ACKNOWLEDGMENT

The author is very grateful to Dr. H. Yamamoto, Dr. S. Hirasawa, and Dr. T. Niinomi for the inspiring discussion on the ARQ schemes and to Dr. V. B. Balakirsky for pointing my attention to the work by Dr. Kudryashov. Moreover, he is also grateful to anonymous referees for valuable comments.

REFERENCES

[1] H. Yamamoto and K. Itoh, "Viterbi decoding algorithm for convolutional codes with repeat request," *IEEE Trans. Inform. Theory*, vol. IT-26, pp. 540–547, Sept. 1980.
[2] R. J. F. Fang, "Lower bounds on reliability functions of variable-length nonsystematic," *IEEE Trans. Inform. Theory*, vol. IT-17, pp. 161–171, Mar. 1971.
[3] S. Lin and D. J. Costello, Jr., *Error Control Coding: Fundamentals and Application.* Englewood Cliffs, NJ: Prentice-Hall, 1983.
[4] A. Drukarev and D. J. Costello, Jr., "Hybrid ARQ error control using sequential decoding," *IEEE Trans. Inform. Theory*, vol. IT-29, pp. 521–535, July, 1983.
[5] B. D. Kudryashov, "Viterbi algorithm for convolutional code decoding in a system with decision feedback," *Probl. Peredach. Inform.*, vol. 20, pp. 18–26, no. 2, 1984.
[6] G. D. Forney, Jr., "Exponential error bounds for erasure, list, and decision feedback scheme," *IEEE Trans. Inform. Theory*, vol. IT-14, pp. 206–220, Mar. 1968.
[7] H.-A. Loeliger, "A practical reliability metric for block codes used on binary-input channels," *IEEE Trans. Commun.*, vol. 38, pp. 405–408, Apr. 1990.
[8] G. D. Forney, Jr., "Convolutional codes II. Maximum-likelihood decoding," *Inform. Contr.*, vol. 25, pp. 223–265, 1974.
[9] V. B. Balakirsky, a series of discussions, 1990.
[10] R. G. Gallager, *Information Theory and Reliable Communication.* New York: Wiley, 1968.
[11] G. D. Forney, Jr., *Concatenated Codes.* Cambridge, MA: M.I.T. Press, 1966.
[12] A. J. Viterbi and J. K. Omura, *Principles of Digital Communication and Coding.* New York: McGraw-Hill, 1979.
[13] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems.* New York: Academic, 1981.
[14] T. Hashimoto, "A coded ARQ scheme with the generalized Viterbi algorithm," *IEEE Trans. Inform. Theory*, to appear.
[15] F. Jelinek, *Probabilistic Information Theory: Discrete and Memoryless Models.* McGraw-Hill, 1968.

Asymptotic Entropy-Constrained Performance of Tessellating and Universal Randomized Lattice Quantization

Tamás Linder and Kenneth Zeger

***Abstract*—Two results are given. First, using a result of Csiszár, the asymptotic (i.e., high-resolution/low distortion) performance for entropy-constrained tessellating vector quantization, heuristically derived by Gersho, is proven for all sources with finite differential entropy. This implies, using Gersho's conjecture and Zador's formula, that tessellating vector quantizers are asymptotically optimal for this broad class of sources, and generalizes a rigorous result of Gish and Pierce from the scalar to the vector case. Second, the asymptotic performance is established for Zamir and Feder's randomized lattice quantization. With the only assumption that the source has finite differential entropy, it is proven that the low-distortion performance of the Zamir–Feder universal vector quantizer is asympotically the same as that of the deterministic lattice quantizer.**

I. Introduction

Let Q_N^k denote an N-level k-dimensional vector quantizer, and let X^k be the k-dimensional random vector to be quantized. Let the rth power quantization distortion be defined in the usual way,

$$D_r(Q_N^k(X^k)) = \frac{1}{k} E\left\|X - Q_N^k(X^k)\right\|^r$$

where $\|\cdot\|$ denotes the Euclidean norm, and $r > 0$. Denote the Shannon entropy of Q_N^k by $H(Q_N^k)$, and for $H > 0$ let

$$D_e(H, k, r) = \inf_N \inf_{H(Q_N^k(X^k)) \le H} D_r(Q_N^k(X^k)), \tag{1}$$

the distortion of an optimal k-dimensional vector quantizer with entropy H. More precisely, $D_e(H, k, r)$ is the smallest distortion approachable arbitrarily by quantizers with finitely many levels with entropy-constraint H. It is not hard to see that we can allow quantizers with infinitely many levels if $E\|X\|^r < \infty$, in which case the value of $D_e(H, k, r)$ remains the same.

The quantity $D_e(H, k, r)$ was first investigated by Zador in two unpublished works [11] and [12]. His results later appeared in [13]. Zador found that for an X^k with a density f,

$$\lim_{H \to \infty} D_e(H, k, r) 2^{(r/k)H} = c_{k,r} 2^{(r/k)h(f)} \tag{2}$$

where $h(f) = -\int f \log f$ is the differential entropy of f, and $c_{k,r}$ is a constant that depends only on k and r. Unfortunately, the conditions needed for the validity of (2) are not precisely given in [13]. For this reason let us denote by $\mathcal{C}$ the class of densities for which (2) holds. A fundamental property of Zador's result is that once the precise asymptotic behavior of $D_e(H, k, r)$ is determined for any density in $\mathcal{C}$, (e.g., constant density over a convex bounded set) the constant $c_{k,r}$ is determined.

Gish and Pierce [5] investigated the low distortion behavior of entropy and resolution constrained quantizers for a certain class of

Manuscript received March 2, 1992; revised June 10, 1993. This work was supported in part by Hewlett-Packard and the National Science Foundation under Grants NCR-90-09766 and NCR-91-57770. This paper was presented in part at the IEEE International Symposium on Information Theory, Budapest, Hungary, June 1991.
T. Linder is with the Technical University of Budapest, Hungary.
K. Zeger is with the Department of Electrical Engineering, University of Illinois, Urbana, IL 61801 USA.
IEEE Log Number 9400305.

difference distortion measures. They outlined a rigorous proof of the claim that if the quantized random variable has a uniformly continuous density and finite differential entropy, then the infinite-level uniform step size entropy-constrained quantizer is asymptotically optimal for mean squared distortion. Let us denote the Δ step size infinite level uniform scalar quantizer by Q_Δ. Gish and Pierce also proved that if the density f of X is continuous except at finitely many points, and in the neighborhood of a discontinuity point the density behaves regularly enough, and its tail decreases fast enough, then

$$\lim_{\Delta\to 0}\left[H(Q_\Delta)+\frac{1}{2}\log 12 D_2(Q_\Delta)\right]=h(f). \tag{3}$$

These two results immediately give $c_{1,2}=\frac{1}{12}$ in Zador's formula (2). A straightforward extension of this argument gives $c_{1,r}=1/(r+1)2^r$.

For $k\geq 2$ the value of $c_{k,r}$ is unknown. Gersho [4] conjectured that for a uniform distribution over a convex bounded set in $\mathcal{R}^k$ the optimal entropy constrained vector quantizer will asymptotically have a partition whose regions are congruent with some tessellating convex polytope P. (Recall that a polytope P is *tessellating* if there exists a partition of $\mathcal{R}^k$ consisting of translated and/or rotated copies of P.) A quantizer of this type is called a *tessellating quantizer.* To present Gersho's conjecture more precisely let P be a k-dimensional *convex polytope* (a closed and bounded convex set in $\mathcal{R}^k$, which is the finite intersection of k-dimensional half-spaces) and let $\hat{y}$ be its centroid, i.e.,

$$\int_P \| x-\hat{y}\|^r\,dx=\inf_{y\in P}\int_P \| x-y\|^r\,dx.$$

The *normalized rth moment* of P is defined by

$$l(P)=\frac{1}{k}\frac{\int_P \| x-\hat{y}\|^r\,dx}{[\lambda(P)]^{(k+r)/k}} \tag{4}$$

where $\lambda(\cdot)$ denotes the k-dimensional volume (Lebesgue measure). The polytope P is *admissible* if: a) P is tessellating, and b) The Voronoi partition induced by the centroids of the copies coincides with the above partition. Gersho's conjecture (on entropy constrained asymptotic quantization) is, that in Zador's formula (2),

$$c_{k,r}=\inf l(P)\stackrel{\text{def}}{=}C(k,r) \tag{5}$$

where the infimum is taken over all k-dimensional admissible polytopes. A polytope for which the infimum is achieved is called *optimal.* The optimal polytope for $k=1$ is the interval, this being the only convex polytope in one-dimension, giving $C(1,r)=1/(r+1)2^r$. Thus Gersho's conjecture is in fact true in one dimension by the Gish–Pierce result.

A special case is when the admissible polytope is the basic Voronoi cell of a k-dimensional lattice. Thus, every lattice quantizer is a tessellating quantizer. On the other hand (as one can see by considering regular triangles) not all tessellating quantizers are lattice quantizers, so the validity of Gersho's conjecture would not imply that lattice quantizers are asymptotically optimal.

Since a rescaled admissible polytope is also admissible, the quantizer with quantization regions $P_\alpha=\{\alpha x: x\in P\}$, $\alpha>0$, is a tessellating quantizer if P is admissible. Denote this quantizer by $Q^k_{\alpha,P}$. In [4] Gersho gave a heuristic derivation of the asymptotic performance of these quantizers. He found that if $Q^k_{d,P}$ denotes the tessellating quantizer with rth power distortion d, then

$$\lim_{d\to 0} d2^{(r/k)H(Q^k_{d,P})}=l(P)2^{(r/k)h(f)}. \tag{6}$$

Yamada *et al.* in [10] took the same heuristic approach to extend Gersho's results to seminorm based distortion measures, i.e., distortion measures of the form $d(x,y)=L(\| x-y\|)$ where $\|\cdot\|$ is a seminorm and L is a "nice" function. To date however, no precise conditions for the validity of (6) have been determined.

In Section II, using a result of Csiszár, we prove (6) in great generality. Our Theorem 1 says that (6) holds whenever the quantized random vector has a density with finite differential entropy, and there exists at least one partition of $\mathcal{R}^k$ into regions of finite volume such that its entropy is finite. In particular, this theorem establishes the asymptotic entropy constrained performance of lattice quantizers without any smoothness or compact support condition on the density. Thus the often quoted formula (3) on the asymptotics of uniform quantizers is proved for all densities such that Q_Δ has finite Shannon entropy for some step size Δ and $h(f)<\infty$, strengthening Gish and Pierce's result. Assume now that Gersho's conjecture is true. Then the tessellating quantizer with the optimal polytope is asymptotically optimal for all source densities for which Zador's formula (2) holds and which satisfy the conditions of Theorem 1. Since our conditions are extremely general, this asymptotic optimality mostly depends on the validity of Zador's formula. In a similar vein, Na and Neuhoff [7] have recently strengthened Gersho's heuristic development of resolution constrained asymptotic vector quantization.

Section III deals with randomized lattice quantization. This quantization scheme was introduced by Ziv [15], who gave an upper bound on the difference between the rate of such a quantizer of dimension k and the rate of the optimal k-dimensional entropy constrained quantizer of the same mean squared distortion. This bound is valid for all source statistics and distortion levels, and gives 0.754 bits for cubic lattices. Zamir and Feder in [14] strengthened this result by showing the validity of the same upper bound on the difference between the rate of the randomized lattice quantizer and the kth order rate distortion function for any source having a density. Zamir and Feder considered the lattice quantizer $Q^k_{\alpha,V}$. This is a tessellating quantizer based on the admissible polytope V, the basic Voronoi cell of a lattice Λ. Their dithered lattice quantizer estimates the k-dimensional random vector X^k as

$$\hat{X}^k=Q^k_{\alpha,V}(X^k+Z^k_\alpha)-Z^k_\alpha \tag{7}$$

where the dither signal Z^k_α is uniformly distributed over the rescaled basic lattice cell αV, and is independent of X^k. The per sample rth power distortion of this quantizer is independent of the source statistics [14] and is seen to be

$$\frac{1}{k}E\|Q^k_{\alpha,V}(X^k+Z^k_\alpha)-Z^k_\alpha-X^k\|^r =\alpha^r\frac{1}{k\lambda(V)}\int_V \| t\|^r\,dt\stackrel{\text{def}}{=}d_\alpha. \tag{8}$$

In fact, in [14] more general distortion measures are considered, but the developed bounds are explicitly evaluated for rth power distortions. The scheme assumes that the decoder knows the values of the dither signal. Accordingly, the per sample average rate of the quantizer is given by the conditional entropy

$$\frac{1}{k}H(Q^k_{\alpha,V}(X^k+Z^k_\alpha)\mid Z^k_\alpha). \tag{9}$$

Zamir and Feder defined the *redundancy* of this randomized quantizer by

$$\rho_k(d_\alpha)=\frac{1}{k}H(Q^k_{\alpha,V}\mid Z^k_\alpha)-R_k(d_\alpha) \tag{10}$$

where $R_k(d)$ is the rate-distortion function of X^k. They showed in [14] that if X^k has a density, then for mean squared error

$$\rho_k(d_\alpha)\leq\frac{1}{2}\log 4\pi e G_k \tag{11}$$

where G_k is the usual notation for the normalized second moment of the lattice. They also observed that for high rates this bound can be improved. A derivation is given for the result

$$\limsup_{\alpha\to 0} \rho_k(d_\alpha) = \frac{1}{2}\log 2\pi e G_k \tag{12}$$

which improves the bound (11) by 1/2 bit. In [14] the derivation of (12) assumes that the density f of X^k satisfies the following conditions: a) f is bounded, and b) f is smooth enough in the sense that for all $\epsilon > 0$ there exists a $\delta > 0$ such that for all $x \in \mathcal{R}^k$

$$\frac{|f(x) - f(y)|}{f(x)} < \epsilon \tag{13}$$

whenever $\| x - y \| < \delta$. The boundedness condition excludes a large class of densities, and b) unfortunately is even more restrictive. Many continuous densities often used in modeling real data fail to satisfy (13). For example, it is easy to check that (13) is violated by Gaussian, Rayleigh, gamma, and beta one-dimensional source densities.

In Section III we prove that (12) holds for a large class of densities including all of the above listed cases. In Theorem 2 we show that if X^k has finite differential entropy, then for rth power distortion, the asymptotic performance of the randomized lattice quantizer (given by Theorem 1) and the asymptotic performance of the ordinary lattice quantizer (i.e., the lattice quantizer without randomization) are the same. Then Shannon's lower bound on the rate distortion function implies (12) for all the source distributions which satisfy this general condition.

II. Asymptotic Entropy-Constrained Performanceof Tessellating Quantizers

Our first lemma in this section determines the asymptotic rth power distortion of the tessellating vector quantizer $Q^k_{\alpha,P}$ defined in the Introduction.

Lemma 1: If the input random vector X^k has a density, then

$$\lim_{\alpha\to 0} \frac{D_r(Q^k_{\alpha,P}(X^k))}{\alpha^r} = \frac{1}{k\lambda(P)}\int_P \| x - \hat{y} \|^r \, dx. \tag{14}$$

Proof: It is not hard to see that the theorem holds when X^k has a uniform distribution over a compact set. We might proceed using uniformly continuous densities, which are approximately constant over the quantization regions, and then approximate an arbitrary density this way. However, there is a shorter and more elegant way to prove (14).

Let $P_{i,\alpha}$ and $y_{i,\alpha}, i = 1, 2, \cdots$ be enumerations of the polytopal quantization regions and the corresponding levels. Define the density f_α by

$$f_\alpha(x) = \frac{1}{\lambda(P_{i,\alpha})}\int_{P_{i,\alpha}} f(y)\,dy \qquad \text{if } x \in P_{i,\alpha}$$

for $i = 1, 2, \cdots$. Let X^k_α be a random variable with density f_α. Then

$$\begin{aligned} D_r(Q^k_{\alpha,P}(X^k_\alpha)) &= \frac{1}{k}\sum_i \int_{P_{i,\alpha}} \| x - y_{i,\alpha} \|^r f_\alpha(x)\,dx \\ &= \frac{1}{k}\sum_i \frac{\Pr\{X^k \in P_{i,\alpha}\}}{\lambda(P_{i,\alpha})}\int_{P_{i,\alpha}} \| x - y_{i,\alpha} \|^r \, dx. \end{aligned} \tag{15}$$

Now the fact that $\lambda(P_{i,\alpha}) = \alpha^k \lambda(P)$ and a simple change of variables show that

$$\int_{P_{i,\alpha}} \| x - y_{i,\alpha} \|^r \, dx = \int_P \alpha^{r+k} \| x - \hat{y} \|^r \, dx,$$

hence

$$D_r(Q^k_{\alpha,P}(X^k_\alpha)) = \frac{\alpha^r}{k\lambda(P)}\int_P \| x - \hat{y} \|^r \, dx. \tag{16}$$

On the other hand,

$$\begin{aligned} &\frac{1}{\alpha^r}\left| D_r(Q^k_{\alpha,P}(X^k_\alpha)) - D_r(Q^k_{\alpha,P}(X^k))\right| \\ &\quad \le \frac{1}{k\alpha^r}\sum_i \int_{P_{i,\alpha}} \| x - y_{i,\alpha} \|^r |f(x) - f_\alpha(x)|\,dx \\ &\quad \le \frac{1}{k}[\text{diam}\,(P)]^r \int_{\mathcal{R}^k} |f(x) - f_\alpha(x)|\,dx \end{aligned} \tag{17}$$

where diam (P) denotes the diameter of P. From Lebesgue's differentiation theorem [9], $f_\alpha \to f$ as $\alpha \to 0$ almost everywhere, from which via Scheffé's theorem [2], (17) tends to zero. From this and (16) we obtain

$$\begin{aligned} \lim_{\alpha\to 0} \frac{1}{\alpha^r} D(Q^k_{\alpha,P}(X^k)) &= \lim_{\alpha\to 0} \frac{1}{\alpha^r} D(Q^k_{\alpha,P}(X^k_\alpha)) \\ &= \frac{1}{k\lambda(P)}\int_P \| x - \hat{y} \|^r \, dx, \end{aligned}$$

and the lemma is proved. □

When $k = 1$ and a random variable X is quantized with a Δ step size uniform quantizer, then Lemma 1 gives the well-known formula

$$\lim_{\Delta\to 0} \frac{D_r(Q_\Delta(X))}{\Delta^r} = \frac{1}{(r+1)2^r},$$

for all source densities. Note that we don't require that the density f behave "sufficiently well" as is usually the case in the asymptotic theory.

To determine the asymptotic entropy of the quantizers $Q^k_{\alpha,P}$, we will use a result by Csiszár [3]. Following the work by Rényi [8], Csiszár investigated the entropy of partitions of abstract measure spaces. The following theorem is a special case of his general result.

Lemma 2: (Csiszár [3]) Let $Z = (Z_1, \cdots, Z_k)$ be an $\mathcal{R}^k$ valued random vector with density f_Z. Suppose that there exists some Borel measurable partition $\mathcal{B}_0 = \{B_1, B_2, \cdots\}$ of $\mathcal{R}^k$ into sets of finite Lebesgue measure such that

$$-\sum_n \Pr\{Z \in B_n\}\log \Pr\{Z \in B_n\} < \infty.$$

Suppose furthermore, that for some $\rho > 0$, some positive integers s, and for all k, the distance of B_k from any other B_l is greater than ρ for all but at most s indexes l. Let $\mathcal{A} = \{A_0, A_1, \cdots\}$ be a measurable partition with equal Lebesgue measure, i.e., $\lambda(A_i) = \epsilon$, $i = 1, 2, \cdots$, and let us denote the supremum of the diameters of the sets A_i by $\delta(\mathcal{A})$. Then we have

$$\lim_{\delta(\mathcal{A})\to 0} [H_{\mathcal{A}}(Z) + \log \epsilon] = h(f_Z)$$

where

$$H_{\mathcal{A}}(Z) = -\sum_n \Pr\{Z \in A_n\}\log \Pr\{Z \in A_n\},$$

and

$$h(f_Z) = -\int_{\mathcal{R}^k} f_Z(x)\log f_Z(x)\,dx,$$

the differential entropy of Z. Moreover, if Z has no density, then the above limit is $-\infty$. It should be mentioned that with the above conditions $h(f_Z)$ is always well-defined and $h(f_Z) < \infty$.

Taking $Z = X^k$, $\mathcal{A} = \{P_{1,\alpha}, P_{2,\alpha}, \cdots\}$ and $\epsilon = \lambda(P_\alpha)$ in Lemma 2, it follows that

$$\lim_{\alpha\to 0} [H(Q^k_{\alpha,P}) + \log \lambda(P_\alpha)] = h(f), \tag{18}$$

whenever $| h(f) | < \infty$ and $H(Q^k_{\alpha,P}(X)) < \infty$ for some $\alpha > 0$. Since (14) can be rewritten in the form

$$\lim_{\alpha\to 0} \frac{D(Q^k_{\alpha,P})}{\alpha^r} = l(P)[\lambda(P)]^{r/k}, \tag{19}$$

from (18) and (19) we obtain

$$\lim_{\alpha\to 0}\left[H(Q^k_{\alpha,P})+\frac{k}{r}\log D(Q^k_{\alpha,P})\right]=h(f)+\frac{k}{r}\log l(P). \quad (20)$$

It can be easily checked that for d small enough there exists at least one $Q^k_{\alpha,P}$ with $D(Q^k_{\alpha,P})=d$. This follows from the continuity of $D(Q^k_{\alpha,P})$, which can be shown by a standard argument, and from the fact that $D_r(Q^k_{\alpha,P})\to 0$ as $\alpha\to 0$. Denote such a quantizer by $Q^k_{d,P}$.

Theorem 1: If $|h(f)|<\infty$ and $H(Q^k_{\alpha,P}(X^k))<\infty$ for some $\alpha>0$, then

$$\lim_{d\to 0} d2^{(r/k)H(Q^k_{d,P})}=l(P)2^{(r/k)h(f)}. \quad (21)$$

Furthermore, if Zador's formula holds for f, $l(P)=C(k,r)$ (the optimal admissible polytope is used), and Gersho's conjecture (5) holds, then

$$\lim_{d\to 0}\frac{D_e(H(Q^k_{d,P}),k,r)}{d}=1, \quad (22)$$

i.e., the quantizer $Q^k_{d,P}$ is asymptotically optimal.

Proof: By the condition that there exists a tessellating quantizer $Q^k_{\alpha,P}$ whose output entropy is finite, Csiszár's lemma applies and therefore (18) holds. Noticing the obvious fact that as $d\to 0$ the scaling factor α for the corresponding $Q^k_{d,P}$ goes to zero, (21) follows directly from (20). We get the second statement upon simply substituting (5) and (21) into (2). □

Remark: When X^k has no density (21) is no longer valid. In this case, by Lemma 2, $H(Q^k_{\alpha,P})+\log\lambda(P_\alpha)\to-\infty$ as $\alpha\to 0$. Since $D(Q^k_{\alpha,P})\le\alpha^r\,\mathrm{diam}(P)$, it follows that

$$\lim_{\alpha\to 0}D(Q^k_{\alpha,P})2^{(r/k)H(Q^k_{\alpha,P})}=0. \quad (23)$$

We can say more than (23) when the distribution of X^k is known to be the mixture of a distribution with a density and a discrete distribution. Specifically, let the distribution of X^k be given by $\beta P_1+(1-\beta)P_2$ where $0<\beta<1$, P_1 is a probability measure with a density f, and P_2 is a discrete probability measure. Assume that f satisfies the conditions of Theorem 1, and P_2 is concentrated on a finite set of vectors $\{x_1,\cdots,x_n\}$ with probabilities $\{p_1,\cdots,p_n\}$. With a slight modification of Lemma 2 it can be proved that

$$\lim_{\alpha\to 0}[H(Q^k_{\alpha,P})+\beta\log\lambda(P_\alpha)] =\beta h(f)+(1-\beta)H(P_2)+H(\beta) \quad (24)$$

where $H(P_2)$ is the Shannon entropy of $\{p_1,\cdots,p_n\}$ and $H(\beta)=-\beta\log\beta-(1-\beta)\log(1-\beta)$. From this we have

$$\lim_{\alpha\to 0}D(Q^k_{\alpha,P})2^{\frac{1}{\beta}\frac{r}{k}H(Q^k_{\alpha,P})} \le C(P)l(P)2^{\frac{r}{k}(\beta h(f)+(1-\beta)H(P_2)+H(\beta))} \quad (25)$$

where $C(P)$ is a constant depending on the ratio of the diameter and the volume of P.

A standard technique using the vector Shannon lower bound on the rate-distortion function (see Gray [6]) can be used to compare the performance of the quantizers $Q^k_{\alpha,P}$ to the rate-distortion bound. For $r=2$ the Shannon lower bound on the kth-order rate-distortion function for random vectors with a density f is

$$R_k(d)\ge\frac{1}{k}h(f)-\frac{1}{2}\log 2\pi ed. \quad (26)$$

This, in combination with (21), gives

$$\limsup_{d\to 0}\left[\frac{1}{k}H(Q^k_{d,P})-R_k(d)\right]\le\frac{1}{2}\log 2\pi eG_k \quad (27)$$

where G_k denotes the normalized second moment of P. The condition for (27) to hold is that $E\,\|X^k\|^2<\infty$, $|h(f)|<\infty$, and $H(Q^k_{\alpha,P}(X^k))<\infty$ for some $\alpha>0$. For "sufficiently nice" densities there is equality in (27), since in this case the difference between $R_k(d)$ and Shannon's lower bound vanishes as d does to zero (c.f., Theorem 4.3.5 in [1]). However, the authors are not aware of any result general enough asserting this convergence for all densities satisfying the above conditions. This result generalizes statements by Gish and Pierce [5], Gersho [4], and Yamada *et al.* [10] (since any nonnegative continuous function of a seminorm would work in (14)–(17)).

III. Asymptotic Entropy-Constrained Performance of Randomized Lattice Quantizers

As was mentioned in the Introduction, the rth power distortion $\overline{D}_r(Q^k_{\alpha,V})$ of the randomized lattice quantizer is given by (8). The bar above D indicates the randomized distortion as opposed to the distortion of the deterministic lattice quantizer. In fact, the uniformly distributed dither signal makes the derivation of (8) straightforward and the formula is true for *arbitrary source statistics*. Comparing (8) and Lemma 1 shows that the randomized and nonrandomized distortion of the lattice quantizer $Q^k_{\alpha,V}$ are asymptotically the same whenever the source has a density.

The next lemma shows that the randomized lattice quantizer has the same asymptotic entropy as the deterministic one.

Lemma 3: Suppose that X^k has a density, $|h(f)|<\infty$, and $H(Q^k_{\alpha,V}(X^k))<\infty$ for some $\alpha>0$. Then

$$\lim_{\alpha\to 0}[H(Q^k_{\alpha,V}(X^k+Z^k_\alpha)\mid Z^k_\alpha)+\log\lambda(V_\alpha)]=h(f). \quad (28)$$

Proof: Let the density of Z^k_α be denoted by f_{Z_α}. Then we have

$$H(Q^k_{\alpha,V}(X^k+Z^k_\alpha)\mid Z^k_\alpha)+\log\lambda(V_\alpha) =\int_{V_\alpha}[H(Q^k_{\alpha,V}(X^k+z))+\log\lambda(V_\alpha)]f_{Z_\alpha}(z)\,dz. \quad (29)$$

Now $Q^k_{\alpha,V}(X^k+z)$ clearly has the same entropy as the shifted lattice quantizer with tessellating partition $\alpha\Lambda-z$. Then by (18), for any fixed z we have

$$\lim_{\alpha\to 0}[H(Q^k_{\alpha,V}(X^k+z))+\log\lambda(V_\alpha)]=h(f).$$

But Lemma 2 readily implies that this convergence is uniform in z, i.e., $|H(Q^k_{\alpha,V}(X^k+z))+\log\lambda(V_\alpha)-h(f)|\le\epsilon(\alpha)$ for all $z\in\mathcal{R}^k$, for some $\epsilon(\alpha)\to 0$ as $\alpha\to 0$. This and (29) prove the lemma. □

Remark: Zamir and Feder [14] showed that

$$H(Q^k_{\alpha,V}(X^k+Z^k_\alpha)\mid Z^k_\alpha)+\log\lambda(V_\alpha)=h(f_{X-Z_\alpha})$$

for all $\alpha>0$ where f_{X-Z_α} is the density of $X^k-Z^k_\alpha$. In view of Lemma 3 we can conclude that

$$\lim_{\alpha\to 0}h(f_{X-Z_\alpha})=h(f)$$

whenever the conditions of Lemma 3 hold.

Now we are in a position to relate the asymptotic performance of the randomized lattice quantizer to that of the deterministic lattice quantizer. By (8) for any $d>0$ if $\alpha=d^{1/r}l(V)^{-1/r}\lambda(V)^{-1/k}$, then $\overline{D}_r(Q^k_{\alpha,V})=d$. Denote this quantizer by $Q^k_{d,V}$, and its rate by $\overline{H}(Q^k_{d,V})$.

Theorem 2: Suppose that X^k has a density, $|h(f)|<\infty$, and $H(Q^k_{\alpha,V}(X^k))<\infty$ for some $\alpha>0$. Then the rate of the randomized lattice quantizer with rth power distortion d satisfies

$$\lim_{d\to 0}d2^{\frac{r}{k}\overline{H}(Q^k_{d,V})}=l(V)2^{\frac{r}{k}h(f)}, \quad (30)$$

i.e., the asymptotic performance of the randomized lattice quantizer is the same as the asymptotic performance of the ordinary (non-randomized) lattice quantizer given by (21). Note that $l(V)$ is defined by (4).

Proof: The substitution of expressions for the distortion (8) into (28), using expression (4), readily gives (30). □

Corollary 1: For $r = 2$, with the conditions of Theorem 2, we have

$$\limsup_{d\to 0}\left[\frac{1}{k}\overline{H}(Q^k_{d,V}) - R_k(d)\right] \leq \frac{1}{2}\log 2\pi e G_k. \quad (31)$$

IV. Conclusion

We have established the asymptotic equivalence between entropy-constrained lattice quantization and a universal quantization scheme based on dithering. Very unrestricted assumptions on the source density are imposed. Although our derivations assumed an rth power distortion measure, the proofs can easily be extended to more general distortions such as continuous functions of seminorms.

Acknowledgment

The authors wish to express their gratitude to L. Györfi for helpful discussions and to the referees for making useful suggestions. In particular, the simple proof of Lemma 3 was suggested by an anonymous referee.

References

[1] T. Berger, *Rate Distortion Theory.* Englewood Cliffs, NJ: Prentice-Hall, 1971.
[2] P. Billingsley, *Convergence of Probability Measures.* New York: Wiley, 1968.
[3] I. Csiszár, "Generalized entropy and quantization problems," in *Trans. Sixth Prague Conf. Inform. Theory, Statist. Decision Functions, Random Processes*, (Prague), pp. 29–35, Akademia, 1973.
[4] A. Gersho, "Asymptotically optimal block quantization," *IEEE Trans. Inform. Theory*, vol. IT-25, pp. 373–380, July 1979.
[5] H. Gish and J. N. Pierce, "Asymptotically efficient quantizing," *IEEE Trans. Inform. Theory*, vol. IT-14, pp. 676–683, Sept. 1968.
[6] R. M. Gray, *Source Coding Theory.* Boston, MA: Kluwer, 1990.
[7] S. Na and D. L. Neuhoff, "Bennett's integral for vector quantizers," preprint 1991.
[8] A. Rényi, "On the dimension and entropy of probability distributions," *Acta Math. Academic Sci., Hungary*, vol. 10, pp. 193–215, 1959.
[9] R. L. Wheeden and A. Z. Zygmund, *Measure and Integral.* New York: Marcel-Dekker, 1977.
[10] 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.
[11] P. Zador, "Development and evaluation of procedures for quantizing multivariate distributions," Ph.D. dissertation, Stanford Univ., 1964; Univ. Microfilm 64-9855.
[12] ——, "Topics in the asymptotic quantization of continuous random variables," unpublished memorandum, Bell Lab., Murray Hill, NJ, Feb. 1966.
[13] ——, "Asymptotic quantization error of continuous signals and the quantization dimension," *IEEE Trans. Inform. Theory*, vol. IT-28, pp. 139–149, Mar. 1982.
[14] R. Zamir and M. Feder, "On universal quantization by randomized uniform/lattice quantizers," *IEEE Trans. Inform. Theory*, vol. IT-38, Mar. 1992.
[15] J. Ziv, "On universal quantization," *IEEE Trans. Inform. Theory*, vol. IT-31, pp. 344–347, May 1985.

Lower Bounds for the Complexity of Reliable Boolean Circuits with Noisy Gates

Péter Gács and Anna Gál

***Abstract*— We prove that the reliable computation of any Boolean function with sensitivity s requires $\Omega(s \log s)$ gates if the gates fail independently with a fixed positive probability. This theorem was stated by Dobrushin and Ortyukov in 1977, but their proof was found by Pippenger, Stamoulis, and Tsitsiklis to contain some errors.**

***Index Terms*—Reliable computation, noisy gates, Boolean functions.**

I. Introduction

In this paper, we prove lower bounds on the number of gates needed to compute Boolean functions by circuits with noisy gates. We say that a gate fails if its output is incorrect. Let us fix a bound $\epsilon \in (0, 1/2)$ on the failure probability of the gates and a bound $p \in (0, 1/2)$ on the probability that the value computed by the circuit is incorrect. These parameters will be held constant throughout the paper, and dependence on them will not be explicitly indicated either in the defined concepts like redundancy, or in the $O(\,)$ and $\Omega(\,)$ notation.

A *noisy gate* fails with a probability bounded by ϵ. A *noisy circuit* has noisy gates that fail independently.

A noisy circuit is *reliable* if the value computed by the circuit on any given input is correct with probability $\geq 1 - p$. The size of a reliable noisy circuit has to be larger than the size needed for circuits using only correct gates. By the *noisy complexity* of a function we mean the minimum number of gates needed for the reliable computation of the function. Note that in this model the circuit cannot be more reliable than its last gate. For a given function, the ratio of its noisy and noiseless complexities is called the *redundancy* of the noisy computation of the function.

The following upper bounds are known for the noisy computation of Boolean functions. The results of von Neumann [9], Dobrushin and Ortyukov [3], and Pippenger [11] prove that if a function can be computed by a noiseless circuit of size L, then $O(L \log L)$ noisy gates are sufficient for the reliable computation of the function. Pippenger [11] proved that any function depending on n variables can be computed by $O(2^n/n)$ noisy gates. Since the noiseless computation of almost all Boolean functions requires $\Omega(2^n/n)$ gates (Shannon [15], Muller [8]), this means that for almost all functions the redundancy of their noisy computation is just a constant. Pippenger [11] also exhibited specific functions with constant redundancy. For the noisy computation of any function of n variables over a complete basis Φ, Uhlig [16] proved upper bounds arbitrarily close to $\rho(\Phi)2^n/n$ as $\epsilon \to 0$ where $\rho(\Phi)$ is a constant depending on Φ, and $\rho(\Phi)2^n/n$ is the asymptotic bound for the noiseless complexity of almost all Boolean functions of n variables (Lupanov [7]).

Manuscript received April 24, 1992. P. Gács was supported in part by the NSF under Grant CCR-9002614. A. Gál was supported in part by the NSF under Grant CCR-8710078 and by OTKA under Grant 2581. This paper was presented in part at the 32nd IEEE Symposium on the Foundations of Computer Science, San Juan, Puerto Rico, October 1991.

P. Gács is with the Department of Computer Science, Boston University, Boston, MA 02215 USA.

A. Gál is with the Department of Computer Science, The University of Chicago, Chicago, IL 60637 USA.

IEEE Log Number 9400229.