$, and that future increments \{w(t) - w(s), 0 \leq s < t \leq T\} are independent of the $\varepsilon$-field generated by the past of $\xi$ and $w$. Then, as shown in Section IV, the asymptotic behavior of $H_s(\xi)$ is given by

$$H_s(\xi) \approx \frac{2T^2\beta^2}{\pi^2} \frac{1}{e^2}, \quad \varepsilon \to 0. \quad (22)$$

**Example 2:** Let the process $\xi_m$ be defined by

$$\xi_2(t) = \int_0^t \xi_1(s) ds,$$

$$\xi_m(t) = \int_0^t \xi_{m-1}(s) ds, \quad 0 \leq t \leq T \quad (23)$$

where $\xi_1$ stands for the process $\xi$ in Example 1. Then (Section IV) as $\varepsilon \to 0$

$$H_s(\xi_m) \approx \frac{1}{2} \beta^2 (2m-1) \left( \frac{2T}{\pi} \right)^{2m/(2m-1)} \cdot \left( 2m - 1 \right)^{-1/(2m-1)} e^{2/(2m-1)}. \quad (24)$$

The following theorem relates the rate-distortion function $R_s(D)$ of a process $\xi = \{\xi(t), 0 \leq t < \infty\}$ to the rate-distortion function of the Gaussian process $\xi_g$ and the entropy rate $h_g(\xi_g(t))$, where

$$\mathcal{H}_g(\xi) = \lim_{T \to \infty} \frac{1}{T} \mathcal{H}_0^T (\xi_0 \mathcal{T}).$$

Its proof follows directly from (17).

**Theorem 3:** Let $\xi = \{\xi(t), 0 \leq t < \infty\}$ be a random process. If there exists a Gaussian process $g = \{g(t), 0 \leq t < \infty\}$ such that $\mathcal{F}_g(\xi) < \infty$, then for every $D > 0$

$$R_s(D) - \mathcal{R}_g(\xi) \leq R_s(\xi) - \mathcal{R}_g(\xi) \leq R_s(\xi) - R_s(\xi_g). \quad (25)$$

**Example 3:** Let us assume that there exists a stationary solution $\xi$ of the stochastic equation (20). Suppose that $\xi$ is passed through a linear filter with frequency characteristic $G(j\omega)$, and the filter output is denoted by $\eta$. If $|G(j\omega)|^2 \approx c \omega^{-2r}$, $|\omega| \to \infty \quad (26)$

where $c$ and $r$ are positive constants, then (Section IV) as $D \to 0$

$$R_s(D) \approx \frac{1}{2} \left( \frac{2r + 2}{\pi} \right)^{(2r+2)/(2r+1)} (2r+1)^{1/(2r+1)} \cdot (2r+1)^{-1/(2r+1)} \left( \frac{1}{D} \right)^{1/(2r+1)}. \quad (27)$$

The $e$-Entropy of Certain Processes with $\mathcal{H}_s(\xi) = \infty$

In the following we consider mean-continuous processes $\xi$ such that $\mathcal{H}_s(\xi) = \infty$. In this case there does not exist any Gaussian process $g$ for which $\mathcal{F}_g(\xi) < \infty$, and therefore the previous lower bounds on $H_s(\xi)$ are meaningless.

Let $K_s(t,t')$, the covariance function of $\xi$ and of $\xi_g$, be continuous in $[0,T] \times [0,T]$ and denote by $\{\phi_i(\xi)\}$ and $\{\lambda_i\}$ the eigenfunctions and the eigenvalues of $K_s$. We assume that the eigenvalues are arranged in nonincreasing, monotonic order. The variables $\{\xi_1, \xi_2, \cdots\}$ and $\{\xi_{g1}, \xi_{g2}, \cdots\}$ are defined through $\{\phi_i(\xi)\}$ by

$$\xi_i = \int_0^T \xi(t) \phi_i(t) dt,$$

$$\xi_{gi} = \int_0^T \xi(t) \phi_i(t) dt, \quad i = 1, 2, \cdots \quad (28)$$

**Theorem 4:** Let $\xi$ be a mean-continuous process with $E[\xi(t)^2] < \infty$.

a) If $\mathcal{F}_g(\xi_{g1}, \cdots, \xi_{gL}) \to o(L)$ and if the eigenvalues of the process satisfy

$$\lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} i \ln \frac{\lambda_i}{\lambda_{i+1}} > 0 \quad (29)$$

then

$$H_s(\xi) \approx H_s(\xi_g), \quad \varepsilon \to 0. \quad (30)$$

b) Suppose that $\mathcal{F}_g(\xi_{g1}, \cdots, \xi_{gL}) = o(\xi_g)$ and if the eigenvalues of the process satisfy

$$\lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} i \ln \frac{\lambda_i}{\lambda_{i+1}} = \infty \quad (29)$$

then

$$H_s(\xi) \approx H_s(\xi_g), \quad \varepsilon \to 0. \quad (30)$$

**Remarks:** 1) A sufficient condition for (29) is that the rate of the decrease of the eigenvalues is greater than that of the sequence $\{1/i^a\}$, where $a > 0$. The condition I of b) is fulfilled if the rate of decrease of the eigenvalues is greater than that of $\{1/i^a\}$, for every $a > 0$. 2) Note that the set of mean-continuous processes satisfying $\mathcal{H}_g(\xi_{g1}, \cdots, \xi_{gL})(\xi_{g1}, \cdots, \xi_{gL}) = o(L)$ and $\mathcal{H}_g(\xi) = \infty$, is not empty. As an example let us define $\{\xi_1, \xi_2, \cdots\}$ as follows. The variables $\xi_i, i = 1, 2, \cdots$ are statistically independent, with $E\xi_i^2 = \lambda_i$, where $\{\lambda_i\}$ is an infinite set of summable positive numbers. Also, $\xi_i, i = 1, 10, 10^2, 10^3, \cdots$ are distributed according to

$$p(x) = \frac{1}{\sqrt{2\lambda_i}} e^{-\frac{x^2}{2\lambda_i}}$$

and all other $\xi_i$ are Gaussian. Then simple calculation yields

$$\mathcal{H}_g(\xi_{g1}, \cdots, \xi_{gL})(\xi_{g1}, \cdots, \xi_{gL}) = \sum_{i=1}^{L} \mathcal{H}_g(\xi_{gi}) = o(L) \to \infty,$$

$$L \to \infty. \quad (30)$$

3) Part II of b) is a generalization of a result of Kazi [19].

**IV. PROOF OF RESULTS**

**Proof of Theorem 1:** By the remark to Lemma 2.2 we may assume, without loss of generality, that $\xi$ is a mean-continuous process. Let $\xi_i, i = 1, 2, \cdots$ be the Karhunen-
Loève expansion coefficients of $\xi$, and let the eigenvalues $\lambda_i = E\xi_i^2$, $i = 1, 2, \cdots$ be arranged in nondecreasing, monotonic order. From Lemmas 2.1 and 2.2

$$H_2(\xi) = H_2(\xi_1, \xi_2, \cdots) = \lim_{L \to \infty} H_2(\xi_1, \cdots, \xi_L).$$

(31)

Therefore, we have to prove the lower bound in (13) only for the finite-dimensional random variable $\xi = (\xi_1, \cdots, \xi_L)$, where $E\xi_i^2 > 0$, for $i \neq j$. If $\mathcal{H}_2(\xi) = \infty$, the lower bound is trivial; therefore in the following proof of the left side of (13) we assume that

$$\mathcal{H}_2(\xi) < \infty.$$ (32)

The $\alpha$-entropy of $(\xi_1, \cdots, \xi_L)$ is defined by $I((\xi_1, \cdots, \xi_L), (\eta_1, \cdots, \eta_L))$, where the infimum is taken over all pairs $(\xi, \eta)$ with distortion vectors $(e_1, \cdots, e_L)$ such that $\sum_{i=1}^L e_i^2 \leq \epsilon^2$.

Note that we may restrict the family of distortion vectors by the additional condition $e_i^2 \leq \lambda_i$, for all $i$. Otherwise, if $e_i^2 > \lambda_i$, even for one $i$, we have [14, eq. 2.2.7]

$$I((\xi_1, \cdots, \xi_L), (\eta_1, \cdots, \eta_L)) \geq I((\xi_1, \cdots, \xi_{i-1}, \xi_{i+1}, \cdots, \xi_L), (\eta_1, \cdots, \eta_{i-1}, \eta_{i+1}, \cdots, \eta_L)).$$

and therefore by neglecting the $i$th component we can decrease the amount of information and still keep the overall distortion less than $\epsilon^2$. Therefore

$$H_2(\xi_1, \cdots, \xi_L) = \inf I((\xi_1, \cdots, \xi_L), (\eta_1, \cdots, \eta_L))$$

where the infimum is taken over all pairs $(\xi, \eta)$ with distortion vectors $(e_1^2, \cdots, e_L^2)$ satisfying

$$\sum_{i=1}^L e_i^2 \leq \epsilon^2$$

and $e_i^2 \leq \lambda_i$.

By (32) and [14] (Theorem 2.4.2), the variable $(\xi_1, \cdots, \xi_L)$ has a well-defined density function. Since for any $\eta$ the pair $(\xi, \eta)$ can be approximated, with an arbitrary degree of “accuracy,” by another pair of random variables whose distribution is absolutely continuous (see [7]), it is sufficient to restrict ourselves to those pairs which have a joint density function. Now, for a given distortion vector we use a known lower bound on the information between $\xi$ and $\eta$ [4] to obtain

$$H_2(\xi_1, \cdots, \xi_L) \geq \inf \left\{ \sum_{i=1}^L \frac{1}{2} \ln \frac{1}{2\pi e\lambda_i} + h(\xi_1, \cdots, \xi_L) \right\}$$

(34)

where the infimum is taken over all distortion vectors satisfying (33) and

$$h(\xi_1, \cdots, \xi_L) = -\int p_\xi(x_1, \cdots, x_L) \ln p_\xi(x_1, \cdots, x_L) dx_1 \cdots dx_L.$$ Then, we also have

$$H_2(\xi_1, \cdots, \xi_L) \geq \inf \left\{ \sum_{i=1}^L \frac{1}{2} \ln \frac{1}{2\pi e\lambda_i} + h(\xi_1, \cdots, \xi_L) - \frac{1}{2} \sum_{i=1}^L \ln 2\pi e\lambda_i \right\}.$$ (35)

By the convexity of $\ln x$, for all positive numbers $x_1$, $x_2$, $x_1'$, $x_2'$ and $x_1 \geq x_1' \geq x_2' \geq x_2$

$$\ln \frac{1}{x_1} + \ln \frac{1}{x_2} \geq \ln \frac{1}{x_1'} + \ln \frac{1}{x_2'}.$$ Therefore, the infimum in (35) over all vectors $(e_1^2, \cdots, e_L^2)$ satisfying (33) is attained by the following distortion vector, given by

$$e_i^2 = \begin{cases} \theta, & i = 1, \cdots, n \quad \text{and} \\ \lambda_i, & i = n + 1, \cdots, L \end{cases}$$

(36)

where $\theta$ is a positive parameter, determined from the equation

$$\epsilon^2 = \sum_{i=1}^L \min \left( \theta, \lambda_i \right)$$

(37)

and $n$ such that $\lambda_{n+1} < \theta \leq \lambda_n$.

Substituting (36) into (35) and using the known result for $H_2(\xi_{g_1}, \cdots, \xi_{g_L})$ in the Gaussian case [7], we obtain

$$H_2(\xi_1, \cdots, \xi_L) = H_2(\xi_{g_1}, \cdots, \xi_{g_L}) + h(\xi_1, \cdots, \xi_L)$$

$$- \frac{1}{2} \sum_{i=1}^L \ln 2\pi e\lambda_i.$$ (38)

Finally, if we denote by $p_\xi(x_1, \cdots, x_L)$ the density function of $(\xi_1, \cdots, \xi_L)$ and by $p_{g_\xi}(x_1, \cdots, x_L)$ the Gaussian density function of $(\xi_{g_1}, \cdots, \xi_{g_L})$, then

$$-H_2(\xi_1, \cdots, \xi_L) = -\int p_\xi(x_1, \cdots, x_L) \ln \frac{p_\xi(x_1, \cdots, x_L)}{p_{g_\xi}(x_1, \cdots, x_L)} dx_1 \cdots dx_L$$

$$= h(\xi_1, \cdots, \xi_L) - \frac{1}{2} \sum_{i=1}^L \ln 2\pi e\lambda_i,$$

which together with (38) yields the lower bound.

We turn now to the proof of (14) and the right side of (13). Let $\theta$ be the positive parameter determined by the equation

$$\epsilon^2 = \sum_{i=1}^\infty \min \left( \theta, \lambda_i \right)$$

(39)

and let $m$ be such that $\lambda_{m+1} < \theta \leq \lambda_m$. We define the variables $\eta = (\eta_1, \cdots, \eta_m)$ and $\eta_g = (\eta_{g_1}, \cdots, \eta_{g_m})$ by

$$\eta_i = a_i \xi_i + \xi_i,$$

$$\eta_{g_i} = a_i \xi_{g_i} + \xi_i, \quad i = 1, \cdots, m$$

(40)

where $a_i = 1 - \theta/\lambda_i$ and $\xi_i, i = 1, \cdots, m$ are independent Gaussian random variables with $E\xi_i^2 = \theta$ and are independent of $\xi$ and $\xi_g$.

Since by (40)

$$E \sum_{i=1}^m (\xi_i - \eta_i)^2 + E \sum_{i=m+1}^\infty \xi_i^2 = \epsilon^2$$

we have by the definition of $H_2(\xi_1, \xi_2, \cdots)$

$$H_2(\xi_1, \xi_2, \cdots) \leq I((\xi_1, \xi_2, \cdots), (\eta_1, \cdots, \eta_m, 0, \cdots))$$

$$= I((\xi_1, \xi_2, \cdots), (\eta_1, \cdots, \eta_m)).$$

(41)
Observe that by (40) the variables \((\xi_{m+1}, \cdots)\) and \((\eta_1, \cdots, \eta_m)\)\) are conditionally independent, conditioned on \((\xi_1, \cdots, \xi_m)\). Therefore, as in (11)

\[
I((\xi_1, \cdots, \xi_m), (\eta_1, \cdots, \eta_m)) = I((\xi_1, \cdots, \xi_m), (\eta_1, \cdots, \eta_m))
\]

and by (40)

\[
I((\xi_1, \cdots, \xi_m), (\eta_1, \cdots, \eta_m)) = h(\eta_1, \cdots, \eta_m) - \sum_{i=1}^{m} \frac{1}{2} \ln 2\pi e \beta_i.
\]

(43)

Since the \(\varepsilon\)-entropy of \(\xi_\varepsilon\) is given by \([7]\)

\[
H_\varepsilon(\xi_\varepsilon) = \sum_{i=1}^{\infty} \frac{1}{2} \ln \frac{1}{\varepsilon}
\]

and since by (40)

\[
-\mathcal{H}(\eta_{g_1}, \cdots, \eta_{g_m}) = h(\eta_1, \cdots, \eta_m) - \sum_{i=1}^{m} \frac{1}{2} \ln 2\pi e \beta_i (45)
\]

we obtain from (31) and (41)-(45)

\[
H_\varepsilon(\xi) \leq H_\varepsilon(\xi_\varepsilon) - \mathcal{H}(\eta_{g_1}, \cdots, \eta_{g_m}) (\eta_1, \cdots, \eta_m).
\]

(46)

The upper bound on the right side of (13) follows at once from (46) since the entropy of one measure with respect to another is always nonnegative \([14]\).

Note that both \(\mathcal{H}(\eta_{g_1}, \cdots, \eta_{g_m}) (\eta_1, \cdots, \eta_m)\) and \(m\) depend on \(\varepsilon\), and \(m \to \infty\) as \(\varepsilon \to 0\). Since

\[
\mathcal{H}_\varepsilon(\xi) = \lim_{\varepsilon \to 0} \mathcal{H}(\eta_{g_1}, \cdots, \eta_{g_m}) (\xi_1, \cdots, \xi_k)
\]

it follows that for any \(\delta > 0\) there exists an integer \(N\) such that

\[
\mathcal{H}_\varepsilon(\xi) \leq \mathcal{H}_\varepsilon(\xi_\varepsilon) - \mathcal{H}(\eta_{g_1}, \cdots, \eta_{g_m}) (\eta_1, \cdots, \eta_N) + \delta
\]

(47)

Let \(\varepsilon_0\) be sufficiently small such that \(m(\varepsilon) \geq N\), for \(0 < \varepsilon \leq \varepsilon_0\). Since

\[
\mathcal{H}(\eta_{g_1}, \cdots, \eta_{g_m}) (\eta_1, \cdots, \eta_m) \leq \mathcal{H}(\eta_{g_1}, \cdots, \eta_{g_m}) (\eta_1, \cdots, \eta_N)
\]

for \(m \geq N\) \([14, \text{ ch. 2}]\), (46) yields for \(\varepsilon \leq \varepsilon_0\)

\[
H_\varepsilon(\xi) \leq H_\varepsilon(\xi_\varepsilon) - H_\varepsilon(\eta_{g_1}, \cdots, \eta_{g_m}) (\eta_1, \cdots, \eta_N).
\]

(48)

Finally, we show that

\[
\lim_{\varepsilon \to 0} \mathcal{H}_\varepsilon(\eta_1, \cdots, \eta_N) = \mathcal{H}_\varepsilon(\eta_1, \cdots, \eta_N) = \mathcal{H}_\varepsilon(\eta_1, \cdots, \eta_N)
\]

(49)

By (40), for every \(\varepsilon \leq \varepsilon_0\) we have \([14, \text{ eq. 2.4.8}]\)

\[
\mathcal{H}(\eta_{g_1}, \cdots, \eta_{g_m}) (\eta_1, \cdots, \eta_N)
\]

\[
\leq \mathcal{H}(\xi_1, \cdots, \xi_N)\quad (\xi_1, \cdots, \xi_N)
\]

(50)

Since by (40) the random variables \((\eta_1, \cdots, \eta_N)\) and \((\eta_{g_1}, \cdots, \eta_{g_m})\) converge (as \(\varepsilon \to 0\)) in the mean-square sense to the variables \((\xi_1, \cdots, \xi_N)\) and \((\xi_{g_1}, \cdots, \xi_{g_m})\), respectively. The measures induced by \((\eta_1, \cdots, \eta_N)\) and \((\eta_{g_1}, \cdots, \eta_{g_m})\) (on the real-valued \(N\)-dimensional Borel space) converge weakly, as \(\varepsilon\) goes to zero, to the measures of \((\xi_1, \cdots, \xi_N)\) and \((\xi_{g_1}, \cdots, \xi_{g_m})\), respectively, and we have \([20]\)

\[
\lim_{\varepsilon \to 0} \mathcal{H}(\eta_{g_1}, \cdots, \eta_{g_m}) (\eta_1, \cdots, \eta_N) \geq \mathcal{H}(\xi_{g_1}, \cdots, \xi_{g_m}) (\xi_1, \cdots, \xi_N).
\]

(51)

Remark: The result (51) is derived in \([20]\) for mutual information only; however, the proof given there goes over directly to the case of entropy. From (50) and (51) we obtain (49). Therefore, from (47)-(49)

\[
H(\xi) \leq H(\xi_\varepsilon) - H(\xi_\varepsilon) + o(1), \quad \varepsilon \to 0
\]

(52)

which together with the left side of (13) yields (14).

Proof of Lemma 3.1: Let \((t_1, \cdots, t_m)\) be an arbitrary finite set in \([0,T]\). First, let us assume, for simplicity, that the density functions \(p_x(t_1, \cdots, t_m)\) of \((\xi(t_1), \cdots, \xi(t_m))\), \(p_y(t_1, \cdots, t_m)\) of \((g(t_1), \cdots, g(t_m))\) and \(p_y(t_1, \cdots, t_m)\) of \((\zeta(t_1), \cdots, \zeta(t_m))\) exist. Then

\[
\int_{R^m} p_x(x_1, \cdots, x_m) \ln \frac{p_x(x_1, \cdots, x_m)}{p_y(x_1, \cdots, x_m)} \, dx_1 \cdots dx_m
\]

(53)

Since the function in \(p_x(x_1, \cdots, x_m)/(p_y(x_1, \cdots, x_m)\) is a quadratic form of the variables \(x_1, \cdots, x_m\). Therefore, we may replace the second term on the right side of (53) with

\[
\int_{R^m} p_x(x_1, \cdots, x_m) \ln \frac{p_x(x_1, \cdots, x_m)}{p_y(x_1, \cdots, x_m)} \, dx_1 \cdots dx_m
\]

(54)

Since \(t_1, \cdots, t_m\) and \(m\) were arbitrary, (16) is proved. Also, by (15) and (16), both \(\mathcal{H}_\varepsilon(\xi)\) and \(\mathcal{H}_\varepsilon(\eta_{g_1}, \cdots, \eta_{g_m})\) are finite, and thus \(\xi\) and \(g\) are equivalent processes \([14, \text{ p. 134}]\).

Finally, in the case where the distributions \(P_{\xi(t_1), \cdots, \xi(t_m)}\) and \(P_{\eta(t_1), \cdots, \eta(t_m)}\) are not absolutely continuous, we define another set of variables as follows. Let \(\xi_i, i = 1, \cdots, m\) be independent Gaussian random variables with the same variance \(\delta^2\), which are independent of \(\xi_1, \cdots, \xi_m\) and \(g\), and let \((\xi_i(t_1), \cdots, \xi_i(t_m))\), \((\xi_{i_1}(t_1), \cdots, \xi_{i_m}(t_m))\), and \((\eta(t_1), \cdots, \eta(t_m))\) be defined by

\[
\xi_i(t_i) = \xi_i + \xi_i
\]

(55)

and \(\eta(t_i)\) are absolutely continuous, and \(P\) converges weakly to \(P\) as \(\delta\) goes to zero. The result of Lemma 3.1 follows now by the same arguments as in the proof of (49).
We turn now to the proof of Theorem 4 (the proofs of Theorems 2 and 3 follow directly from the discussion in Section III).

**Proof of Theorem 4**: a) By a property of mutual information [14, eq. 2.2.7] and the definition of $s$-entropy,

$$H_s(\xi) = H_s(\xi_1, \xi_2, \ldots, \xi_L) \geq H_s(\xi_1, \xi_2, \ldots, \xi_L)$$

for every $L$. Using Theorem 1 and our assumption on the asymptotic behavior of $H_s(\xi_1, \xi_2, \ldots, \xi_L)$, we get

$$H_s(\xi) \geq H_s(\xi_1, \xi_2, \ldots, \xi_L) - o(L).$$

For fixed $L$, the $s$-entropy in the Gaussian case is given by

$$H_s(\xi_1, \xi_2, \ldots, \xi_L) = \sum_{i=1}^{L} \frac{1}{2} \ln \frac{\lambda_i}{\theta^s}$$

where $\theta^s$ and $n^s$ satisfy

$$\theta^s = \frac{L}{\sum_{i=1}^{n^s} \frac{1}{i} \ln \frac{\lambda_i}{\theta^s}}, \quad \lambda_{n^s+1} < \theta^s \leq \lambda_n.$$ (58)

Let $k > 1$ be an integer. Then, for any $\epsilon$ satisfying $0 < \epsilon^2 < \sum_{i=1}^{k} \lambda_i$, we can always find $n^s$ and $\theta^s$ satisfying

$$\epsilon^2 = n^s \theta^s + \sum_{i=n^s+1}^{k} \lambda_i, \quad \lambda_{n^s+1} < \theta^s \leq \lambda_n.$$ (59)

Now set $L = k \cdot n^s$. By (56), (57), and (59), we have

$$H_s(\xi) \geq \sum_{i=1}^{k} \frac{1}{2} \ln \frac{\lambda_i}{\theta^s} - ko(n^s).$$ (60)

By (59)

$$\frac{\epsilon^2}{\sum_{i=1}^{k} \frac{1}{i} \ln \frac{\lambda_i}{\theta^s}} \leq \sum_{i=n^s+1}^{k} \frac{1}{2} \ln \frac{\lambda_i}{\theta^s} = \sum_{i=1}^{n^s-1} \frac{1}{2} \sum_{i=n^s+1}^{k} i \ln \frac{\lambda_i}{\lambda_{i+1}}.$$ (61)

Therefore, since $n^s$ goes to infinity as $\epsilon$ goes to zero, (60), (61), and (29) yield

$$H_s(\xi) \geq \sum_{i=1}^{k} \frac{1}{2} \ln \frac{\lambda_i}{\theta^s}, \quad \epsilon \rightarrow 0.$$ (62)

Finally, since $k$ was arbitrary, we take the limit in (62) for $k \rightarrow \infty$ to obtain

$$H_s(\xi) \geq \sum_{i=1}^{\infty} \frac{1}{2} \ln \frac{\lambda_i}{\theta^s}, \quad \epsilon \rightarrow 0.$$ (63)

where $n$ and $\theta$ satisfy

$$\epsilon^2 = \sum_{i=1}^{n} \min (\theta^s, \lambda_i), \quad \lambda_{n+1} < \theta^s \leq \lambda_n.$$ (64)

Since $H_s(\xi_{i}) = \sum_{i=1}^{L} \frac{1}{2} \ln (\lambda_i/\theta^s)$ (and since by Theorem 1 $H_s(\xi) \leq H_s(\xi_{i})$), the proof of part a) is completed.

b) The proof of part I of b) is similar to that of part a).

II. If $H(\xi_{i}) = o(L)$, then there exists a constant $M > 1$ such that for every $L$

$$H(\xi_{i}) \leq \frac{L}{2} \ln M^2.$$ (65)

Therefore, by (55) and Theorem 1

$$H_s(\xi_{i}) \geq H(\xi_{i_1}, \xi_{i_2}, \ldots, \xi_{i_L}) - \frac{L}{2} \ln M^2$$

for every $L$. Let $\epsilon$ be small enough that $0 < \epsilon^2 < 1/M^2$

$$\sum_{i=1}^{\infty} \lambda_i.$$ We choose $L$ and $\theta$ satisfying

$$\ln \epsilon^2 / L \leq \ln \lambda_n, \quad \epsilon \rightarrow 0.$$ (66)

Since $\epsilon^2 / L \leq \ln \lambda_n$, the $s$-entropy of $(\xi_{i_1}, \ldots, \xi_{i_L})$ is given by

$$H_s(\xi_{i_1}, \ldots, \xi_{i_L}) = \frac{L}{2} \sum_{i=1}^{L} \ln \frac{\lambda_i}{\epsilon^2 / L}.$$ (67)

From (66) and (67) we obtain

$$H_s(\xi) \leq \frac{L}{2} \ln \frac{\lambda_L}{\epsilon^2 / L} = \frac{1}{2} \ln \frac{\lambda_L}{\epsilon^2 / L}.$$ (68)

Next, define $\epsilon_i^2$ by

$$\epsilon_i^2 = (\Delta \theta)^2 + \sum_{i=L+1}^{\infty} \lambda_i.$$ (69)

Since by (66) and (69)

$$\epsilon_i^2 = \Delta \theta + \sum_{i=L+1}^{\infty} \lambda_i,$$

it follows that

$$\sum_{i=L+1}^{\infty} \ln \frac{\lambda_i}{\Delta \theta} = H_s(\xi_{i}).$$ (70)

Observe that for $\{\lambda_i\}$ as in b), II we have

$$\sum_{i=1}^{\infty} \lambda_i = o(L \lambda_L), \quad L \rightarrow \infty.$$ (72)

and therefore by (69) and (70) $\epsilon_i^2 \approx (B \lambda \epsilon)^2$ for $\epsilon \rightarrow 0$, where $B$ is a constant (e.g., for $\alpha = 2$, which is the case of the diffusion process discussed by Kazi [19], we have $B \lambda = 2$). Substituting the relation between $\epsilon_i^2$ and $\epsilon^2$ into (71) and using (68), we obtain

$$H_s(\xi) \geq H_s(\xi_{i}), \quad \epsilon \rightarrow 0.$$ (73)

Finally, it is easy to show that the $s$-entropy of a Gaussian process $\xi_{i}$ with eigenvalues that are going to zero as an inverse power of $i$ satisfies

$$H_s(\xi_{i}) \approx (\text{constant}) \epsilon^{-2/(2\alpha - 1)}, \quad \epsilon \rightarrow 0$$

(compare also (24)). Therefore we obtain the following desired result:

$$H_s(\xi_{i}) \approx kH_s(\xi_{i}), \quad \epsilon \rightarrow 0$$

where $k$ is a constant.

**Proof of Examples**: We start with the proof of Example 1. Let the Gaussian process $g$ be defined as

$$g(t) = g(0) + \beta \xi(t), \quad 0 \leq t \leq T$$

where $g(0) = \xi(0)$, and let $P$ and $P_{\beta}$ be the measures (in function space) induced by $\xi$ and $g$, respectively. Since the asymptotic behavior of $H(\xi)$ is given by (22) [10], we have only to show that $H(\xi)$ is finite. In this case the Radon-Nikodym derivative $dP_{\beta}/dP$ is given by [12], [13]

$$\ln \frac{dP_{\beta}}{dP} = \ln \frac{dP_{\beta}(0)}{dP(0)} + \frac{1}{\beta^2} \int_{0}^{T} \xi(t) \xi(t) \, dt$$

$$+ \frac{1}{\beta} \int_{0}^{T} \xi(t) \xi(t) \, dt - \frac{1}{2\beta^2} \int_{0}^{T} \xi(t) \xi(t) \, dt.$$ (74)
(almost surely $P_{\xi}$), where $\xi(t)$ is defined by

$$\xi(t) = E_{\eta}[\xi(t) \mid \xi(t)]. \quad \tau < t.$$ 

Taking the expectation of both sides of (74) with respect to $P_{\xi}$ yields [14, theorem 2.4.2]

$$H_{\eta}(\xi) = H_{\eta}(\xi(0)) + \frac{1}{2\beta^2} E \left[ \int_0^T \xi(t)^2 dt \right].$$

Since $E[\xi(t) - \xi(t)]^2 \leq 0$, (75) yields

$$H_{\eta}(\xi) \leq H_{\eta}(\xi(0)) + \frac{1}{2\beta^2} E \left[ \int_0^T \xi^2(t) dt \right] < \infty$$

which completes the proof of Example 1.

Example 2 is a particular case of the following generalization: Let $\xi$ be a random process and let $g$ be a Gaussian process such that $H_{\eta}(\xi) < \infty$. Let $\xi$ and $g$ be defined by $\xi = A\xi$, $g = Ag$, where $A$ is a measurable transformation such that $g$ is a Gaussian process with $E \int_0^T g^2(t) dt < \infty$. Then

$$H_{\eta}(H) \approx H_{\eta}(H), \quad \varepsilon \to 0.$$ (77)

Since we have $H_{\eta}(\xi) \geq H_{\eta}(\xi(0))$ [14, ch. 2], (77) follows directly from Theorem 1.

Now, the proof of Example 2 follows from (77), (76), and the result (24) for the Gaussian process defined by $m - 1$ integrations of the Wiener process [10].

Finally, we turn to the proof of Example 3. Let the Gaussian process $g$ be as in the proof of Example 1. Then $H_{\eta}(\xi) < \infty$, and therefore $H_{\eta}(\xi(0)) < \infty$, where $g$ is the output of the same linear filter $G(j\omega)$ operating on the process $g$. Therefore, by Theorem 3,

$$R_{\eta}(D) \approx R_{\eta}(D), \quad D \to 0.$$ (78)

It has been shown in [21], under some general conditions, that the asymptotic behavior of the power spectral density of $\xi$ is given by

$$S_{\xi}(\omega) \approx \beta^2 \omega^{-2}, \quad |\omega| \to \infty.$$ (79)

Therefore, the power spectral density of $\eta$ satisfies

$$S_{\eta}(\omega) = c\beta^2 \omega^{-2} + o(\omega^{-2}).$$ (80)

Using the known result (e.g., [16, ch. 9]) on the rate-distortion function of a Gaussian process with spectral density $S(f)$

$$\left\{ \begin{array}{l}
R(D) = \frac{1}{2} \int \max \left( \frac{S(f)}{\theta}, 1 \right) df \\
D = \int \min \left[ S(f), \theta \right] df.
\end{array} \right.$$ (81)

It is easy to show that the asymptotic behavior of the rate-distortion function of a Gaussian process with spectral density satisfying (78) is given by (27). Thus the proof is completed.

REFERENCES