The Interpretation of Spectral Entropy Based Upon Rate Distortion Functions

Jaewoo Jung
Department of Electrical and Computer Eng.
University of California, Santa Barbara
Santa Barbara, CA 93106-9560, USA
Email: jwjung@ece.ucsb.edu

Jerry D. Gibson
Department of Electrical and Computer Eng.
University of California, Santa Barbara
Santa Barbara, CA 93106-9560, USA
Email: gibson@ece.ucsb.edu

***Abstract*— In 1960 Campbell derived a quantity that he called coefficient rate which is expressible in terms of the entropy of the process power spectral density. Later, Yang, et al showed that the spectral entropy is proportional to the logarithm of the equivalent bandwidth of the smallest frequency band containing most of the energy. Gibson et al also showed that for discrete time AR(1) sequences, Campbell's coefficient rate and Shannon's entropy rate power are equal but that the equality does not hold for higher order AR processes. In this paper, we derive a new expression for Campbell's coefficient rate in terms of the parametrized version of the rate distortion function of a Gaussian random process with a given power spectral density subject to the MSE fidelity criterion. We also derive expressions for the entropy rate power and coefficient rate in terms of the slope of the rate distortion function for the given source and for a source with flat power spectral density.**

I. Introduction

The coefficient rate of a random process was first derived and defined by Campbell in 1960[1]. Campbell considered the product of N sample functions of a random process, and showed, using an AEP(Asymptotic Equipartition Property)-like argument, that a Karhunen-Loève expansion of this product could be separated into two sets - one set with average power very close to that of the product and the other set having very low average power. Asymptotically in the number of sample functions forming the product and in the support interval of the process, he showed that the average number of terms in the high power set approached a quantity that he interpreted as a coefficient rate given by

$$Q_2 = \exp\Big[-\int_{-\infty}^{\infty} S(f)\log S(f)df\Big], \tag{1}$$

where we denote the quantity in the exponent as the spectral entropy (we use Q_2 for the coefficient rate and reserve Q_1 for Shannon's entropy rate power[2],[3],[4]).

Campbell did not investigate further the implications of coefficient rate for signal processing or for source compression. Abramson also considered Campbell's coefficient rate and noted that an approach to source compression using coefficient rate did not exist and was not immediately obvious[5].

Gibson, et al studied the relationship of coefficient rate and Shannon's entropy rate power and proved that for first order autoregressive processes (AR(1)), these quantities are equal, but that for higher order AR processes, equality does not hold[6]. Using a different development based upon the AEP for continuous random variables, Yang, et al showed that the equivalent rate per dimension of a random process falls within a small region about the coefficient rate defined by Campbell. Additionally, they showed that the number of times that a particular coefficient in the Karhunen-Loève expansion appears in the high energy set is proportional to the variance of that coefficient[7].

In this paper, we investigate Campbell's coefficient rate within the context of rate distortion theory and develop two significant results. First, based upon the classical expression for the rate distortion function in terms of the parameter θ, we derive a new expression for coefficient rate in terms of the bandwidth occupied by the process as a function of θ. Second, we derive an expression for the spectral entropy in terms of the ratio of the values of θ of a uniform PSD to an AR PSD for a given value of average distortion D, where this ratio is averaged over all values of distortion greater than the distortion associated with the minimum of the PSD. This result is contrasted with a similar result for the entropy rate power.

The paper is organized as follows. Section II contains rate distortion theory basics used in the remainder of the paper. Section III considers the averaging of the parametrized rate distortion function for a Gaussian source with respect to the MSE fidelity criterion over all values of the parameter θ and presents the resulting new expression for coefficient rate in terms of the bandwidth as a function of θ. Section IV contains the derivation of new expressions for the spectral entropy in terms of the ratio of the values of θ for a uniform PSD and an AR PSD at a given average distortion and in terms of the ratio of the slopes of the rate distortion functions corresponding to these two PSDs. A related result for the entropy rate power is also derived. Section V contains conclusions and future research directions.

II. Rate-Distortion Curve

For use in subsequent sections, we note the following theorem.

Theorem (Berger [2]) *Let* $\{X_t, t = 0, \pm 1, \cdots,\}$ *be a time-discrete stationary Gaussian source with spectral density* *function*

$$S(f)=\sum_{k=-\infty}^{\infty}\phi_k e^{-jk2\pi f}. \tag{2}$$

Then the MSE rate distortion function of $\{X_t\}$ *has the parametric representation*

$$D_\theta=\int_{-0.5}^{0.5}\min[\theta,S(f)]df, \tag{3}$$

and

$$R(D_\theta)=\frac{1}{2}\int_{-0.5}^{0.5}\max[0,\log\frac{S(f)}{\theta}]df. \tag{4}$$

The nonzero portion of the $R(D)$ *curve is generated as the parameter* θ *transverses the interval* $0\le\theta\le\Delta=$ $ess\ sup\ S(f)$.

The entropy power, or entropy rate power, of a stationary Gaussian sequence $\{X_t\}$ is

$$Q_1=\exp\Big[\int_{-0.5}^{0.5}\log S(f)df\Big]. \tag{5}$$

The right side of Eq. (5) is the formula for the one-step ahead prediction error of a stationary process with power spectral density (PSD) $S(f)$[2].

III. Expected Value of Rate

In transform coding, the actual bandwidth of a coded sequence, i.e., the number of coefficients that need to be coded, depends on the distortion level θ. We assume that the distribution of θ is uniform over $[0,a]$ where a is the maximum value of the PSD, and calculate the expected value of $R(D_\theta)$ in the above Theorem as

$$\begin{aligned}E[R(D_\theta)] &= \int_0^a R(D_\theta)d\theta \qquad (6)\\ &= \int_0^a\Big\{\int_{-0.5}^{0.5}\max[0,\frac{1}{2}\log\frac{S(f)}{\theta}]df\Big\}d\theta\\ &= \int_{-0.5}^{0.5}\Big\{\int_0^a\max[0,\frac{1}{2}\log\frac{S(f)}{\theta}]d\theta\Big\}df,\end{aligned}$$

by changing the order of integration.
For a fixed f, the max function of the integrand is

$$\max[0,\frac{1}{2}\log\frac{S(f)}{\theta}]=\begin{cases}\frac{1}{2}\log\frac{S(f)}{\theta} & 0\le\theta\le S(f)\\ 0 & S(f)\le\theta\le a.\end{cases}$$

The range of integration can be divided into two parts,

$$E[R(D_\theta)]=\int_{-0.5}^{0.5}\Big\{\int_0^{S(f)}\frac{1}{2}\log\frac{S(f)}{\theta}d\theta+\int_{S(f)}^{a}0d\theta\Big\}df$$

$$=\int_{-0.5}^{0.5}\Big\{\int_0^{S(f)}\frac{1}{2}[\log S(f)-\log\theta]d\theta\Big\}df \tag{7}$$

$$=\frac{1}{2}\int_{-0.5}^{0.5}S(f)\log S(f)df+\frac{1}{2}\int_0^a 2W(\theta)(\log\frac{1}{\theta})d\theta, \tag{8}$$

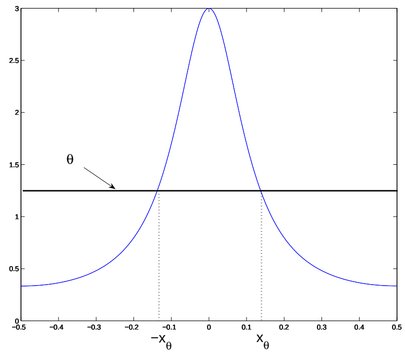

Fig. 1. Threshold θ and PSD of an AR(1) process (r=0.5)

where $2W(\theta)$ is the bandwidth where the magnitude of the PSD is greater than the threshold θ. In the case of the AR(1) process in Fig. 1, $2W(\theta)=2x_\theta$.
The first term of the right-hand side of Eq. (8) has the form of the spectral entropy. If we multiply Eq. (8) by 2,

$$-\int_{-0.5}^{0.5}S(f)\log S(f)df=-2E[R(D_\theta)]+\int_0^a 2W(\theta)(\log\frac{1}{\theta})d\theta.$$

The expected value of rate in Eq. (7) is

$$\begin{aligned}E[R(D_\theta)] &= \int_{-0.5}^{0.5}\Big\{\frac{1}{2}[\log S(f)\theta-(\theta\log\theta-\theta)]\Big|_0^{S(f)}\Big\}df\\ &= \int_{-0.5}^{0.5}\Big\{\frac{1}{2}S(f)\Big\}df.\end{aligned}$$

Assuming the source is unit-variance Gaussian and $\int_{-0.5}^{0.5}S(f)df=1$,

$$E[R(D_\theta)]=\frac{1}{2}. \tag{9}$$

The expected value of the rate is always 1/2, regardless of the shape of the PSD.
From the expected value of rate, we can derive the spectral entropy in another form as

$$-\int_{-0.5}^{0.5}S(f)\log S(f)df=-1+\int_0^a 2W(\theta)(\log\frac{1}{\theta})d\theta. \tag{10}$$

The coefficient rate Q_2 is

$$Q_2=\exp\{-1+\int_0^a 2W(\theta)(\log\frac{1}{\theta})d\theta\}. \tag{11}$$

The bandwidth $2W(\theta)$ affects Q_2. We can rewrite Q_2 as

$$Q_2=\frac{\exp\{\int_0^a 2W(\theta)(\log\frac{1}{\theta})d\theta\}}{\exp\{\int_0^1(\log\frac{1}{\theta})d\theta\}}. \tag{12}$$

The limits of the integral in the denominator are the same as those in the numerator in the case of uniform power spectral density, i.e.,$a=1$ and $2W(\theta)=1$. We infer from Eq. (12) that the uniform power spectral density is a reference power spectral density for Q_2.

IV. Coefficient Rate and Entropy Power from Rate-Distortion Curves

To investigate the physical meaning of the spectral entropy, the coefficient rate and entropy power are derived from the rate-distortion curves.

A. Rate, Distortion, and slope in term of threshold θ

The distortion of an AR(1) process in Fig. 1 is

$$\begin{aligned} D_\theta &= \int_{-0.5}^{0.5} \min[\theta, S(f)]df \\ &= \theta 2x_\theta + 2\int_{x_\theta}^{0.5} S(f)df. \end{aligned}$$

The derivative of D_θ with respect to θ is

$$\frac{\partial D_\theta}{\partial \theta} = 2x_\theta + 2\theta\frac{\partial x_\theta}{\partial \theta} - 2S(x_\theta)\frac{\partial x_\theta}{\partial \theta}.$$

Since $\theta = S(x_\theta)$,

$$\frac{\partial D_\theta}{\partial \theta} = 2x_\theta. \tag{13}$$

The rate of the AR(1) process in Fig. 1 is

$$\begin{aligned} R(D_\theta) &= \frac{1}{2}\int_{-0.5}^{0.5} \max[0, \log\frac{S(f)}{\theta}]df \\ &= \frac{1}{2}\int_{-x_\theta}^{x_\theta} \log\frac{S(f)}{\theta}df \\ &= \int_{0}^{x_\theta} \log S(f)df - x_\theta \log\theta. \end{aligned}$$

The derivative of $R(D_\theta)$ with respect to θ is

$$\frac{\partial R(D_\theta)}{\partial \theta} = \log S(x_\theta)\frac{\partial x_\theta}{\partial \theta} - (\frac{\partial x_\theta}{\partial \theta}\log\theta + x_\theta\frac{1}{\theta}).$$

Since $\theta = S(x_\theta)$,

$$\frac{\partial R(D_\theta)}{\partial \theta} = -x_\theta\frac{1}{\theta}. \tag{14}$$

From Eqs. (13) and (14),

$$\frac{\partial R(D_\theta)}{\partial D_\theta} = \frac{\partial R(D_\theta)}{\partial \theta}\frac{\partial \theta}{\partial D_\theta} = -x_\theta\frac{1}{\theta}\times\frac{1}{2x_\theta} = -\frac{1}{2\theta}, \tag{15}$$

where $-\frac{1}{2\theta}$ is the slope of the rate-distortion curve at the given threshold θ.

In the case of a memoryless source

$$\frac{\partial R(D_\theta)}{\partial D_\theta} = -\frac{1}{2\theta} = -\frac{1}{2D_\theta}. \tag{16}$$

For a fixed distortion D, the absolute value of the slope $|-\frac{1}{2\theta}|$ decreases as the coefficient r of the AR(1) process increases.

We now consider an AR(2) process. If the PSD of an AR(2) process is bimodal, there are two frequencies, $x_{1\theta}$ and $x_{2\theta}$, where the PSD meets the threshold θ. The distortion of the AR(2) process is

$$\begin{aligned} D_\theta &= \int_{-0.5}^{0.5} \min[\theta, S(f)]df \\ &= \theta 2[x_{2\theta} - x_{1\theta}] + 2\int_{0}^{x_{1\theta}} S(f)df + 2\int_{x_{2\theta}}^{0.5} S(f)df. \end{aligned}$$

The derivative of D_θ with respect to θ is

$$\begin{aligned} \frac{\partial D_\theta}{\partial \theta} &= 2[x_{2\theta} - x_{1\theta}] + 2\theta[\frac{\partial x_{2\theta}}{\partial \theta} - \frac{\partial x_{1\theta}}{\partial \theta}] \\ &\quad + 2S(x_{1\theta})\frac{\partial x_{1\theta}}{\partial \theta} - 2S(x_{2\theta})\frac{\partial x_{2\theta}}{\partial \theta}. \end{aligned}$$

Since $\theta = S(x_{1\theta}) = S(x_{2\theta})$,

$$\frac{\partial D_\theta}{\partial \theta} = 2[x_{2\theta} - x_{1\theta}]. \tag{17}$$

The rate of the AR(2) process is

$$\begin{aligned} R(D_\theta) &= \frac{1}{2}\int_{-0.5}^{0.5} \max[0, \log\frac{S(f)}{\theta}]df \\ &= \int_{x_{1\theta}}^{x_{2\theta}} \log\frac{S(f)}{\theta}df \\ &= \int_{x_{1\theta}}^{x_{2\theta}} \log S(f)df - [x_{2\theta} - x_{1\theta}]\log\theta. \end{aligned}$$

The derivative of $R(D_\theta)$ with respect to θ is

$$\begin{aligned} \frac{\partial R(D_\theta)}{\partial \theta} &= \log S(x_{2\theta})\frac{\partial x_{2\theta}}{\partial \theta} - \log S(x_{1\theta})\frac{\partial x_{1\theta}}{\partial \theta} \\ &\quad -([\frac{\partial x_{2\theta}}{\partial \theta} - \frac{\partial x_{1\theta}}{\partial \theta}]\log\theta + [x_{2\theta} - x_{1\theta}]\frac{1}{\theta}). \end{aligned}$$

Since $\theta = S(x_{1\theta}) = S(x_{2\theta})$,

$$\frac{\partial R(D_\theta)}{\partial \theta} = -[x_{2\theta} - x_{1\theta}]\frac{1}{\theta}. \tag{18}$$

From Eqs. (17) and (18),

$$\begin{aligned} \frac{\partial R(D_\theta)}{\partial D_\theta} &= \frac{\partial R(D_\theta)}{\partial \theta}\frac{\partial \theta}{\partial D_\theta} \\ &= -[x_{2\theta} - x_{1\theta}]\frac{1}{\theta}\cdot\frac{1}{2[x_{2\theta} - x_{1\theta}]} = -\frac{1}{2\theta} \end{aligned} \tag{19}$$

We can consider $2[x_{2\theta} - x_{1\theta}]$ for an AR(2) process as the bandwidth where the magnitude of the PSD is greater than the threshold θ. In other words, $2W(\theta)$ in Eq. (10) is equal to $2[x_{2\theta} - x_{1\theta}]$ in an AR(2) process.

Similarly, $2W(\theta)$ in Eq. (10) can be interpreted as a bandwidth for higher order AR processes and $-\frac{1}{2\theta}$ is the slope of the rate-distortion curve at the given threshold θ.

B. Spectral Entropy from Rate-Distortion Curves

From Eq. (13),

$$2x_\theta d\theta = dD. \tag{20}$$

The second term of the right-hand side of the spectral entropy in Eq. (10) can be written using Eq. (20) as

$$\int_0^a 2x_\theta(\log\frac{1}{\theta})d\theta = \int_0^1 \log\frac{1}{\theta(D)}dD, \tag{21}$$

where $\theta(D)$ is a function of D.

The first term of the right-hand side of the spectral entropy in Eq. (10) can be expressed in terms of D

$$-1 = -\int_0^1 \log\frac{1}{D}dD. \tag{22}$$

Therefore we can rewrite the spectral entropy of the AR(1) process in Eq. (10) as

$$-\int_{-0.5}^{0.5} S(f)\log S(f)df = \int_0^1 \log\frac{D}{\theta(D)}dD, \tag{23}$$

where D is equal to θ when the power spectral density is uniform. We can interpret the ratio $\frac{D}{\theta(D)}$ as the ratio of thresholds of a uniform PSD and an AR PSD for a given distortion D,

$$\frac{D}{\theta(D)} = \frac{\text{threshold of uniform PSD given } D}{\text{threshold of PSD given } D}. \tag{24}$$

This ratio of two thresholds gives us partial information regarding the compactness of the PSD for a given distortion D. We can interpret the compactness as a bandwidth within which most of the energy of the PSD is contained.

Since $2x_\theta$ in Eq. (20) is the bandwidth where the magnitude of the PSD is greater than the threshold θ, the derived spectral entropy in Eq. (23) can be applied to higher order AR processes.

We can rewrite the spectral entropy in Eq. (23) in terms of a slope s of the MSE rate-distortion curve,

$$\int_0^1 \log\frac{D}{\theta(D)}dD = \int_0^1 \Big[\log\frac{1}{2\theta(D)} - \log\frac{1}{2D}\Big]dD.$$

From Eqs. (15) and (16), $\frac{1}{2\theta(D)}$ is the absolute value of the slope of the rate-distortion curve for the PSD of interest, and $\frac{1}{2D}$ is the slope of the rate-distortion curve for the uniform PSD. The spectral entropy is represented by the integral of the log of the ratio of the slopes,

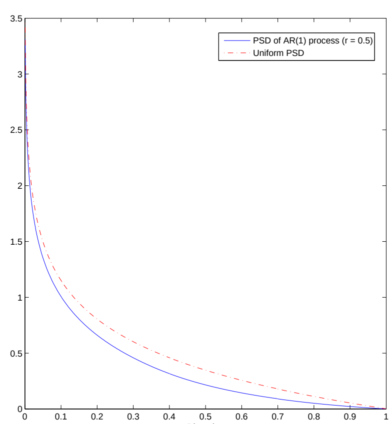

Fig. 2. rate-distortion curve for the PSD of AR(1) process and the uniform PSD

$$\int_0^1 \log\frac{D}{\theta(D)}dD = \int_0^1 \Big[\log|s_{S(f)}| - \log|s_{U(f)}|\Big]dD,$$

$$\int_0^1 \log\frac{D}{\theta(D)}dD = \int_0^1 \log\Big|\frac{s_{S(f)}}{s_{U(f)}}\Big|dD. \tag{25}$$

This ratio of the slopes contains information about the shape of the power spectral density for a given amount of distortion, and the integral over the entire range of possible distortion gives us information about the shape of the source power spectral density.

Let D^* denote the distortion where the threshold θ is the same as the minimum value of the PSD. The slope $s_{S(f)}$ and the slope $s_{U(f)}$ are the same when the distortion is in the range $[0, D^*]$. Actually, the bandwidth of a uniform PSD is constant for all possible levels of distortion, but, the bandwidth of the AR process starts to decrease as the distortion level increases beyond the point D^*. Therefore, when $0 \le D \le D^*$, the two PSDs have the same bandwidth and the ratio of two slopes $\frac{s_{S(f)}}{s_{U(f)}}$, as well as the ratio of thresholds $\frac{D}{\theta(D)}$ in Eq. (25) is unity. Therefore the distortion range $[0, D^*]$ does not affect the spectral entropy. The spectral entropy is given by

$$\begin{aligned} -\int_{-0.5}^{0.5} S(f)\log S(f)df &= \int_0^1 \log\frac{D}{\theta(D)}dD \\ &= \int_0^1 \log\Big|\frac{s_{S(f)}}{s_{U(f)}}\Big|dD \\ &= \int_{D^*}^1 \log\frac{D}{\theta(D)}dD \\ &= \int_{D^*}^1 \log\Big|\frac{s_{S(f)}}{s_{U(f)}}\Big|dD. \end{aligned}$$

The slopes in the range $[D^*, 1]$ determine the spectral entropy. The shape of the PSD determines the slope for each value of distortion, as well as a corresponding bandwidth.
Therefore,

$$Q_2 = \exp\Big\{\int_{D^*}^1 \log\frac{D}{\theta(D)}dD\Big\} = \exp\Big\{\int_{D^*}^1 \log\Big|\frac{s_{S(f)}}{s_{U(f)}}\Big|dD\Big\}.$$

Since the coefficient rate Q_2 is the exponential of the spectral entropy, the coefficient rate is a geometric mean of the ratio $\frac{s_{S(f)}}{s_{U(f)}}$ over all possible distortions. This ratio of two slopes gives us information of the bandwidth for a given distortion D, and the coefficient rate is interpreted as an indicator of a bandwidth in which most of the energy is contained. This agrees with the equivalent bandwidth derived by Yang, Gibson and He[7].

The spectral entropy is also expressed as a divergence of the AR PSD and the uniform PSD,

$$\begin{aligned} -\int_{-0.5}^{0.5} S(f)\log S(f)df &= -\int_{-0.5}^{0.5} S(f)\log\frac{S(f)}{U(f)}df \\ &= -D(S(f)||U(f)). \end{aligned}$$

Likewise, the divergence can be interpreted as the energy compactness of the AR PSD.

C. Entropy Power from Rate-Distortion Curves

Let D^* denote the distortion where the threshold θ is the same as the minimum value of the PSD of AR processes.

We know that

$$s_{S(f)} = s_{U(f)} \quad \text{if } D \le D^*. \tag{26}$$

The entropy power is expressed as

$$Q_1 = \exp\{2[R_{S(f)}(D) - R_{U(f)}(D)]\} \quad \text{if} \quad D \leq D^*. \quad (27)$$

(Proof)
The rates of $S(f)$ and $U(f)$ at $D = D^*$ are

$$R_{S(f)}(D^*) = \frac{1}{2}\int_{-0.5}^{0.5} \log \frac{S(f)}{D^*} df,$$

and

$$R_{U(f)}(D^*) = \frac{1}{2} \log \frac{1}{D^*}.$$

The difference between the two rates at D^* is

$$R_{S(f)}(D^*) - R_{U(f)}(D^*) = \frac{1}{2}\int_{-0.5}^{0.5} \log S(f) df.$$

Since from Eq. (26), the slopes are averaged for $D \leq D^*$,

$$R_{S(f)}(D) - R_{U(f)}(D) = \frac{1}{2}\int_{-0.5}^{0.5} \log S(f) df \quad \text{if} \quad D \leq D^*.$$

Therefore,

$$Q_1 = \exp\{2[R_{S(f)}(D) - R_{U(f)}(D)]\} \quad \text{if} \quad D \leq D^* \qquad \blacksquare$$

We conclude that the entropy power of any order AR process is expressed in terms of the difference between $R_{S(f)}(D)$ and $R_{U(f)}(D)$ under the condition $D \leq D^*$.
As an example, if we substitute the general distortion value D^* with a value $\frac{1-r}{1+r}$, which is the minimum value of the PSD of an AR(1) process, the entropy power of an AR(1) process in Equation (27) can be expressed as

$$Q_1 = \exp\{2[R_{S(f)}(D) - R_{U(f)}(D)]\} \quad \text{if} \quad D \leq \frac{1-r}{1+r}. \quad (28)$$

In general, the entropy power Q_1 has information about the PSD because

$$\begin{aligned} \log Q_1 &= 2[R_{S(f)}(D^*) - R_{U(f)}(D^*)] \\ &= \int_{D^*}^{1} \left|2\ s_{S(f)}\right| dD - \int_{D^*}^{1} \left|2\ s_{U(f)}\right| dD \\ &= \int_{D^*}^{1} \frac{1}{\theta(D)} dD - \int_{D^*}^{1} \frac{1}{D} dD. \end{aligned}$$

$\theta(D)$ contains the information about the PSD. However $\theta(D)$ is an integrand, and the integral value determines the entropy power. Each $\theta(D)$ is not meaningful, but the integral value is meaningful in the entropy power.

The coefficient rate Q_2 is a geometric mean of the ratio $\frac{D}{\theta(D)}$ or the ratio $\frac{s_{S(f)}}{s_{U(f)}}$. Each value contributes to the coefficient rate. Each value also has information about the shape of the PSD. Therefore, the coefficient rate has more information about the shape of the PSD than the entropy power.

V. Conclusions

The coefficient rate derived by Campbell is a representational result and not a compression result. However, in at least one special case, coefficient rate can be related to entropy power, which plays an important role in rate distortion theory. In this paper, we have derived a new expression for coefficient rate by averaging the parameterized rate distortion function over the parameter θ. Furthermore, we have obtained an expression for coefficient rate in terms of the log of the ratio of the slopes of the rate distortion function of the given source to that of a uniform source, where the logarithm of this ratio is averaged over large distortions.

Acknowledgment

This research was supported by NSF Grants CCR-0243332, CCR-0087568, and CCF-0429884.

References

[1] L. Campbell, "Minimum coefficient rate for stationary random processes," *Information and Control*, Vol. 3, pp. 360-371, Dec. 1960.
[2] T. Berger, *Rate Distortion Theory,* Prentice-Hall, Englewood Cliffs, NJ. 1968.
[3] C. E. Shannon, "A Mathematical Theory of Communication," *Bell System Technical Journal*, Vol. 27, pp. 379-423 Jul. 1948.
[4] C. E. Shannon, "A Mathematical Theory of Communication," *Bell System Technical Journal*, Vol. 27, pp. 623-656 Oct. 1948.
[5] N. Abramson, "Information theory and information storage," in *Proc. of Symp. on System Theory*, New York, NY, April 20-22, 1965, vol. XV, pp. 207-213, Polytechnie Institute of Brooklyn, Brooklyn, NY.
[6] J. D. Gibson S.P. Stanners, and S.A. McClellan, "Spectral entropy and coefficient rate for speech coding," in *Proc. of Asilomar Conference on Signals, Systems and Computers*, Nov. 1993.
[7] W. Yang J. D. Gibson and T. He, "Coefficient rate and lossy source coding," *IEEE Trans. on Information Theory*, Vol. 51, No. 1, pp. 381-386, Jan. 2005.