TR-85-32

PERFORMANCE OF ENTROPY-CONSTRAINED BLOCK TRANSFORM QUANTIZERS

by

N. FArvardin and F.Y. Lin

Performance of Entropy-Constrained Block Transform Quantizers‡

by

N. Farvardin and F.Y. Lin

Electrical Engineering Department
and
Systems Research Center
University of Maryland
College Park, Maryland 20742

Abstract

An analysis of the rate-distortion performance of an optimum entropy-constrained block transform quantization scheme operating on discrete-time stationary autoregressive processes is presented. Uniform-threshold quantization is employed to quantize the transform coefficients. An algorithm for optimum stepsize (or, equivalently, entropy) assignment among the quantizers is developed. A simple asymptotic formula indicating the high rate performance of the block transform quantization scheme is presented. Finally, specific results determining the rate-distortion performance of the entropy-constrained block transform quantization scheme operating upon first-order Gauss-Markov and Laplace-Markov sources are presented and appropriate comparisons with the Huang and Schultheiss block transform quantization, vector quantization and predictive encoding are rendered.

‡ This work was supported in part by a grant from Martin Marietta Laboratories, in part by a grant from the Minta Martin Fund for Aeronautical Research, and in part by the National Science Foundation under grant No. OIR-85-00108.

I. Introduction:

There are numerous techniques available for analog-to-digital conversion or data compression of discrete time sources. They range from simple scalar (zero-memory) quantization or PCM, to more sophisticated schemes such as predictive encoding, tree encoding and multi-dimensional quantization. It is well-known that the potential gains of data compression are most noticeable for highly correlated sources.

One of the most popular data compression schemes for encoding of correlated sources is block transform quantization. The main advantages associated with the block transform quantization scheme are: (i) good performance results, and (ii) ease of implementation.

In a typical block transform quantization scheme blocks of length L of the source output are operated upon by a transformation. Then, the resulting transform coefficients are quantized and encoded for transmission over a noiseless channel. In the receiver another transformation will operate upon the received transform coefficients to generate a replica of the source output vector.

Transform coders were developed by Kramer and Mathews [19] and later refined by Huang and Schultheiss [1]. Huang and Schultheiss restricted the transformation to be linear. They also assumed that the transform coefficients are quantized by means of Lloyd-Max quantizers [4], [5]. They then determined the optimal linear transformation, which turned out to be the Karhunan-Loeve transformation. They also found the approximately optimal integral bit allocation among the transform coefficients and then quantized each component with a Lloyd-Max quantizer. Specific examples determining the performance of this scheme on Gauss-Markov sources are provided in [1].

Certain refinements of Huang and Schultheiss' results are given in [2] by Segall. Also, various adaptive schemes for block transform quantization of speech and image sources are developed and reported in [3], [7] and [8].

All aforementioned results are based upon the assumption that the transform coefficients are quantized by means of Lloyd-Max quantizers which implicity assume that the number of bits needed to represent the output of each quantizer is given by the base 2 logarithm of the corresponding number of levels. However, recent results on entropy-constrained quantization reveals that substantial performance improvements can be obtained by using optimum entropy-constrained quantizers instead of Lloyd-Max quantizers [6].

In this paper, we will study the performance of block transform quantization schemes in which entropy-constrained zero-memory quantizers are used to encode the transform coefficients. This system, as we will show in subsequent sections, will result in a better rate-distortion performance as compared to the Huang-Schutheiss scheme. The reasons for this performance improvement are two. First, zero-memory entropy-constrained quantizers perform better than zero-memory Lloyd-Max quantizers. Second, in the Huang-Schultheiss scheme the rate associated with each quantizer is constrained to be an integer. The absence of this restriction in the entropy-constrained quantization situation results in additional performance improvements.

In this paper we present an analysis of an entropy-constrained block transform quantization scheme. We develop an algorithmic approach for the design of the quantizers and the entropy assignment among the quantizers. We examine the rate-distortion performance of this scheme on Gauss-Markov and Laplace-Markov sources and present comparisons against the rate-distortion bounds, as well as the performance of other source coding schemes such as Huang-Schultheiss' block transform coding, vector quantization [11] and predictive encoding schemes [10]. Furthermore, we develop a high rate asymptotic analysis of the rate-distortion performance of the entropy-constrained block transform quantization scheme which provides a simple formula for the high bit rate region in terms of the differential entropies of the transform coefficients.

The organization of this paper is as follows. In Section II, we present a description of the block transform encoding scheme. Then, we formulate the entropy-constrained block transform quantization problem in Section III. Also, in this section an algorithm for entropy (rate) assignment and quantizer design is presented. In Section IV the Guass-Markov and Laplace-Markov sources, for which the system performance is obtained, are described and the density of the transform coefficients are computed. This is followed by a presentation of the high bit rate asymptotic results in Section V. In Section VI, numerical results describing the efficacy of the entropy-constrained block transform quantization scheme are presented. Finally, in Section VII a summary and suggestions for future research are included.

II. Preliminaries and Notation:

In the sequel we assume that the source to be encoded can be modeled as a discrete-time stationary first-order autoregressive procss with zero mean and variance σ_X^2, described by

$$X_n = \rho X_{n-1} + W_n \quad ; \quad n=1,2,\ldots \quad , \tag{II.1}$$

where $\{W_n\}$ is a zero-mean sequence of independent and identically distributed random variables with probability density function $p_W(x)$, and variance σ_W^2. Furthermore, we assume that the initial state X_0 is chosen appropriately to insure stationarity of $\{X_n\}$.

This model has been chosen both because it is often a good mathematical model for real-world data (e.g., speech and images), and because it provides a well-understood standard for comparison [1], [9]-[11].

Let us consider a block of length L of consecutive source outputs defined by

$$\underline{X}_n = (X_{(n-1)L+1}, X_{(n-1)L+2}, \ldots, X_{nL})^T \quad , \quad n=1,2,\ldots \tag{II.2}$$

Upon defining the source autocorrelation function

$$\phi_k = E\{X_n X_{n+k}\} \quad ; \quad k=0,\pm 1,\ldots \quad , \tag{II.3}$$

the moment matrix of the vector process $\underline{X}_n$ is defined by

$$\underline{\underline{\Phi}}_X = E\{\underline{X}_n \underline{X}_n^T\} = \begin{bmatrix} \phi_0 & \phi_1 & \cdots & \phi_{L-1} \\ \phi_1 & \phi_0 & \cdots & \phi_{L-2} \\ \vdots & & & \\ \phi_{L-1} & \phi_{L-2} & \cdots & \phi_0 \end{bmatrix} . \tag{II.4}$$

Also, the source power spectral density $\Phi_X(\omega)$, is defined as

$$\Phi_X(\omega) = \sum_{k=-\infty}^{\infty} \phi_k e^{-j\omega k} \quad , \quad \omega \in [-\pi, \pi] \quad , \tag{II.5}$$

with $\phi_k = \phi_{-k}$, $k = 1,2,\ldots$.

In a typical block transform quantization scheme, L successive source outputs are collected to form a block $\underline{X}_n$, or simply $\underline{X}$, as n, due to stationarity, bears no significance in our analysis. The source L-vector $\underline{X}$ is then operated upon by a nonsingular (LxL) transformation matrix $\underline{\underline{A}}$ to generate an L-vector $\underline{Y}$ with uncorrelated components given by

$$\underline{Y} \overset{\Delta}{=} (Y_1, Y_2, \ldots, Y_L)^T = \underline{\underline{A}}\underline{X} \quad . \tag{II.6}$$

The intuitive idea behind this transformation is to exploit the correlation between the input samples $\{X_n\}$ by first generating a set of uncorrelated variables $Y_1, Y_2, \ldots, Y_L$ and then encoding these sample by sample. Thus, each component of $\underline{Y}$ is quantized and encoded by means of a separate zero-memory quantizer. The quantizer operating upon the ith transform coefficient $q_i(\cdot)$, $i=1,2,\ldots,L$ is described by

$$q_i(x) = Q_\ell^{(i)} \quad \text{if} \quad x \in (T_{\ell-1}^{(i)}, T_\ell^{(i)}], \ \ell=1,2,\ldots,N_i, \tag{II.7}$$

in which $T_0^{(i)} < T_1^{(i)} < \ldots < T_{N_i}^{(i)}$ and $Q_1^{(i)}, Q_2^{(i)}, \ldots, Q_{N_i}^{(i)}$ are the <u>threshold levels</u> and the <u>quantization levels</u> associated with the ith quantizer, respectively. Here, the number of levels associated with the ith quantizer is denoted by N_i.

Let us denote the vector of quantized transform coefficients by $\hat{\underline{Y}} = (\hat{Y}_1, \hat{Y}_2, \ldots, \hat{Y}_L)^T$ in which

$$\hat{Y}_i = q_i(Y_i) \ , \quad i=1,2,\ldots,L \ . \tag{II.8}$$

The $\hat{Y}_i$'s, after being coded as a sequence of binary digits, are transmitted via a noiseless channel. In the receiver, an (LxL) transformation matrix $\underline{\underline{B}}$ operates on $\hat{\underline{Y}}$ to generate a replica of $\underline{X}$, say $\hat{\underline{X}}$, given by

$$\hat{\underline{X}} = \underline{\underline{B}}\hat{\underline{Y}} \quad . \tag{II.9}$$

At this point, the problem can be described as follows. Given a fixed number of bits available for representing a source sample, say R, our goal is to minimize the average per-symbol squared-error distortion incurred in the above operation, described by

$$D = \frac{1}{L} E\{ \| \underline{X} - \hat{\underline{X}} \|^2 \} , \tag{II.10}$$

in which $E\{\cdot\}$ is the expectation operation and $\|\underline{X}\|$ denotes the norm of the vector $\underline{X} \in R^L$. The minimization of D will be achieved by adjustment of the matrices $\underline{\underline{A}}$ and $\underline{\underline{B}}$, and the appropriate design of the quantizers $q_i(\cdot)$, $i=1,2,\ldots,L$.

The above problem was first studied by Huang and Schultheiss [1], in which the block transform quantization problem was investigated for stationary Gaussian sources with the assumption that the quantizers used for encoding the transform coefficients are Lloyd-Max quantizers. In [1], the optimum choices of the transformations $\underline{\underline{A}}$ and $\underline{\underline{B}}$ are determined. More specifically, it is shown that, (i) the best transformation matrix $\underline{\underline{B}}$ is the inverse of the matrix $\underline{\underline{A}}$, and (ii) the best transformation matrix $\underline{\underline{A}}$ is the matrix whose rows are the orthonormalized eigenvectors of $\underline{\underline{\Phi}}_X$, the covariance matrix of $\underline{X}$ (i.e., the Karhunen-Loeve transformation).

Due to the fact that the Karhunen-Loeve transformation is a unitary transformation (i.e., $\underline{\underline{A}}^T = \underline{\underline{A}}^{-1}$), we have

$$D = \frac{1}{L} E\{ \| \underline{X} - \hat{\underline{X}} \|^2 \} = \frac{1}{L} E\{ \| \underline{Y} - \hat{\underline{Y}} \|^2 \} =$$

$$= \frac{1}{L} \sum_{i=1}^{L} \lambda_i D_i , \tag{II.11}$$

where λ_i is the ith eigenvalue of $\underline{\underline{\Phi}}_X$ (or, the variance of the ith transform coefficient) and D_i is the normalized (to the variance of the ith coefficient) average squared-error associated with the quantization of the ith coefficient.

If R_i, $i=1,2,\ldots,L$, designates the rate associated with the ith quantizer, the problem of quantizer design is that of minimizing D given by (II.11) while the overall rate requirement is satisfied, i.e.,

$$\frac{R_1 + R_2 + \ldots + R_L}{L} = R \ . \qquad \text{(II.12)}$$

Assuming that Lloyd-Max quantizers are used for encoding the transform coefficients, i.e., R_i is the logarithm of the number of levels of the quantizer, Huang and Schultheiss [1] developed an algorithm to find <u>approximately</u> the best bit assignment among the L quantizers‡.

While Huang and Schultheiss', as well as all subsequent work on block transform quantization [2], [7], [8] are based on the use of Lloyd-Max quantizers, recent work on optimum entropy-constrained zero-memory quantization [6] has revealed that, from a rate-distortion theoretic point of view, substantial performance improvements can be obtained by optimum entropy-constrained quantization compared to the Lloyd-Max scheme. This has become our motivation for studying the performance of block transform quantization schemes which employ entropy-constrained quantizers to encode the transform coefficients. In what follows, we provide a precise formulation of the problem and develop an algorithm for optimum system design.

‡ Note that in [1] only Gaussian sources are considered for which the transform coefficients are also Gaussian and hence the problem of optimum quantizer design reduces to that of optimum bit assignment, since the quantizer design with a known number of levels for Gaussian sources is known [4], [5].

III. Problem Statement and Algorithm:

Let us suppose that the ith transform coefficient is quantized by means of a zero-memory optimum entropy-constrained quantizer with output entropy H_i bits/sample. Thus, according to Shannon's noiseless source coding arguments [15], the average number of bits necessary to represent the output of the ith quantizer is H_i bits/sample. Let $D_i(H_i)$ be the variance-normalized average squared-error distortion associated with quantizing the ith coefficient. The overall average distortion is then given by‡

$$D(\underline{H}) = \frac{1}{L} \sum_{i=1}^{L} \lambda_i D_i(H_i) \quad , \tag{III.1}$$

in which $\underline{H} = (H_1, H_2, \ldots, H_L)$. We wish to put a constraint, say R, on the overall average output entropy. Therefore, our problem is to find an optimum vector $\underline{H}^* = (H_1^*, H_2^*, \ldots, H_L^*)$ that minimizes

$$D(\underline{H}) = \frac{1}{L} \sum_{i=1}^{L} \lambda_i D_i(H_i) \ , \tag{III.2.a}$$

subject to

$$\frac{1}{L} \sum_{i=1}^{L} H_i \leq R \ . \tag{III.2.b}$$

Notice that the vector of optimum entropy (rate) assignment $\underline{H}^*$, plays the same role as the optimum bit assignment vector in [1]. Solving the above nonlinear constrained programming problem, in general, is a formidable task. This is, primarily, due to the fact that explicit relationship between $D_i(H_i)$ and H_i is not known. To alleviate this problem it is possible to express both the average distortion and the output entropy associated with individual coefficients in

‡ In (III.1) and subsequent sections we will assume that the Karhunen-Loeve transformation is used and hence λ_i, i=1,2,...,L, is the variance of the ith coefficient. Extension to other non-optimal transformations is straightforward.

terms of their respective quantization thresholds $T_\ell^{(i)}$, $\ell=0,1,\ldots,N_i$ and quantization levels $Q_\ell^{(i)}$, $\ell=1,2,\ldots,N_i$. Solving this problem involves the determination of $\sum_{i=1}^{L}(2N_i-1)$ threshold levels and quantization levels. Considering the fact that to get good performance results some of the N_i's need to be fairly large, and also to fully exploit the source correlation, large values of L might be required, it becomes obvious that, in general, the number of variables could become prohibitively large which, in turn, makes the problem extremely difficult.

To alleviate the aforementioned problem we have made use of some observations about the performance of zero-memory entropy-constrained quantizers. Specifically, it is shown by Gish and Pierce [13] that at high bit rates and for large number of quantization levels, the optimum quantizer has uniformly spaced levels. Furthermore, experimental results of Farvardin and Modestino [6] have revealed that, for a wide class of memoryless sources, even at low bit rates, uniform-threshold quantizers (i.e., N-level quantizers with $T_\ell - T_{\ell-1} = \Delta$, $\ell=2,3,\ldots,N-1$) perform very close to the optimum performance when N is sufficiently large. Indeed, in the absence of a limit on the number of levels, the difference between the performance of the optimum uniform-threshold quantizer and the optimum quantizer is observed to be negligible [6]. The above observation and the need for simplification of our problem has led us to consider a block transform quantization scheme with entropy constraint in which zero-memory uniform-threshold quantizers are employed in quantization of the transform coefficients.

Let us suppose that the variance-normalized average squared-error distortion and output entropy associated with the uniform-threshold quantizer with stepsize Δ_i operating on the ith transform coefficient are given by $\hat{D}_i(\Delta_i)$ and $\hat{H}_i(\Delta_i)$, bits/sample, respectively. Then, the constrained nonlinear programming problem described by (III.2) will be translated to finding a vector of optimum stepsizes $\underline{\Delta}^* = (\Delta_1^*, \Delta_2^*, \ldots, \Delta_L^*)$ that minimizes

$$\hat{D}(\underline{\Delta}) = \frac{1}{L} \sum_{i=1}^{L} \lambda_i \hat{D}_i(\Delta_i) \quad , \tag{III.3.a}$$

subject to

$$\hat{H}(\underline{\Delta}) = \frac{1}{L} \sum_{i=1}^{L} \hat{H}_i(\Delta_i) \leq R \quad . \tag{III.3.b}$$

For comparison purposes, we will denote the optimum distortion-rate performance of this scheme by $\hat{D}(R) \equiv \hat{D}(\underline{\Delta}^*)$, while reserving $D(R)$ to denote the limiting source distortion-rate function [14].

Necessary conditions for the solution of the problem described in (III.3) follow from the Kuhn-Tucker Theorem [12]. Specifically, letting $\underline{\Delta}^*$ be a <u>local solution</u> of (III.3), there must exist a scalar $\lambda \geq 0$ such that‡

$$\nabla\hat{D}(\underline{\Delta}^*) + \lambda\nabla\hat{H}(\underline{\Delta}^*) = \underline{0} \quad , \tag{III.4.a}$$

and

$$\lambda[\hat{H}(\underline{\Delta}^*) - R] = 0 \quad , \tag{III.4.b}$$

where λ is the Lagrange multiplier.

Notice that for the ith N-level‡‡ quantizer the explicit expressions for $\hat{D}_i(\Delta_i)$ and $\hat{H}_i(\Delta_i)$ are simply written as

$$\hat{D}_i(\Delta_i) = \sum_{\ell=1}^{N} \int_{T_{\ell-1}^{(i)}}^{T_\ell^{(i)}} (x - Q_\ell^{(i)}(\Delta_i))^2 \, p_i(x)dx \; , \tag{III.5.a}$$

‡ The operator $\nabla(\cdot)$ indicates the gradient of a scalar function with respect to its arguments.

‡‡ Notice that so long as the number of levels is sufficiently large, its actual value does not play an important role. Therefore, we will assume, from now on, that all L quantizers have the same number of levels N.

and

$$\hat{H}_i(\Delta_i) = -\sum_{\ell=1}^{N} P_\ell^{(i)}(\Delta_i)\log_2 P_\ell^{(i)}(\Delta_i), \text{ bits/sample,} \qquad \text{(III.5.b)}$$

in which

$$\begin{cases} T_\ell^{(i)} = (\ell - N/2)\Delta_i \quad , & \ell=1,2,\ldots,N-1, \\ T_N^{(i)} = -T_0^{(i)} = \infty \quad , & \end{cases} \qquad \text{(III.5.c)}$$

and

$$Q_\ell^{(i)}(\Delta_i) = \frac{\int_{T_{\ell-1}^{(i)}}^{T_\ell^{(i)}} x p_i(x)dx}{\int_{T_{\ell-1}^{(i)}}^{T_\ell^{(i)}} p_i(x)dx} \quad , \ell=1,2,\ldots,N \ , \qquad \text{(III.5.d)}$$

where $p_i(x)$ is the probability density function (p.d.f) of the ith transform coefficient and $P_\ell^{(i)}(\Delta_i)$ is the probability that the ith transform coefficient falls in the ℓth bin of its respective quantizer. It must be mentioned that in (III.5) and all subsequent sections, following the arguments in [6], the uniform-threshold quantizers are assumed to be symmetric with respect to the origin.

For $\lambda > 0$, (III.4) yields a set of L <u>necessary</u> conditions for optimality described by

$$f_i(\Delta_i^*;\lambda) \triangleq \left.\frac{\partial \hat{D}_i(\Delta_i)}{\partial \Delta_i}\right|_{\Delta_i=\Delta_i^*} + \frac{\lambda}{\lambda_i} \cdot \left.\frac{\partial \hat{H}_i(\Delta_i)}{\partial \Delta_i}\right|_{\Delta_i=\Delta_i^*} = 0 \ , \ i=1,2,\ldots,L \ . \qquad \text{(III.6)}$$

Using (III.5), an expression for $f_i(\Delta_i;\lambda)$ in terms of Δ_i and λ can be found which is included in Appendix A. It is important to note that for λ=0 the problem reduces to that of minimizing the average distortion $\hat{D}(\underline{\Delta})$ with no constraint on $\hat{H}(\underline{\Delta})$. This, essentially, corresponds to the minimization of the average distortion in each coordinate. The resulting $\hat{D}(\underline{\Delta})$ and $\hat{H}(\underline{\Delta})$ yield the minimum distortion point (for a fixed value of N) on the distortion-rate performance curve $\hat{D}(R)$.

With the above description, our algorithm works in the following manner.

<u>Algorithm</u>:

i) Fix the number of quantization levels N, and the block size L.

ii) Compute the variance λ_i and the p.d.f. $P_i(x)$ of the transform coefficients, $i=1,2,\ldots L$ (see Section IV). Set $\lambda=0$.

iii) Solve (III.6) for the optimum Δ_i, say Δ_i^*, $i=1,2,\ldots,L$. This can be done using numerical techniques for computing the root(s) of a nonlinear equation. Compute $\hat{D}(\underline{\Delta}^*)$ and $\hat{H}(\underline{\Delta}^*)$. If $\hat{H}(\Delta^*) > H_t$ go to (iv). Otherwise stop. Here, $H_t>0$ is a small rate used for terminating the algorithm.

iv) Set $\lambda = \lambda + \delta$ ($\delta>0$ is the value by which λ is incremented at each iteration). Go to (iii).

The collection of points $(\hat{D}(\underline{\Delta}^*), \hat{H}(\underline{\Delta}^*))$ obtained in step (iii) of the algorithm determine the system's distortion-rate performance curve.

One issue that we have neglected so far is that of computing the p.d.f. of the transform coefficients. In the following section, we shall confine attention to two different source distributions for which the numerical results are obtained, and will describe the procedure for computing the p.d.f. of the transform coefficients.

IV. Source Models:

As mentioned in Section II, our attention in studying the entropy-constrained block transform quantization scheme is focused on stationary first-order autoregressive sources described by (II.1). In particular, we are interested in first-order Gauss-Markov and Laplace-Markov sources. Let us consider the two sources separately.

A.) Gauss-Markov Source:

The Gauss-Markov source is defined according to (II.1) where $\{W_n\}$ is a zero-mean sequence of Gaussian random variables with variance $\sigma_W^2 = 1$. It turns out that, under appropriate initial conditions, $\{X_n\}$ is a stationary zero-mean Gaussian random process with variance $\sigma_X^2 = 1/(1-\rho^2)$. Since the Karhunen-Loeve transformation $\underline{A}$ is a linear transformation, the transform coefficients will be Gaussian variates with zero-mean and variance λ_i, $i=1,2,\ldots,L$. Thus, the p.d.f. of the ith transform coefficient, is given by:

$$p_i(x) = \frac{1}{\sqrt{2\pi\lambda_i}} \exp\{-x^2/2\lambda_i\}, \; i=1,2,\ldots,L, \; -\infty < x < \infty \; . \qquad \text{(IV.1)}$$

B.) Laplace-Markov Source:

The stationary Laplace-Markov process is defined in [17]. By this we mean a first-order Markov process with a Laplacian marginal distribution. It is straightforward to determine the p.d.f. of the process $\{W_n\}$ generating the Laplace-Markov source.

Upon taking characteristic functions in (II.1), we have

$$\psi_X(z) = \psi_X(\rho z)\psi_W(z) \quad , \qquad \text{(IV.2)}$$

in which $\psi_X(\cdot)$ and $\psi_W(\cdot)$ denote the characteristic functions of X_n and W_n, respectively. If, we require

$$p_X(x) = \frac{1}{2} e^{-|x|} \quad , \quad -\infty < x < \infty \quad , \qquad \text{(IV.3.a)}$$

as assumed, we have

$$\psi_X(z) = \frac{1}{1+z^2}, \quad -\infty < z < \infty \quad , \tag{IV.3.b}$$

which, in turn, combined with (IV.2) yields

$$\psi_W(z) = \frac{1+\rho^2 z^2}{1+z^2} = \rho^2 + (1-\rho^2)\frac{1}{1+z^2} \quad , \tag{IV.4.a}$$

and hence

$$p_W(x) = (1-\rho^2)\,\frac{1}{2}\,e^{-|x|} + \rho^2\delta(x) \ , \ -\infty < x < \infty \ . \tag{IV.4.b}$$

That is, $\{X_n\}$ represents a source generating a random variable whose value is either zero with probability ρ^2 or Laplacian distributed with probability $(1-\rho^2)$. Here, again under appropriate initial conditions, the process $\{X_n\}$ is a stationary zero-mean Laplacian process with variance $\sigma_X^2 = 2$.

The significance of studying Laplace-Markov sources is the observation, made by several researchers (e.g., [18]) that speech signals possess a marginal density reasonably close to a Laplacian density.

Let us now proceed to compute the marginal p.d.f. of the transform coefficients for the Laplace-Markov input. Consider a typical source L-vector $\underline{X} = (X_1, X_2, \ldots, X_L)$. Using (II.1) it is straightforward to show that

$$X_j = \rho^j X_0 + \sum_{k=1}^{j} \rho^{j-k} W_k \ , \ j=1,2,\ldots,L \ . \tag{IV.5}$$

Then, using (II.6) and denoting the (i,j)th element of $\underline{\underline{A}}$ by $a_{i,j}$, $i=1,2,\ldots,L$, $j=1,2,\ldots,L$, we have

$$Y_i = \Big[\sum_{j=1}^{L} a_{ij}\rho^j\Big]X_0 + \sum_{j=1}^{L} a_{ij}\sum_{k=1}^{j}\rho^{j-k}W_k, \ i=1,2,\ldots,L \quad , \tag{IV.6}$$

which can be further simplified as

$$Y_i = \alpha_i X_0 + \sum_{k=1}^{L} \beta_{k,i} W_k, \quad i = 1,2,\ldots,L, \tag{IV.7.a}$$

where

$$\alpha_i \triangleq \sum_{j=1}^{L} a_{ij}\rho^j \quad , \; i=1,2,\ldots,L \quad , \tag{IV.7.b}$$

and

$$\beta_{k,i} \triangleq \sum_{j=k}^{L} a_{ij}\rho^{j-k} \quad , \; k=1,2,\ldots,L, \; i=1,2,\ldots,L \; . \tag{IV.7.c}$$

Now, using the fact that X_0, and W_k, $k=1,2,\ldots,L$ are all independent from each other, we can write

$$\psi_{Y_i}(z) = \psi_{X_0}(\alpha_i z) \cdot \prod_{k=1}^{L} \psi_W(\beta_{k,i} z), \quad i=1,2,\ldots,L \; . \tag{IV.8}$$

Therefore, using (IV.3.b) and (IV.4.a), we have the following explicit formula for the characteristic function of Y_i:

$$\psi_{Y_i}(z) = \frac{1}{1+\alpha_i^2 z^2} \cdot \prod_{k=1}^{L} \frac{1+\rho^2\beta_{k,i}^2 z^2}{1+\beta_{k,i}^2 z^2} \quad , \; i=1,2,\ldots,L \; . \tag{IV.9}$$

Now, noting that $\alpha_i = \rho\beta_{1,i}$, $i=1,2,\ldots,L$, (IV.9) simplifies to

$$\psi_{Y_i}(z) = \frac{\prod_{k=2}^{L} (1+\rho^2\beta_{k,i}^2 z^2)}{\prod_{k=1}^{L} (1+\beta_{k,i}^2 z^2)} , \quad i=1,2,\ldots,L \; . \tag{IV.10}$$

Upon assuming that $\beta_{k,i}^2 \neq \beta_{\ell,i}^2$ ‡ for all $k\neq\ell$, we can rewrite (IV.10) as the following sum:

‡ If this condition is not satisfied, then multiple poles exist in (IV.10) which should be dealt with properly. An example of this, is when $L=4$ and $\rho=0.5$ in which case three of the poles coincide.

$$\psi_{Y_i}(z) = \sum_{k=1}^{L} \frac{A_{k,i}}{1+\beta_{k,i}^2 z^2} , \qquad i=1,2,\ldots,L , \tag{IV.11.a}$$

in which

$$A_{\ell,i} \triangleq (1 + \beta_{\ell,i}^2 z^2)\psi_{Y_i}(z) \Big|_{z^2 = -1/\beta_{\ell,i}^2} , \quad \ell = 1,2,\ldots,L. \tag{IV.11.b}$$

Having expressed the characteristic function of the random variable Y_i as (IV.11.a), its p.d.f. can easily be written as

$$p_i(x) = \sum_{k=1}^{L} A_{k,i} \frac{1}{2|\beta_{k,i}|} e^{-|x|/|\beta_{k,i}|} , \quad i=1,2,\ldots,L . \tag{IV.12}$$

In other words, the p.d.f. of the transform coefficients can be expressed as a weighted sum of the p.d.f.'s of L zero-mean Laplacian p.d.f.'s with variances $2\beta_{k,i}^2$, k=1,2,...,L. The simple formula for the p.d.f. of the ith transform coefficient is extremely useful in simplifying the computation of the performance of entropy-constrained block transform quantization scheme. Before presenting the numerical results, in the following section we present a high rate asymptotic analysis of the system performance.

V. Asymptotic Results:

It is shown by Gish and Pierce [13] that in zero-memory quantization of memoryless sources, when the number of quantization levels is large and when the output entropy is high, under certain mild conditions on the source p.d.f., the variance-normalized average distortion and output entropy can be expressed as a function of the quantizer stepsize through the following simple relationship

$$\hat{D}_i(\Delta_i) = \frac{\Delta_i^2}{12\lambda_i} \quad , \tag{V.1.a}$$

and

$$\hat{H}_i(\Delta_i) = h_i - \log_2 \Delta_i \ , \text{ bits/sample,} \tag{V.1.b}$$

respectively, where h_i is the differential entropy associated with the ith transform coefficient given by

$$h_i = -\int_{-\infty}^{\infty} p_i(x)\log_2 p_i(x)dx, \text{ bits/sample.} \tag{V.1.c}$$

Furthermore, eliminating Δ_i in (V.1) yields the following high rate asymptotic performance for zero-memory entropy-constrained quantizers

$$\hat{D}_i(\hat{H}_i) = \frac{1}{12\lambda_i} 2^{2(h_i - \hat{H}_i)} \ . \tag{V.2}$$

Replacing (V.1.a) and (V.1.b) in (III.6) yields

$$\Delta_i^* = \left[\frac{6\lambda}{\ln 2}\right]^{1/2} , \ i=1,2,\ldots,L. \tag{V.3}$$

Noting that from (III.4.b), $\lambda > 0$ implies $\hat{H}(\underline{\Delta}^*) = R$, we must have

$$\hat{H}(\underline{\Delta}^*) = \frac{1}{L} \sum_{i=1}^{L} (h_i - \log_2 \Delta_i^*) = R \ , \tag{V.4}$$

which, in turn, implies

$$\lambda = \frac{\ln 2}{6} 2^{2(h-R)} \quad , \tag{V.5}$$

in which h is defined by

$$h \triangleq \frac{1}{L} \sum_{i=1}^{L} h_i \ . \tag{V.6}$$

Combining (V.1.a), (V.3), (V.5) and (III.3.a) yields the following expression for the distortion-rate performance of our system at the high bit rate region,

$$\hat{D}(R) = \frac{1}{12} \; 2^{2(h-R)} \quad . \tag{V.7}$$

It is interesting to note that $\hat{D}(R)$ in (V.7) is very similar to $D_i(\hat{H}_i)$ in (V.2) with the exception of h defined by (V.6) which, in a sense, is the average per-symbol differential entropy.

Equation (V.7) is very useful in assessing the distortion-rate performance of the block transform coding scheme against the optimal source distortion rate-rate function and the optimal performance of other source coding schemes such as Huang and Schultheiss' block transform quantization [1] scheme or optimum predictive encoding [10] at high bit rates.

In general, as indicated in Section IV, computation of the p.d.f. and hence the differential entropy of the transform coefficients is difficult. However, for the Gaussian source, the transform coefficients remain Gaussian and, therefore, a good deal of analytical results can be obtained, which is presented in the following.

Let us first compute $\hat{D}(R)$ for the Gaussian source. The differential entropies h_i, i=1,2,...,L are given by

$$h_i = \frac{1}{2} \log_2 2\pi e \lambda_i \ , \ i=1,2,\ldots,L \quad , \text{ bits/sample,} \tag{V.8}$$

and hence

$$h = \frac{1}{2} \log_2 2\pi e \left[\prod_{i=1}^{L} \lambda_i \right]^{1/L} . \tag{V.9}$$

Therefore, (V.7) simplifies to

$$\hat{D}(R) = \frac{\pi e}{6} \left[\prod_{i=1}^{L} \lambda_i \right]^{1/L} 2^{-2R} . \tag{V.10}$$

On the other hand, the high bit rate (or, low distortion, i.e., $D < \lambda_i$, $i=1,2,\ldots,L$) distortion-rate function of the source is given by

$$D_L(R) = \left[\prod_{i=1}^{L} \lambda_i \right]^{1/L} 2^{-2R} , \tag{V.11}$$

where the subscript L is used to indicate that the distortion-rate function is obtained based on source blocks of length L, [14], [16]. Of course, as L tends to infinity $D_L(R)$ converges to the source distortion-rate function.

To provide a basis for comparison with Huang and Schultheiss' scheme [1], in what follows we present an asymptotic analysis of the performance of the scheme described in [1].

In [1], Lloyd-Max quantizers are used to encode the transform coefficients. The distortion-rate performance of the ith Lloyd-Max quantizer at high bit rates can be approximated by means of companding arguments [3]. Denoting the variance-normalized distortion-rate performance of the ith Lloyd-Max quantizer by $\tilde{D}_i(R_i)$, we have [3]

$$\tilde{D}_i(R_i) = \frac{2}{3\lambda_i} \left[\int_0^{\infty} [p_i(x)]^{1/3} dx \right]^3 2^{-2R_i} . \tag{V.12}$$

For the Gaussian source, however, we have

$$p_i(x) = \frac{1}{\sqrt{2\pi\lambda_i}} \exp\{-x^2/2\lambda_i\} , \quad i=1,2,\ldots,L , \tag{V.13}$$

which results in the following expression for $\tilde{D}_i(R_i)$,

$$\tilde{D}_i(R_i) = \frac{\pi\sqrt{3}}{2} 2^{-2R_i} . \tag{V.14}$$

Now the problem of minimizing the distortion is that of finding $\underline{R}^* = (R_1^*, R_2^*, \ldots, R_L^*)$ which minimizes

$$\tilde{D}(\underline{R}) = \frac{1}{L} \sum_{i=1}^{L} \frac{\pi\sqrt{3}}{2} \lambda_i \, 2^{-2R_i} \quad , \qquad \text{(V.15.a)}$$

subject to

$$\frac{1}{L} \sum_{i=1}^{L} R_i = R \quad . \qquad \text{(V.15.b)}$$

Using the same approach as before, it can be shown that the best R_i's are given approximately‡ by

$$R_i^* = R + \frac{1}{2} \log_2 \frac{\lambda_i}{\left[\prod_{i=1}^{L} \lambda_i\right]^{1/L}} \quad , \quad i=1,2,\ldots,L \ , \qquad \text{(V.16)}$$

which, in turn, results in the best performance $\tilde{D}(R) \triangleq \tilde{D}(\underline{R}^*)$ given by

$$\tilde{D}(R) = \frac{\pi\sqrt{3}}{2} \left[\prod_{i=1}^{L} \lambda_i\right]^{1/L} 2^{-2R} \quad . \qquad \text{(V.17)}$$

Comparison of (V.10), (V.11) and (V.17) reveals the fact that for a fixed transmission rate R, the difference between the source rate-distortion function and the entropy-constrained block transform coding scheme is $10\log_{10} \frac{\pi e}{6} = 1.53$ dB, while the gain of entropy-constrained block transform over Huang and Schultheiss' scheme is $10\log_{10} \frac{\pi\sqrt{3}/2}{\pi e/6} = 2.81$ dB. Equivalently, these differences, for a fixed distortion D, are 0.255 bits/sample and 0.467 bits/sample, respectively.

Finally, it is worth mentioning that similar differences hold regardless of the actual value of L. Also, the limiting results for $L \to \infty$ can be obtained using the Toeplitz distribution theorem [14], [16]. More precisely, to obtain the

‡ The R_i^*'s given by (V.16) are not necessarily integers. Therfore, these R_i^*'s should be somehow adjusted to their nearest integer values [1], [7].

distortion-rate performance of the optimum entropy-constrained block transform quantization scheme in the high bit-rate region and for the limit of large L, we observe from (V.10) that

$$\log_2 \hat{D}(R) = \log_2 \frac{\pi e}{6} + \frac{1}{L} \sum_{i=1}^{L} \log_2 \lambda_i - 2R , \qquad \text{(V.18)}$$

which then using the Toeplitz distribution theorem [14] implies

$$\lim_{L \to \infty} \log_2 \hat{D}(R) = \log_2 \frac{\pi e}{6} + \frac{1}{2\pi} \int_{-\pi}^{\pi} \log_2 \Phi_X(\omega) d\omega - 2R , \qquad \text{(V.19)}$$

in which $\Phi_X(\omega)$ is the source power spectral density as described in (II.5).

It is instructive to mention, before closing this section, that the limiting value of the expression in (V.19) for first-order autoregressive Gaussian sources coincides with the asymptotic performance of optimum entropy-constrained predictive encoding schemes (equation (63a) in [10]). In other words, for first-order Gauss-Markov sources, the entropy-constrained block transform quantization and the entropy-constrained DPCM yield the same high rate performance.

In the following section we present the numerical results obtained from our algorithm for the first-order Gauss-Markov and Laplace-Markov sources and make appropriate comparisons against other coding schemes.

VI. Numerical Results:

We have used the algorithm described in Section III to obtain the optimum distortion-rate performance of the entropy-constrained block transform quantization scheme with uniform-threshold quantizers. Performance results are obtained for the Gauss-Markov and the Laplace-Markov source for different values of the block size L, the correlation coefficient ρ and the transmission rate R. In the following we present a description of various system parameters and the results.

In the outset, it should be mentioned that in all cases the number of quantization levels is fixed and equal to N=35. This assumption, although rather arbitrary, is based on the results in [6]. It is shown in [6] that for most sources N=35 is sufficiently large for close-to-optimal performance at low bit rates. It is conceivable that at higher bit rates (e.g., R=3 bits/sample) and especially for the Laplace-Markov source for which the transform coefficients possess more broad-tailed densities, better performance results could be obtained with a larger number of quantization levels. It is also important to mention that since different transform coefficients are quantized at different rates, one could use a smaller number of levels for those coefficients that are quantized at lower bit rates. This will be useful in complexity reduction when the system is to be implemented.

In Tables 1-3 we have summarized the performance results of the entropy-constrained block transform quantization scheme for Gauss-Markov sources. Where available, the performance of the Huang-Schultheiss scheme [1], the vector quantization scheme [11],[20] and the optimum entropy-constrained DPCM [10], is also included in these tables. Several comments about these results are in order.

First, it is important to note that in all cases the entropy-constrained scheme outperforms the Huang-Schultheiss scheme. This difference in performance ranges from 0.18 dB for ρ=0.5, L=1, and R=1 bit/sample, to 2.53 dB for ρ=0.5, L=2, and R=3 bits/sample. In general, the performance improvements are more noticeable at higher bit rates. These performance improvements are due to two factors. First, entropy-constrained zero-memory quantizers outperform Lloyd-Max quantizers. Secondly, the Huang-Schultheiss scheme is designed based upon the assumption that the rate associated with each quantizer is an integer. This constraint, which does not exist in the entropy-constrained scheme, results in additional suboptimality and hence inferior performance. It is important to recall, as we established in Section V, that if there is no restriction on the number of quantization levels of the entropy-constrained scheme, then at high bit rates it outperforms the Huang-Schultheiss scheme by 2.81 dB.

The second interesting observation is that in all cases for which the results are available for the full-search vector quantizer performance [20], the entropy-constrained block transform quantization scheme has yielded superior performance results. This observation, together with the fact that block transform quantization schemes are more amenable to implementation further underscores their attractiveness for easily implementable high-performance data compression schemes.

Finally, it is of importance to compare the performance of the entropy-constrained block transform quantization scheme against the optimum entropy-constrained DPCM. This is especially important at the low bit rate region because, as indicated in Section V, for high bit rates the two schemes offer the same performance. Results of Tables 1-3 reveal that at rate R=1 bit/sample the entropy-constrained block transform scheme offers substantial performance improvements over DPCM. Interestingly, this takes place for relatively low values of L. Indeed, in all cases the performance of the system with L=4 supercedes that of the optimum DPCM. The performance improvements are more tangible for higher values of ρ and L. For instance, for ρ=0.9 and L=8, the entropy-constrained block transform scheme at R=1 bit/sample outperforms the DPCM scheme by 1.52 dB. We expect that this difference will become noticeably larger for larger values of L, especially for large values of ρ for which there is more memory in the source.

For the sake of completion, the rate-distortion performance of entropy-constrained block transform quantizers are illustrated in Fig.'s 1-3. In these Figures we have also included the asymptotic performance curves given by (V.10), as well as the source rate-distortion performance curve described in [14].

Careful examination of the results in Figures 1-3 reveals that, in general, there is good agreement between the numerical results and the asymptotic formula in the high bit rate region. Slight deviation from the asymptotic performance (dashed curves in Figures 1-3) for $\rho = 0.8$ and $\rho = 0.9$ is a direct consequence of the limitation on the number of quantization levels. More specifically, it turns out that for large values of ρ, R and L, the entropy assigned to the first component is very large (e.g., for $\rho = 0.9$, $L = 8$ and $R = 3$ bits/sample, the entropy assigned to the first component is $\hat{H}_1 = 4.58$ bits/sample). Therefore, in order to achieve an average distortion close to optimum for this component, a large number of quantization levels (larger than $N = 35$) will become necessary.

Let us now consider similar results for the Laplace-Markov source. These results are summarized in Tables 4-6. In these tables we have also included upper and lower bound on the source rate-distortion function, denoted by $R_L(D)$ and $R_U(D)$. The upper bound is the well-known Gaussian upper bound [14] and the lower bound is based on the autoregressive lower bound [21]. This lower bound is determined in terms of the rate-distortion function of the innovation sequence $\{W_n\}$ described by (IV.4.b). This rate-distortion function, in turn, has been lower bounded by the Carter-Neuhoff composite lower bound [22], the computation of which is based on the Blahut algorithm [23]. Details of this lower bound can be found in [10].

As one could easily see from the results in Tables 4-6, the performance improvements obtained by the entropy-constrained block transform coding over the Huang-Schultheiss scheme are much more noticeable for the Laplace-Markov source. This is, primarily, due to the fact that for the Laplace-Markov source, the transform coefficients possess more broad-tailed densities than for the Gauss-Markov source.

It is shown in [6] that the performance improvement obtained by entropy-constrained quantization is more noticeable for more broad-tailed densities. Also, as our results in Tables 4-6 indicate, noticeable gains over the vector quantization scheme are obtained. However, with the limited available information on the DPCM performance, it appears that at low bit rates (i.e., R = 1 bit/sample) the DPCM offers very good performance, especially for larger values of ρ.

Let us note in Table 4-6 that in some cases the performance of the system with L=8 is lower than that with L=4. This needs some explanation. When we increase the blocksize from L=4 to L=8, the variance of the first transform coefficient inceases which, generally, means a larger value of optimal rate for the quantization of the first coefficient. However, because of the fact that the number of quantization levels is fixed, higher bit rate could result in noticeable deviation from optimal performance which, in turn, could result in an overall performance degradation[6]. This problem can be overcome by increasing the number of quantization levels. Although not included here for the sake of consistency, for the Laplace-Markov source numerical results with larger number of quantization levels have been obtained which have resolved this problem.

In light of all the numerical results, it can be concluded that the entropy-constrainted block transform coding scheme performs better than the Huang-Schultheiss and the vector quantization scheme for both Gauss-Markov and Laplace-Markov sources at least for the cases studied here. For Gauss-Markov sources, its performance coincides with that of DPCM at high bit rates and at low bit rates its performance supercedes that of DPCM. For the Laplace-Markov source, at low bit rates the DPCM system performs very well, and very large values of L may be needed for the block transform coder to yield an equally good performance.

VII. Summary and Conclusions:

In this paper we have studied the problem of optimum entropy-constrained block transform quantization of first-order autoregressive sources. Necessary conditions for the optimality of this scheme are developed and an algorithm for optimal entropy assignment among different quantizers used for encoding the transform coefficients is developed. Also, for first-order Gauss-Markov and Laplace-Markov sources, the distribution of the transform coefficients are obtained.

Asymptotic results, similar to those developed by Gish and Pierce [13] in the memoryless case, are developed. These results that agree favorably with our numerical results in the Gaussian case, imply that at high bit rates there is only 0.255 bits/sample performance penalty, exactly similar to the DPCM scheme [10].

Our numerical results indicate performance improvements over the Huang-Schultheiss scheme and the vector quantization scheme in all cases for which results are available. The performance improvements are much more noticeable for the Laplace-Markov source. To compare our system with the DPCM scheme, we should consider the two sources separately. For the Gauss-Markov source, our results indicate noticeable improvements over DPCM at low bit rates. For the Laplace-Markov source for which the available DPCM results are very limited, the DPCM outperforms our block transform quantization at low bit rates. At high bit rates, results on the DPCM performance are not available.

We should mention here that although the entropy-constrained block transform quantization scheme offers superior performance to most other coding schemes, it suffers from two problems. First, in our analysis of the system we have assumed that the rate associated with each transform coefficient equals the output entropy of the corresponding quantizer. This suggests the use of some type of variable-length coding for transmission of the transform coefficients. But, transmission of variable-length codes introduces two problems. First, it introduces buffer overflow/underflow problems [24], [25], and secondly it results in the propagation of channel errors. To overcome these problems, entropy-constrained quantization can be replaced by permutation coding which is proved to offer the same performance [26]. This work is currently under investigation.

Appendix A

Computation of $f_i(\Delta;\lambda)$

Recall from Section III that $f_i(\Delta;\lambda)$ is given by

$$f_i(\Delta;\lambda) = \frac{\partial \hat{D}_i(\Delta)}{\partial \Delta} + \frac{\lambda}{\lambda_i} \cdot \frac{\partial \hat{H}_i(\Delta)}{\partial \Delta} \quad , \tag{A.1}$$

in which $\hat{D}_i(\Delta)$ and $\hat{H}_i(\Delta)$ are described by (III.5). We will assume throughout that N is an odd integer representing the number of quantization levels for all L quantizers and that all quantizers are symmetric. Then, it is straightforward to come up with

$$\hat{D}_i(\Delta) = \sigma_i^2 - 2 \sum_{\ell=1}^{(N-1)/2} P_\ell^{(i)}(\Delta) \left[Q_\ell^{(i)}(\Delta)\right]^2 \quad , \tag{A.2.a}$$

and

$$\hat{H}_i(\Delta) = - P_{(N+1)/2}^{(i)}(\Delta)\log_2 P_{(N+1)/2}^{(i)}(\Delta) - 2 \sum_{\ell=1}^{(N-1)/2} P_\ell^{(i)}(\Delta)\log_2 P_\ell^{(i)}(\Delta). \tag{A.2.b}$$

Therefore, taking derivatives with respect to Δ, of (A.2.a) and (A.2.b) yields the following expression for $f_i(\Delta;\lambda)$:

$$f_i(\Delta;\lambda) = -2 \sum_{\ell=1}^{(N-1)/2} Q_\ell^{(i)}(\Delta)\left[2 \frac{\partial S_\ell^{(i)}(\Delta)}{\partial \Delta} - Q_\ell^{(i)}(\Delta) \frac{\partial P_\ell^{(i)}(\Delta)}{\partial \Delta}\right] -$$

$$- \frac{\lambda}{\lambda_i \ell n 2} \left\{ \frac{\partial}{\partial \Delta} P_{(N+1)/2}^{(i)}(\Delta)\left[1+\ell n P_{(N+1)/2}^{(i)}(\Delta)\right] + 2 \sum_{\ell=1}^{(N-1)/2} \frac{\partial}{\partial \Delta} P_\ell^{(i)}(\Delta)\right.$$

$$\left.\left[1+\ell n P_\ell^{(i)}(\Delta)\right]\right\} \quad , \tag{A.3}$$

in which

$$S_\ell^{(i)}(\Delta) \triangleq \int_{T_{\ell-1}^{(i)}}^{T_\ell^{(i)}} x p_i(x) dx \quad , \tag{A.4}$$

and

$$\frac{\partial}{\partial\Delta} P_\ell^{(i)}(\Delta) = \begin{cases} (1-N/2)p_i\lfloor(1-N/2)\Delta\rfloor \quad ; & \ell=1 \\ (\ell-N/2)p_i\lfloor(\ell-N/2)\Delta\rfloor-(\ell-1-N/2)p_i\lfloor(\ell-1-N/2)\Delta\rfloor; & \ell=2,\ldots,(N-1)/2, \end{cases} \tag{A.5}$$

and,

$$\frac{\partial}{\partial\Delta} S_\ell^{(i)}(\Delta) = \begin{cases} (1-N/2)^2\Delta p_i\lfloor(1-N/2)\Delta\rfloor \quad , & \ell=1 \\ (\ell-N/2)^2\Delta p_i\lfloor(\ell-N/2)\Delta\rfloor-(\ell-1-N/2)^2\Delta p_i\lfloor(\ell-1-N/2)\rfloor \quad ; & \ell=2,3,\ldots,(N-1)/2. \end{cases} \tag{A.6}$$

References

1. J. J. Y. Huang and P. M. Schultheiss, "Block Quantization of Correlated Gaussian Random Variables", IEEE Trans. Commun. Syst., Vol. CS-11, pp. 289-296, Sept. 1963.

2. A. Segall, "Bit Allocation and Encoding for Vector Sources", IEEE Trans. Inform. Theory, Vol. IT-22, pp. 162-169, March 1976.

3. N. S. Jayant and P. Noll, Digital Coding of Waveforms: Principles and Applications to Speech and Video, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1984.

4. S. P. Lloyd, "Least Squares Quantization in PCM", IEEE Trans. Inform. Theory, Vol. IT-28, pp. 129-137, March 1982.

5. J. Max, "Quantization for Minimum Distortion", IEEE Trans. Inform. Theory, Vol. IT-6, pp. 7-12, March 1960.

6. N. Farvardin and J. W. Modestino, "Optimum Quantizer Performance for a Class of Non-Gaussian Memoryless Sources", IEEE Trans. Inform. Theory, Vol. IT-30, pp. 485-497, May 1984.

7. A. Habibi and P. A. Wintz, "Image Coding by Linear Transformation and Block Quantization," IEEE Trans. Commun. Technol., Vol. COM-19, pp. 50-62, Feb. 1971.

8. R. Zelinski and P. Noll, "Adaptive Transform Coding of Speech Signals", IEEE Trans. Acoust., Speech, Signal Processing, Vol. ASSP-25, pp. 299-309, Aug. 1977.

9. D. S. Arnstein, "Quantization Erros in Predictive Coders", IEEE Trans. Commun., Vol. COM-23, pp. 423-429, April 1975.

10. N. Farvardin and J. W. Modestino, "Rate-Distortion Performance of DPCM Schemes", IEEE Trans. Inform. Theory, Vol. IT-31, pp. 402-418, May 1985.

11. R. M. Gray and Y. Linde, "Vector Quantizers and Predictive Quantizers for Gauss-Markov Sources", IEEE Trans. Commun., Vol. COM-30, pp. 381-389, Feb. 1982.

12. M. S. Bazarra and C. M. Shetty, Nonlinear Programming, Wiley, New York, 1979.

13. H. Gish and J. N. Pierce, "Asymptotically Efficient Quantizing", IEEE Trans. Inform. Theory, Vol. IT-14, pp. 676-683, Sept. 1968.

14. T. Berger, Rate-Distortion Theory: A Mathematical Basis for Data Compression, Prentice-Hall, Englewood Cliffs, New Jersey, 1971.

15. R. G. Gallager, Information Theory and Reliable Communication, Wiley, New York, 1968.

16. A. J. Viterbi and J. K. Omura, Principles of Digital Communication and Coding, McGraw-Hill, New York, 1979.

17. J. W. Modestino, "Nonparametric and Adaptive Detection of Dependent Data", Ph.D. Thesis, Electrical Engineering Department, Princeton University,

Princeton, New Jersey, April 1969.

18. M. D. Paez and T. H. Glisson, "Minimum Mean-Squared-Error Quantization in Speech PCM and DPCM Systems", IEEE Trans. Commun., Vol. COM-20, pp. 225-230, April 1982.

19. H. P. Kramer and M. V. Mathews, "A Linear Coding for Transmitting a Set of Correlated Signals", IRE Trans. Inform. Theory, Vol. IT-2, pp. 41-46, Sept. 1956.

20. J. Davieau, "Vector Quantizer Design for First-Order Autoregressive Gaussian and Laplacian Sources", Report for ENEE 728B, Digital Waveform Coding, Electrical Engineering Department, University of Maryland, May 1985.

21. R. M. Gray, "Information Rates of Autoregressive Processes", IEEE Trans. Inform. Theory, Vol. IT-16, pp. 412-421, July 1970.

22. M. J. Carter and D. L. Neuhoff, "Bounds to the Rate-Distortion Function of Regenerative Composite Sources", in Proc. 20th Allerton Conf. Commun., Control and Comput., Monticello, IL, Oct. 1982, pp. 680-681.

23. R. E. Blahut, "Computation of Channel Capacity and Rate-Distortion Function", IEEE Trans. Inform. Theory, Vol. IT-18, pp. 460-473, July 1972.

24. N. Farvardin and J. W. Modestino, "On Overflow and Underflow Problems in Buffer-Instrumented Variable-Length Coding of Fixed-Rate Memoryless Sources", to appear in IEEE Trans. Inform. Theory.

25. N. Farvardin and J. W. Modestino, "Adaptive Buffer-Instrumented Variable-Length Coded Quantizer Performance for Memoryless Sources", to appear in IEEE Trans. Inform. Theory.

26. T. Berger, "Minimum Entropy Quantizers and Permutation Codes", IEEE Trans. Inform. Theory, Vol. IT-28, pp. 149-157, March 1982.

R, bits/sample	L=1	2	4	8	DPCM	R(D)
1	4.58 (4.40) [4.40]	5.43 (4.71) [5.18]	5.78 (5.30) [5.69]	5.96 (5.41)	5.22	7.27
2	10.51 (9.30)	11.14 (9.33)	11.45 (10.12)	11.60 (10.24)	11.58	13.29
3	16.53 (14.62)	17.15 (14.62)	17.47 (15.34)	17.62 (15.46)	17.77	19.31

Table 1

SNR (in dB) of Optimum Entropy-Constrained Block Transform Quantization and Comparisons with the Huang-Schultheiss Scheme [1], (numbers in the parenthesis), the Vector Quantization Scheme [11], [20], (numbers in the brackets), the Optimum DPCM [10], and the Rate-Distortion Function for a Fist-Order Gauss-Markov Source with $\rho=0.5$.

R, bits/sample	L=1	2	4	8	DPCM	R(D)
1	4.58 (4.40) [4.40]	7.45 (6.87) [6.85]	8.47 (7.90) [8.18]	9.02 (8.39)	7.56	10.46
2	10.51 (9.50)	12.74 (11.29)	13.86 (12.64)	14.41 (13.20)	14.44	16.48
3	16.53 (14.62)	18.75 (16.93)	19.84 (17.78)	20.31 (18.44)	20.96	22.50

Table 2

SNR (in dB) of Optimum Entropy-Constrained Block Transform Quantization and Comparisons with the Huang-Schultheiss Scheme [1], (numbers in the parenthesis), the Vector Quantization Scheme [11], [20], (numbers in the brackets), the Optimum DPCM [10], and the Rate-Distortion Function for a First-Order Gauss-Markov Source with $\rho=0.8$.

R, bits/sample	L=1	2	4	8	DPCM	R(D)
1	4.58 (4.40) [4.40]	8.71 (7.91) [7.87]	10.60 (10.02) [10.16]	11.53 (10.89)	10.01	13.23
2	10.51 (9.30)	14.19 (12.93)	15.98 (14.90)	16.85 (15.69)	NA	19.25
3	16.53 (14.62)	20.14 (18.27)	21.71 (20.14)	22.18 (20.87)	NA	25.27

Table 3

SNR (in dB) of Optimum Entropy-Constrained Block Transform Quantization and Comparisons with the Huang-Schultheiss Scheme [1], (numbers in the parenthesis), the Vector Quantization Scheme [4], [20]; (numbers in the brackets), the Optimum DPCM [9], and the Rate-Distortion Function for a Fist-Order Gauss-Markov Source with ρ=0.9.

R, bits/sample	L=1	2	4	8	DPCM	$R_L(D)$	$R_U(D)$
1	5.76 (3.01) [3.01]	6.22 (4.31) [4.32]	6.25 (4.68) [5.72]	6.19 (5.10)	7.50	9.81	7.27
2	11.31 (7.53)	11.81 (7.89)	11.79 (9.05)	11.73 (9.56)	NA	17.90	13.29
3	17.20 (12.61)	17.68 (12.92)	17.71 (14.06)	17.70 (14.61)	NA	25.96	19.31

Table 4

SNR (in dB) of Optimum Entropy-Constrained Block Transform Quantization and Comparisons with the Huang-Schultheiss Scheme [1], (numbers in the parenthesis), the Vector Quantization Scheme [11], [20], (numbers in the brackets), the Optimum DPCM [10], and the Rate-Distortion Function for a Fist-Order Laplace-Markov Source with ρ=0.5.

R, bits/sample	L=1	2	4	8	DPCM	$R_L(D)$	$R_U(D)$
1	5.76 (3.01) [3.01]	8.36 (5.97) [5.97]	9.34 (7.09) [7.88]	9.53 (7.74)	11.80	21.69	10.46
2	11.31 (7.53)	15.04 (9.45)	15.44 (11.00)	15.18 (11.86)	NA	38.51	16.48
3	17.20 (12.61)	20.46 (14.45)	20.63 (15.52)	20.40 (16.81)	NA	55.22	22.50

Table 5

SNR (in dB) of Optimum Entropy-Constrained Block Transform Quantization and Comparisons with the Huang-Schultheiss Scheme [1], (numbers in the parenthesis), the Vector Quantization Scheme [11], [20], (numbers in the brackets), the Optimum DPCM [10], and the Rate-Distortion Function for a Fist-Order Laplace-Markov Source with $\rho=0.8$.

R, bits/sample	L=1	2	4	8	$R_L(D)$	$R_U(D)$
1	5.76 (3.01) [3.01]	9.56 (6.68) [6.72]	12.07 (8.82) [9.39]	12.55 (10.04)	39.52	13.23
2	11.31 (7.53)	17.74 (11.91)	18.86 (13.03)	18.39 (14.17)	17.21	19.25
3	17.20 (12.61)	21.60 (15.76)	21.77 (17.36)	21.72 (18.43)	102.90	25.27

Table 6

SNR (in dB) of Optimum Entropy-Constrained Block Transform Quantization and Comparisons with the Huang-Schultheiss Scheme [1], (numbers in the parenthesis), the Vector Quantization Scheme [11], [20], (numbers in the brackets), and the Rate-Distortion Function Bounds for a Fist-Order Laplace-Markov Source with $\rho=0.9$.

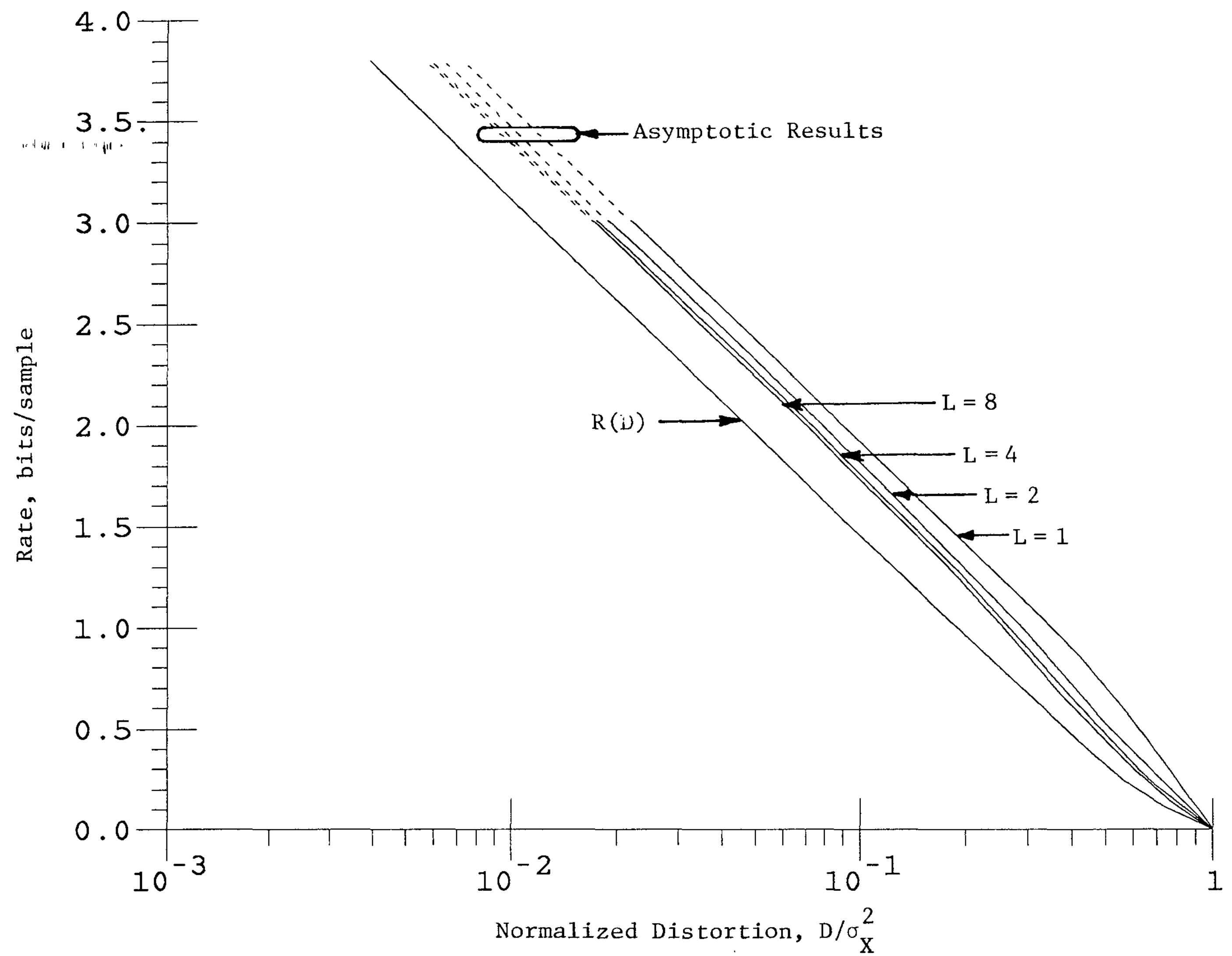

Figure 1 Performance of the Optimum Entropy-Constrained Block Transform Quantization Scheme for a First-Order Gauss-Markov Source with $\rho = 0.5$.

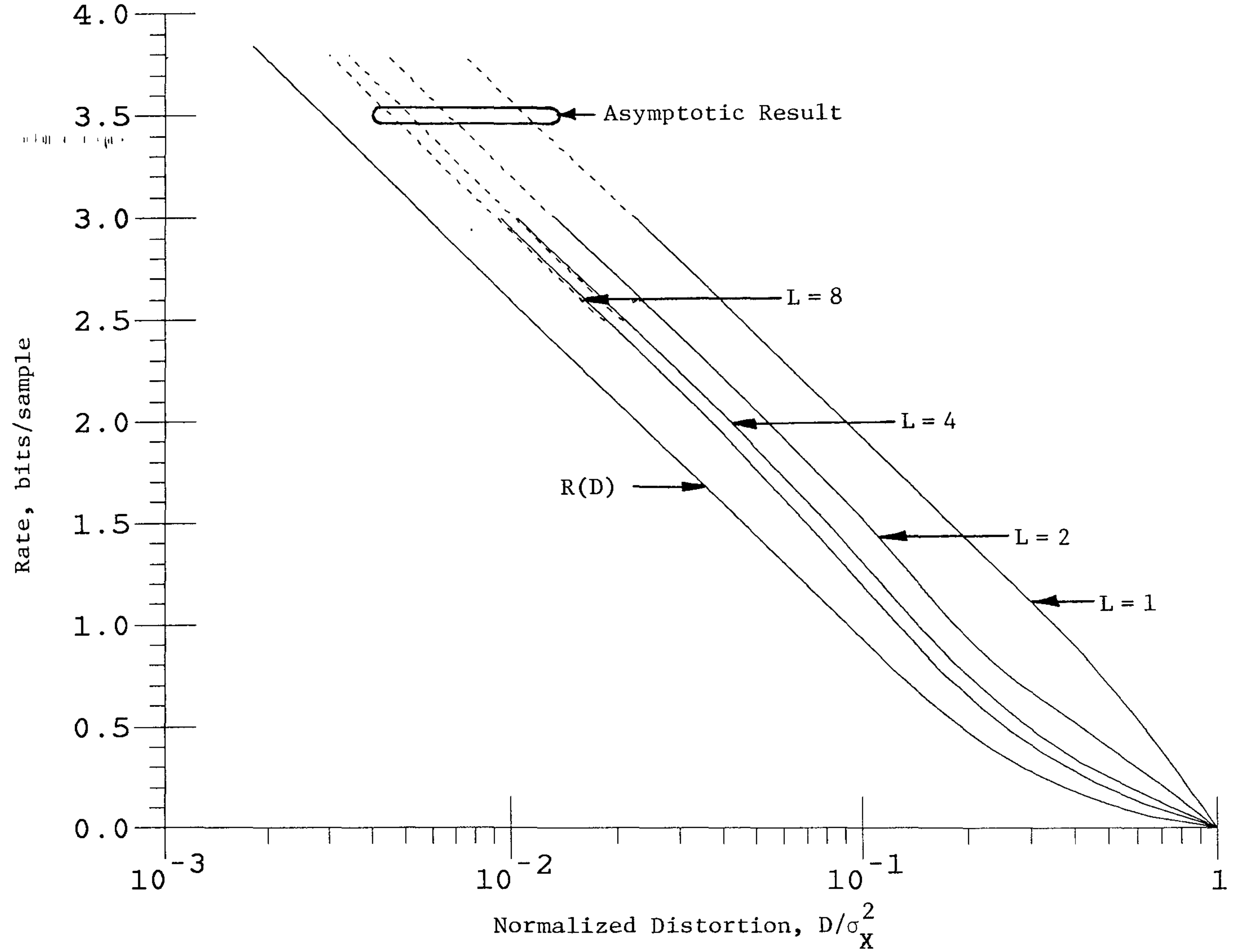

Figure 2 Performance of the Optimum Entropy-Constrained Block Transform Quantization Scheme for a First-Order Gauss-Markov Source with $\rho = 0.8$.

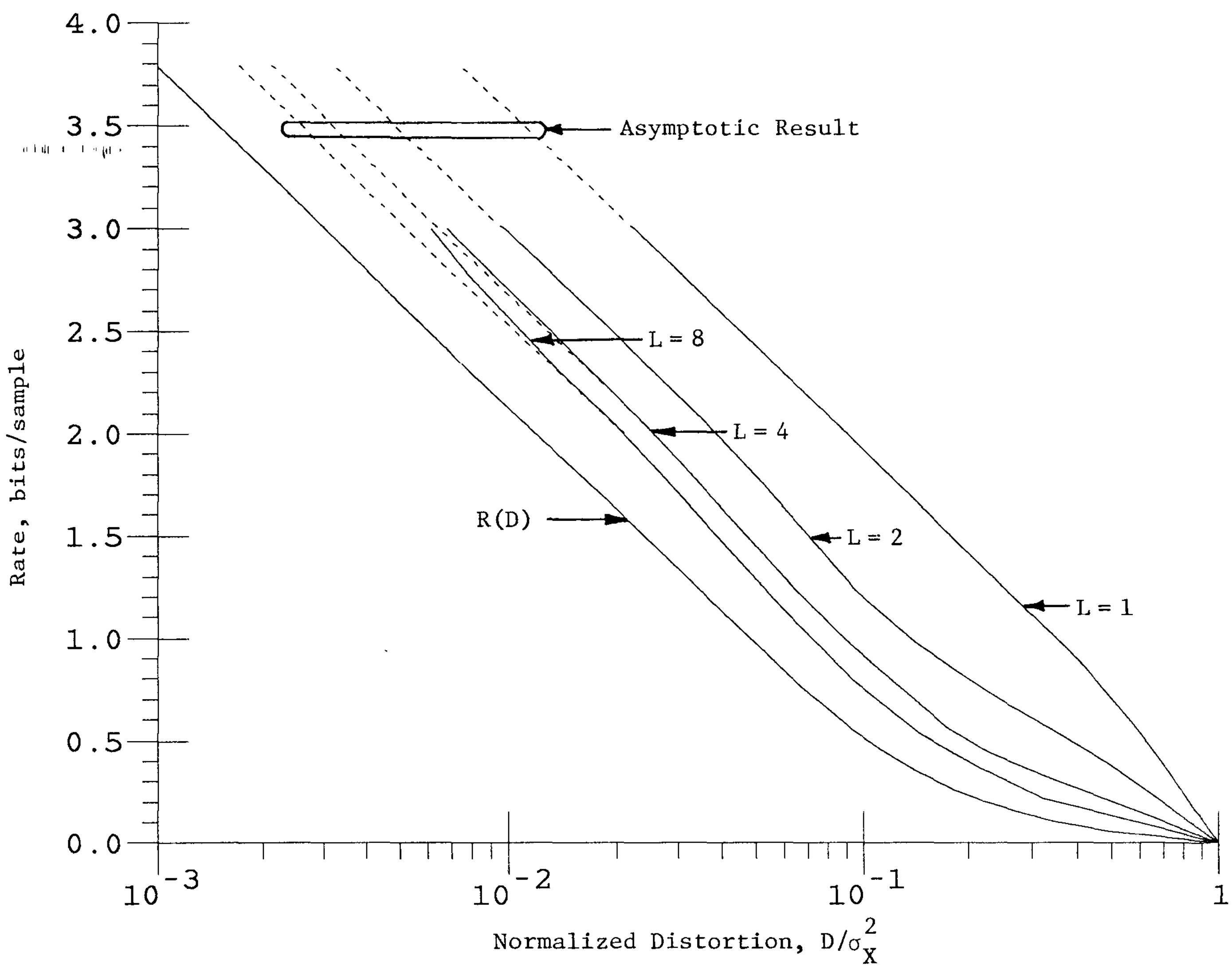

Figure 3 Performance of the Optimum Entropy-Constrained Block Transform Quantization Scheme for a First-Order Gauss-Markov Source with $\rho = 0.9$.