Multidimensional Companding for Non-difference Distortion Measures *

Tamás Linder[1] Ram Zamir[2] Kenneth Zeger[3]

[1,3]Dept. of Electrical and Computer Engineering, University of California, San Diego, La Jolla, CA 92093-0407. {linder,zeger}@code.ucsd.edu

[2] Dept. of Electrical Engineering–Systems Tel Aviv University Tel-Aviv 69978, Israel. zamir@eng.tau.ac.il

Abstract

Entropy-coded vector quantization is studied using high resolution multidimensional companding over a class of non-difference distortion measures. For distortion measures which are "locally quadratic" a rigorous derivation of the asymptotic distortion and entropy coded rate of multidimensional companders is given along with conditions for the optimal choice of the compressor function. Examples are shown for the existence of optimal compressors. The rate distortion performance of the companding scheme is studied using a recently obtained asymptotic expression for the rate distortion function which parallels the Shannon lower bound for difference distortion measures. It is proved that the high resolution performance of the scheme is arbitrarily close to the rate distortion limit for large quantizer dimensions if the compressor function and the lattice quantizer used in the companding scheme are optimal, extending an analogous statement for entropy coded lattice quantization and MSE distortion.

1 Introduction

The high resolution (asymptotic, low distortion) behavior of vector quantizers is relatively well understood for so called difference distortion measures where the distortion is measured by a function of the difference between the source and the reproduction vectors. In particular, for the mean squared error, and more generally for nice functions of a norm-based distance measures, the asymptotic distortion of the optimal quantizer, as well as the asymptotic distortion of sequences of quantizers with a given "point density" have been identified as a function of the codebook size, or as a function of the entropy of the output [1, 2, 3, 4, 5, 6]. These results give insight to the structure of asymptotically optimal quantizers. On the practical side, the expressions for quantizer performance provide useful guidance for quantizer design at even small to moderate rates.

Source coding is less understood when the distortion is not measured by a difference distortion measure. Non-difference distortion measures occur naturally in source coding problems. Prominent examples include the log spectral distortion and the Itakura-Saito distortion, which are used in linear predictive speech coding [7], perceptual distortion measures in image coding, and most distortion measures that arise in noisy (or remote) source coding if the original distortion measure is other than the squared error. Due to the difficulty of such analyses, there exist only a few known results for high resolution quantization with non-difference distortion measures. By assuming the existence of a limiting quantizer point density, a lower bound was calculated in [8] for the high resolution performance of fixed rate optimal vector quantizers for locally quadratic distortion measures. The log spectral distortion and the Itakura-Saito distortion are examples of such measures. A more formal treatment of the same lower bound is given in [9], and a new lower bound on the variable rate performance is developed using optimal point densities. It is also pointed out in [9] that some important "perceptual distortion measures" in image coding are locally quadratic. In [10] an asymptotically tight expression for the rate distortion function is derived for locally quadratic distortion measures. As will be shown in this paper, the expression given in [10] plays the same important role in high resolution quantization for these distortion measures as does the Shannon lower bound in quantizing for squared error loss.

To develop the basics of a high resolution quantization theory for locally quadratic distortion measures, we investigate variable rate (entropy coded) companding vector quantization. Multidimensional companding is a type of structured vector quantization of low complexity where a k-dimensional source vector X is "compressed" by an invertible mapping F (called the compressor function). Then $F(X)$ is quantized by a uniform (or more generally, a lattice) quantizer, and the inverse mapping F^{-1} is

*The research was supported in part by the National Science Foundation.

applied to obtain the reproduction $\hat{X}$. Thus the scheme is

$$X \to F(\cdot) \to Q_U(\cdot) \to F^{-1}(\cdot) \to \hat{X}$$

where Q_U is a uniform or lattice quantizer whose output may or may not be entropy coded.

We have two main reasons for considering multidimensional companders. First, in [10] it has been informally observed that for a large class of non-difference distortion measures the asymptotically optimal forward test channel which realizes Shannon's rate distortion function has a certain structure very similar to that of multidimensional companding quantizers. It has also been conjectured that such a companding scheme, together with entropy coding, performs arbitrarily close to the rate distortion limit. Second, our aim is to develop a rigorous theory with performance bounds relating to the rate distortion function. In high resolution quantization theory a commonly invoked technique to obtain results is to use heuristic, informal reasoning (see, e.g., [11, 2, 3, 4, 8]). While there are a number of important results derived rigorously, it is apparent that the majority of them deal with fixed rate quantization [5, 12, 13, 6, 14]. The reason is that it is especially difficult to construct rigorous proofs using the notion of "point density" for unstructured entropy coded vector quantization. On the other hand, for the purpose of directly comparing quantizer performance to the rate distortion function, one is forced to consider the asymptotic variable rate performance and thus to introduce structure in the coding scheme (see e.g. [15, 16]).

In this paper we consider entropy coded multidimensional companding quantizers with non-difference distortion measures satisfying rather general regularity conditions. The main requirement is a smoothness condition which implies that the distortion $d(x,y)$ between $x, y \in \mathbb{R}^k$ can be approximated as $d(x,y) \approx (x-y)^T M(x)(x-y)$ for y close to x, where $M(x)$ is an input-dependent positive definite matrix. Theorem 1 gives a rigorous derivation of the asymptotic entropy coded rate as a function of the distortion for sources with densities. A general sufficient condition for the optimal choice of the compressor function is derived in Theorem 2 and examples are shown for the existence of optimal compressors, which are determined by the distortion measure and do not depend on the source distribution. Using a result from [10] we prove in Theorem 4 that if the compressor function satisfies the sufficient condition for optimality, and if the lattice quantizer used in the companding scheme is optimal, then the high resolution performance of the scheme is arbitrarily close to the rate distortion limit for large quantizer dimensions. When specialized to mean squared error, this result gives back the well known fact that for large rate and large quantizer dimension, lattice quantizers combined with entropy coding are asymptotically optimal.

2 Preliminaries

A k-dimensional *vector quantizer* Q is a mapping defined by

$$Q(x) = y_i \quad \text{if} \quad x \in B_i,$$

where $B_1, \ldots, B_n$ form a measurable partition of $\mathbb{R}^k$, and the collection of *codepoints* $y_i \in \mathbb{R}^k$, $1 \le i \le n$ is called the *codebook*. We do not eliminate the possibility that $n = \infty$, i.e., the codebook of Q can contain countably infinite number of codepoints. The distortion between x and $Q(x)$ is measured by $d(x, Q(x))$, where $d : \mathbb{R}^k \times \mathbb{R}^k \to [0, \infty)$ is a Borel measurable function. The expected distortion in quantizing a k-dimensional random vector X is

$$D(Q) = \mathbf{E}[d(X, Q(X))],$$

and we assume the expectation is finite. The *rate* of Q will be measured by the Shannon entropy of $Q(X)$,

$$H(Q) = -\sum_i \mathbf{P}\{Q(X) = y_i\} \log \mathbf{P}\{Q(X) = y_i\},$$

where the logarithm is base two. The per dimension rate of the system can by made within $\frac{1}{k}$ of $\frac{1}{k}H(Q)$ by use of entropy coding techniques.

The basic building block in a multidimensional companding quantizer is a lattice quantizer. Let Λ be a k-dimensional nonsingular k-dimensional lattice, and for any $\alpha > 0$ let $\alpha\Lambda$ denote the scaled lattice $\alpha\Lambda = \{\alpha z : z \in \Lambda\}$. The lattice quantizer $Q_{\alpha\Lambda}$ is then defined so that its codepoints are the points of $\alpha\Lambda$ and its quantization regions are the corresponding Voronoi regions of $\alpha\Lambda$, i.e., if $Q_{\alpha\Lambda}(x) = z$, then $\|x - z\| \le \|x - z'\|$ for all $z' \in \alpha\Lambda$, where $\|\cdot\|$ denotes the Euclidean norm. The quantization regions of $\alpha\Lambda$ are the translated and scaled copies of P_0, the basic Voronoi cell of Λ, which is defined by

$$P_0 = \{x \in \mathbb{R}^k : \|x\| \le \|x - z\| \text{ for all } z \in \Lambda\}.$$

An important performance figure of Λ is the (dimensionless) normalized second moment of its basic cell, namely

$$L(P_0) = \frac{\int_{P_0} \|x\|^2 \, dx}{kV(P_0)^{2/k+1}},$$

where $V(P_0)$ denotes the k-dimensional volume of P_0. We call a lattice *optimal* if its normalized second moment is minimum over all k-dimensional lattices [17]. It was proved in [18] that the basic cell of an optimal lattice is *white* in the sense that if $Z = (Z_1, \ldots, Z_k)$ is a random vector uniformly distributed over P_0, then the covariance matrix of Z is

$$\mathbf{E}[ZZ^T] = \sigma^2 I,$$

where I denotes the $k \times k$ identity matrix. In other words, the Z_i are uncorrelated and their second moments are equal. We will assume that the lattice Λ used in the companding scheme has a white basic cell P_0.

The concept of a companding realization of a nonuniform quantizer originates from Bennett [11]. The idea is to apply a nonlinear transformation (called the compressor) to the input, followed by a uniform (more generally a lattice) quantizer and the inverse of the transformation to obtain the reproduction. Let $F : \mathbb{R}^k \to \mathbb{R}^k$ be a one-to-one continuously differentiable mapping whose derivative matrix $F'(x)$ is nonsingular for all x. Then F has an inverse $F^{-1} = G$ which is continuously differentiable on its domain and whose derivative G' is nonsingular. F and G are called the *compressor* and *expander* functions, respectively.

The *companding vector quantizer* realized by the compressor function F and the scaled lattice quantizer $Q_{\alpha\Lambda}$ is defined by

$$Q_{\alpha,F}(x) = G(Q_{\alpha\Lambda}(F(x))), \quad x \in \mathbb{R}^k.$$

Our goal is to analyze the entropy coded rate of $Q_{\alpha,F}$ as a function of its distortion for absolutely continuous source distributions. In general, the analytical evaluation of the rate is not possible for any given $\alpha > 0$, so we are forced to take the asymptotic approach and determine the asymptotic behavior of the rate as the distortion (or equivalently α) tends to zero.

3 Multidimensional Companding

3.1 Asymptotic Performance

Let $x = (x_1, \ldots, x_k) \in \mathbb{R}^k$, $y = (y_1, \ldots, y_k) \in \mathbb{R}^k$ (x and y are regarded as column vectors) and assume that the distortion measure $d(x,y)$ satisfies the following three conditions.

(a) For all fixed $x \in \mathbb{R}^k$, $d(x,y)$ is three times continuously differentiable in the variable y, and the third order partial derivatives

$$\frac{\partial^3 d(x,y)}{\partial y_i \partial y_j \partial y_n}, \quad i,j,n \in \{1, \ldots, k\} \tag{1}$$

are uniformly bounded.

(b) For all $x \in \mathbb{R}^k$, $d(x,y) \geq 0$ with equality if and only if $y = x$.

Condition (b) implies that the gradient of $d(x,y)$ with respect to y is zero at $y = x$. Thus for any fixed x, a second order Taylor expansion of $d(x,y)$ in y gives

$$d(x,y) = (x-y)^T M(x)(x-y) + O(\|x-y\|^3), \tag{2}$$

where $M(x)$ is the $k \times k$ matrix whose ijth element is given by

$$m_{ij}(x) = \frac{1}{2}\frac{\partial^2 d(x,y)}{\partial y_i \partial y_j}\bigg|_{y=x}. \tag{3}$$

$M(x)$ is called the *sensitivity matrix* of d [8]. Since $d(x,y) > 0$ if $y \neq x$, this quadratic approximation implies that $M(x)$ is symmetric and nonnegative definite. In addition to (a) and (b) above, we impose the following condition on the sensitivity matrix.

(c) $M(x)$ is positive definite for all x and its elements $m_{ij}(x)$ are continuous functions.

Remark. (i) Consider an *input weighted quadratic* distortion measure given by

$$d(x,y) = \|W(x)(x-y)\|^2,$$

where $W(x)$ is a nonsingular $k \times k$ matrix depending on the input x [19]. Since $d(x,y) = (x-y)^T W^T(x) W(x)(x-y)$, and since $M(x) = W(x)^T W(x)$ is positive definite, it is easy to see that $d(x,y)$ satisfies condition (a)–(c) if the elements of $W(x)$ are continuous functions of x. (ii) Very similar conditions are used in [9] to heuristically derive lower bounds on the asymptotic distortion of a sequence of fixed rate quantizers with a given point density. Some important measures of image quality [20, 21] satisfy these regularity conditions, for example.

To study the rate of $Q_{\alpha,F}$ as a function of its distortion, one needs to eliminate the scaling factor α. One reasonable way to do this is to chose an $\alpha(D) > 0$, for each $D > 0$, such that

$$D(Q_{\alpha(D),F}) = D.$$

If X has a density, it is not hard to see that $D(Q_{\alpha,F})$ is a continuous function of $\alpha > 0$ which converges to zero as $\alpha \to 0$, so that such $\alpha(D)$ always exists for all sufficiently small $D > 0$. For such values of D we define

$$Q_{D,F} = Q_{\alpha(D),F}.$$

The next theorem, the main result of this paper, determines the asymptotic behavior of the rate of $Q_{D,F}$ as $D \to 0$ for bounded sources.

Theorem 1 *Assume that the source X has a density which is zero outside a bounded subset of $\mathbb{R}^k$ and suppose the distortion function $d(x,y)$ satisfies conditions (a)–(c). If X has a finite differential entropy $h(X)$, then the rate $H(Q_{D,F})$ and the distortion D of the multidimensional companding quantizer $Q_{D,F}$ satisfy*

$$\lim_{D\to 0}\left(H(Q_{D,F}) + \frac{k}{2}\log D\right) = h(X) + \mathbf{E}[\log|\det F'(X)|] + \frac{k}{2}\log\left(L(P_0)\mathbf{E}[\mathrm{tr}\{\Gamma(X)\}]\right)$$

where F is the compressor function, $\mathrm{tr}\{\Gamma(x)\}$ *denotes the trace of the matrix*

$$\Gamma(x) = [F'(x)]^{-T} M(x) [F'(x)]^{-1} \tag{4}$$

and $[F'(x)]^{-T}$ is the inverse transpose of the derivative of $F(x)$.

Remarks. (i) The proof of this result is given in [22]. The only restrictive condition in Theorem 1 is the assumption that the source density has bounded support.

In principle, the result can be proved for source densities with unbounded support, but in that case extra conditions on the compressor function are needed. These conditions are associated with the tail of the source density, leading to a substantially more complicated proof. (ii) The proof of the result can be used to obtain the high-rate distortion of the companding quantizer as a function of the number of codepoints (which is finite since the source is bounded). When specialized to mean squared error ($M(x) = I$), we obtain Bucklew's heuristically derived formula [23] for fixed rate multidimensional companding.

3.2 Optimal Compressor Functions

The question of the optimal choice of the compressor F is considered next. Let us define

$$C_1(F) = \mathbf{E}[\mathrm{tr}\{\Gamma(X)\}]$$

and

$$C_2(F) = \mathbf{E}[\log|\det F'(X)|],$$

where $\Gamma(x)$ is defined in (4). Then the statement of Theorem 1 becomes

$$\lim_{D\to 0}\Big(H(Q_{D,F}) + \frac{k}{2}\log D\Big) = h(X) + \frac{k}{2}\log L(P_0) + C_2(F) + \frac{k}{2}\log C_1(F).$$

It is clear that if F satisfies

$$C_2(F) + \frac{k}{2}\log C_1(F) \le C_2(\hat{F}) + \frac{k}{2}\log C_1(\hat{F})$$

for all allowable $\hat{F}$, then $Q_{D,F}$ asymptotically outperforms all other companding quantizers $Q_{D,\hat{F}}$. Thus to find a best compander one has to minimize the functional

$$C_2(F) + \frac{k}{2}\log C_1(F) \qquad (5)$$

over all one-to-one and continuously differentiable F such that $F'(x)$ is nonsingular for all x.

A lower bound on (5) is obtained next. Since $M(x)$ is positive definite and $F'(x)$ is nonsingular, $\Gamma(x) = [F'(x)]^{-T}M(x)[F'(x)]^{-1}$ is also positive definite. Thus by the arithmetic-geometric mean inequality we have $\mathrm{tr}\{\Gamma(x)\} \ge k(\det\Gamma(x))^{1/k}$ with equality if and only if the eigenvalues of $\Gamma(x)$ are all equal. Therefore

$$\begin{aligned}\frac{k}{2}\log\mathbf{E}[\mathrm{tr}\{\Gamma(X)\}] &\ge \frac{k}{2}\log\mathbf{E}\left[k(\det\Gamma(X))^{1/k}\right] &(6)\\ &\ge \frac{k}{2}\mathbf{E}\left[\log(k(\det\Gamma(X))^{1/k})\right] &(7)\\ &= \frac{k}{2}\log k + \frac{1}{2}\mathbf{E}[\log(\det M(X))]\\ &\quad -\mathbf{E}[\log|\det F'(X)|],\end{aligned}$$

where (7) follows from Jensen's inequality. The above is equivalent to

$$\frac{k}{2}\log C_1(F) + C_2(F) \ge \frac{k}{2}\log k + \frac{1}{2}\mathbf{E}[\log(\det M(X))]. \qquad (8)$$

Let us examine the conditions for achieving the above lower bound. We have equality in (8) iff both (6) and (7) are equalities. Equality holds in (6) iff the eigenvalues of $\Gamma(x)$ are equal a.e. $[\mu_X]$. (μ_X denotes the probability measure induced by X on $\mathbb{R}^k$.) Since $\Gamma(x)$ is positive definite, this implies that $\Phi(x)^T\Gamma(x)\Phi(x) = \beta(x)I$ a.e. $[\mu_X]$ for some $\beta(x) > 0$ and some orthogonal matrix $\Phi(x)$ (i.e., $\Phi(x)^T\Phi(x) = I$). This in turn implies that $\Gamma(x) = \beta(x)I$ a.e. $[\mu_X]$. The condition of equality in (7) is that the determinant of $\Gamma(x)$ be constant a.e. $[\mu_X]$. Thus equality holds in (8) if and only if $\Gamma(x) = \beta I$ a.e. $[\mu_X]$, where $\beta > 0$ is a constant. This is equivalent to $M(x) = \beta F'(x)^T F'(x)$ a.e. $[\mu_X]$. We have thus proved the following sufficient condition for the optimality of a compressor function in terms of a condition involving the sensitivity matrix of d.

Theorem 2 *Assume the compressor F satisfies*

$$F'(x)^T F'(x) = cM(x) \qquad (9)$$

a.e. $[\mu_X]$, where $c > 0$ is a scalar constant. If the conditions of Theorem 1 hold, then

$$\lim_{D\to 0}\left(H(Q_{D,F}) + \frac{k}{2}\log D\right) = h(X) + \frac{k}{2}\log(kL(P_0)) + \frac{1}{2}\mathbf{E}[\log(\det M(X))]$$

and therefore F is an optimal compressor function in the sense that

$$\lim_{D\to 0}\left(H(Q_{D,\hat{F}}) - H(Q_{D,F})\right) \ge 0$$

for all other compressors $\hat{F}$.

Remarks. (i) The proof of Theorem 1 indicates that the sufficient condition of optimality in the above theorem means the following. The optimal compressor function shapes the lattice quantizer so that for small α the weighted quantization error vector $e = W(X)(X - Q_{\alpha,F}(X))$ is approximately white and its conditional power $\mathbf{E}[\|e\|^2|Q_{\alpha,F}(X) = y_{\alpha,i}]$ does not depend on the codepoint $y_{\alpha,i}$. (ii) Note that the optimality condition of Theorem 2 does not depend on the source density. This observation nicely parallels the fact that for mean squared error the asymptotically optimal entropy coded quantizer is an infinite level uniform quantizer [2]. It is also analogous to a widely cited conjecture made by Gersho [3] that the asymptotically optimal entropy coded quantizer has a so called tessellating structure, i.e., its quantization regions are congruent polytopes which tessellate the whole space.

The condition $F'(x)^T F'(x) = M(x)$ is a system of partial differential equations which might not have a solution for a general $M(x)$. Thus, as in the case of fixed rate multidimensional companding for the squared error [3, 23, 24], in general there may not exist a compressor function $F(x)$ satisfying the above condition. The following example shows that the condition of Theorem 2 can be satisfied if d is a single letter distortion measure.

Example Assume that $d(x, y)$ can be written as

$$d(x, y) = \sum_{i=1}^{k} d_i(x_i, y_i), \tag{10}$$

where $x = (x_1, \dots, x_k)$, $y = (y_1, \dots, y_k)$, and the scalar distortion measures $d_i : \mathbb{R} \times \mathbb{R} \to \mathbb{R}^+$, $1 \le i \le k$ satisfy conditions (a)–(c). Namely, for each d_i we must have that $\frac{\partial^3}{\partial u^3} d_i(t, u)$ is uniformly bounded, $d_i(t, u) \ge 0$ with equality iff $u = t$, and

$$m_i(t) = \frac{1}{2} \frac{\partial^2 d_i(t, u)}{\partial u^2}\Big|_{u=t}$$

is positive and continuous for all t. Then $M(x)$ is the diagonal matrix

$$M(x) = \operatorname{diag}\{m_1(x_1), \dots, m_k(x_k)\}.$$

Define $F(x) = (F_1(x), \dots, F_k(x))$ by setting

$$F_i(x) = \int_0^{x_i} m_i(t)^{1/2}\, dt,$$

where we used the convention that $\int_a^b = -\int_b^a$ if $a > b$. Then $F(x)$ is one-to-one and continuously differentiable since each $m_i(t)$ is positive and continuous by assumption. Obviously, $F'(x)$ is diagonal and $[F'(x)]^2 = M(x)$ so the optimality condition of Theorem 2 is satisfied. There is an interesting analogy with fixed rate multidimensional companding for squared error. For this problem it has recently been reported [25] that if the source is stationary and memoryless, then the optimal compressor function compresses each vector component independently, using the scalar compressor optimal for the marginal distribution of the source.

Consider now the more general case when $d(x, y)$ is not a single letter distortion measure, but there exists an orthogonal transformation such that $d(x, y)$ becomes a single letter distortion measure in the transformed space. That is, if

$$d(x, y) = d^*(Vx, Vy)$$

where V is a fixed orthogonal matrix and d^* is in the form of (10), then it is easy to see that the optimal compander F is given by $F(x) = F^*(Vx)$, where F^* is optimal for d^*.

4 Rate Distortion Performance

The rate distortion function of the random vector X is defined by

$$R(D) = \inf\left\{I(X, Y) : \mathbf{E}[d(X, Y)] \le D\right\}, \tag{11}$$

where $I(X; Y)$ denotes the mutual information between the the k-dimensional random vectors X and Y, and the infimum is taken over all joint distributions of the pair (X, Y) such that $\mathbf{E}[d(X, Y)] \le D$. The rate distortion function characterizes the lowest rate achievable by any source coding scheme in coding a memoryless vector source with marginal X at distortion level D. In particular, $R(D)$ is a lower bound on the rate of any vector quantizer for X whose distortion does not exceed D.

In recent work [10] the asymptotic behavior of the rate distortion function was investigated for locally quadratic non-difference distortion measures. Let us assume that in addition to conditions (a)–(c), the distortion measure also satisfies the following natural requirements.

(d) The elements of $M(x)$ are continuously differentiable functions of x.

(e) $\liminf_{\|y\| \to \infty} d(x, y) > 0$ for all $x \in \mathbb{R}^k$.

The next theorem is a specialization of a more general result to our case.

Theorem 3 ([10, Theorem 1]) *Assuming that (d), (e), and the conditions of Theorem 1 hold, the low distortion asymptotic behavior of $R(D)$ is given by*

$$\lim_{D \to 0} \left(R(D) + \frac{k}{2} \log(2\pi e D/k) \right) = h(X) + \frac{1}{2} \mathbf{E}[\log(\det M(X))],$$

where $M(x)$ is the sensitivity matrix of $d(x, y)$ given in (3).

Assume now that there exists an optimal compressor F which satisfies the optimality condition (9) of Theorem 2. Then Theorem 3 implies the following.

Theorem 4 *Assume that an optimal compressor F exists which satisfies $F'(x)^T F'(x) = cM(x)$ a.e. $[\mu_X]$, where $c > 0$ is a scalar constant and $M(x)$ is the sensitivity matrix of $d(x, y)$. Then with the conditions of Theorem 1 and (d) and (e), the low distortion asymptotic behavior of the multidimensional companding quantizer relative to $R(D)$ is given by*

$$\lim_{D \to 0} \left(H(Q_{D,F}) - R(D) \right) = \frac{k}{2} \log(2\pi e L(P_0)). \tag{12}$$

Thus, for low distortion, the per dimension rate of $Q_{F,D}$ is at most $\frac{1}{2}\log(2\pi e L(P_0))$ bits above the rate distortion function.

Remark. This statement has a well known analogue for mean squared error and entropy coded lattice (or tesselating) quantizers [2, 4]. In fact, the same upper bound applies there, but the the result is conceptually much simpler since the well known Shannon lower bound for the squared error can be used in place of Theorem 3.

Let G_k be the minimum value of $L(P_0)$ for any k-dimensional lattice. Based on a result of Poltyrev it was proved in [18, Lemma 1] that as $k \to \infty$, $G_k \to (2\pi e)^{-1}$ at a rate $\log(2\pi e G_k) = O(k^{-1} \log k)$. Thus, for optimal lattices and compressors,

$$\lim_{D \to 0} \frac{1}{k}(H(Q_{D,F}) - R(D)) = O\left(\frac{\log k}{k}\right)$$

which indicates that for high dimension and low distortion, an entropy constrained companding vector quantizer with optimal compressor function can arbitrarily approach the rate distortion performance limit.

References

[1] P. Zador, "Asymptotic quantization error of continuous signals and the quantization dimension," *IEEE Trans. Inform. Theory*, vol. IT-28, pp. 139–149, Mar. 1982.

[2] H. Gish and J. N. Pierce, "Asymptotically efficient quantizing," *IEEE Trans. Inform. Theory*, vol. IT-14, pp. 676–683, Sep. 1968.

[3] A. Gersho, "Asymptotically optimal block quantization," *IEEE Trans. Inform. Theory*, vol. IT-25, pp. 373–380, July 1979.

[4] Y. Yamada, S. Tazaki, and R. M. Gray, "Asymptotic performance of block quantizers with difference distortion measures," *IEEE Trans. Inform. Theory*, vol. IT-26, pp. 6–14, Jan. 1980.

[5] J. A. Bucklew and G. L. Wise, "Multidimensional asymptotic quantization theory with rth power distortion measures," *IEEE Trans. Inform. Theory*, vol. IT-28, pp. 239–247, Mar. 1982.

[6] S. Na and D. L. Neuhoff, "Bennett's integral for vector quantizers," *IEEE Trans. Inform. Theory*, vol. IT-41, pp. 886–900, Jul. 1995.

[7] A. Buzo, A. H. Gray, R. M. Gray, and J. D. Markel, "Speech coding based upon vector quantization," *IEEE Trans. Acoust. Speech, Signal Processing*, vol. ASSP-28, pp. 562–574, Oct. 1980.

[8] W. R. Gardner and B. D. Rao, "Theoretical analysis of the high-rate vector quantization of LPC parameters," *IEEE Trans. Speech, Audio Processing*, vol. 3, pp. 367–381, Sep. 1995.

[9] J. Li, N. Chadda, and R. M. Gray, "Asymptotic performance of vector quantizers with the perceptual distortion measure." preprint, to be presented at the *1997 IEEE Int. Symp. Inform. Theory*, 1996.

[10] T. Linder and R. Zamir, "High-resolution source coding for non-difference distortion measures: the rate distortion function." submitted to *IEEE Trans. Inform. Theory*, 1997.

[11] W. R. Bennett, "Spectrum of quantized signals," *Bell. Syst. Tech. J.*, vol. 27, pp. 446–472, 1948.

[12] J. A. Bucklew, "Two results on the asymptotic performance of quantizers," *IEEE Trans. Inform. Theory*, vol. IT-30, pp. 341–348, Mar. 1984.

[13] S. Cambanis and N. L. Gerr, "A simple class of asymptotically optimal quantizers," *IEEE Trans. Inform. Theory*, vol. IT-29, pp. 664–676, Sep. 1983.

[14] D. L. Neuhoff, "On the asymptotic distribution of errors in vector quantization," *IEEE Trans. Inform. Theory*, vol. IT-42, pp. 461–468, Mar. 1996.

[15] J. Ziv, "On universal quantization," *IEEE Trans. Inform. Theory*, vol. IT-31, pp. 344–347, May 1985.

[16] R. Zamir and M. Feder, "On universal quantization by randomized uniform/lattice quantizers," *IEEE Trans. Inform. Theory*, vol. IT-38, pp. 428–436, Mar. 1992.

[17] J. H. Conway and N. J. A. Sloane, *Sphere packings, Lattices and Groups.* New York: Springer-Verlag, 1988.

[18] R. Zamir and M. Feder, "On lattice quantization noise," *IEEE Trans. Inform. Theory*, vol. IT-42, pp. 1152–1159, July 1996.

[19] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression.* Boston: Kluwer, 1992.

[20] A. M. Eskicioglu and P. S. Fisher, "Image quality measures and their performance," *IEEE Trans. Communications*, vol. COM-43, pp. 2959–2965, Dec. 1995.

[21] N. B. Nill, "A visual model weighted cosine transform for image compression," *IEEE Trans. Communications*, vol. COM-33, pp. 551–557, Jun. 1985.

[22] T. Linder, R. Zamir, and K. Zeger, "High-resolution source coding for non-difference distortion measures: multidimensional companding." submitted to *IEEE Trans. Inform. Theory*, 1997.

[23] J. A. Bucklew, "Companding and random quantization in several dimensions," *IEEE Trans. Inform. Theory*, vol. IT-27, pp. 207–211, Mar. 1981.

[24] J. A. Bucklew, "A note on optimal multidimensional companders," *IEEE Trans. Inform. Theory*, vol. IT-29, p. 279, Mar. 1983.

[25] P. W. Moo and D. L. Neuhoff, "Optimal compressor functions for multidimensional companding." preprint, to be presented at the *1997 IEEE Int. Symp. Inform. Theory*, 1996.