Asymptotically Optimal Block Quantization

ALLEN GERSHO, SENIOR MEMBER, IEEE

***Abstract*—In 1948 W. R. Bennett used a companding model for nonuniform quantization and proposed the formula**

$$D=\frac{1}{12N^2}\int p(x)[E'(x)]^{-2}\,dx$$

for the mean-square quantizing error where N is the number of levels, $p(x)$ is the probability density of the input, and $E'(x)$ is the slope of the compressor curve. The formula, an approximation based on the assumption that the number of levels is large and overload distortion is negligible, is a useful tool for analytical studies of quantization. This paper gives a heuristic argument generalizing Bennett's formula to block quantization where a vector of random variables is quantized. The approach is again based on the asymptotic situation where N, the number of quantized output vectors, is very large. Using the resulting heuristic formula, an optimization is performed leading to an expression for the minimum quantizing noise attainable for any block quantizer of a given block size k. The results are consistent with Zador's results and specialize to known results for the one- and two-dimensional cases and for the case of infinite block length ($k\to\infty$). The same heuristic approach also gives an alternate derivation of a bound of Elias for multidimensional quantization. Our approach leads to a rigorous method for obtaining upper bounds on the minimum distortion for block quantizers. In particular, for $k=3$ we give a tight upper bound that may in fact be exact. The idea of representing a block quantizer by a block "compressor" mapping followed with an optimal quantizer for uniformly distributed random vectors is also explored. It is not always possible to represent an optimal quantizer with this block companding model.

I. Introduction

Digital coding of analog sources is today a subject of considerable importance, yet very little is understood about optimal block quantization. On the one hand, extensive results are available for the one-dimensional (or zero-memory) quantizer. On the other hand, useful bounds are available in the limiting case where the block length approaches infinity. What is needed is a theory of quantization for finite block lengths of arbitrary size. In this note an attempt is made to apply some of the appealing features of the one-dimensional theory to the study of block quantization. A heuristically derived formula is found for the asymptotic case of high-quality quantization. This formula specializes to known results for the one- and two-dimensional cases and for the limiting case of infinite block length.

Manuscript received March 3, 1978; revised January 9, 1979. This work was supported in part by the Electronics Program of the Office of Naval Research while the author was visiting the Department of System Science at the University of California, Los Angeles. This paper was presented at the Information Theory International Symposium, Cornell University, Ithaca, NY, October 10–14, 1977.

The author is with Bell Laboratories, Murray Hill, NJ 07974.

II. Formulation

Let X be a k dimensional random vector with joint density $p(x)=p(x_1,x_2,\cdots,x_k)$. An N point "block" quantizer is a function $Q(x)$ which maps x in R_k into one of N output vectors or "output points" $y_1,y_2,\cdots,y_N$ each in R_k. The quantizer is specified by the values of the output points and by a partition of the space R_k into N disjoint and exhaustive regions $S_1,S_2,\cdots,S_N$, where $S_i=Q^{-1}(y_i)\subset R_k$. Then

$$Q(x)=y_i,\qquad \text{if } x \text{ is in } S_i$$

for $i=1,2,\cdots,N$. The term "block" quantizer is used to indicate that the quantizer operates on a "block" of k random variables, i.e., a k-dimensional random vector.

The performance of such a quantizer can be measured by the distortion:

$$D=\frac{1}{k}E\|X-Q(X)\|^r$$

where $\|\cdot\|$ denotes the usual l_2 norm. We assume that $E\|X\|^r$ is finite. Note that for $r=2$, D is the familiar mean-square "per-letter" distortion measure and for $k=1$ it is the usual rth absolute moment of the quantizing error.

We wish to determine a) the minimum distortion $D_1(N)$ attainable over the set of all N point quantizers and b) the minimum distortion $D_2(H_Q)$ attainable over the set of all quantizers having a fixed output entropy H_Q where

$$H_Q=-\sum_1^N p_i\log p_i\qquad p_i=P\{X\in S_i\}.$$

We consider only the asymptotic case of high quality quantization where N is very large in problem a) or H_Q is very large in problem b).

III. Previous Work

In 1948 W. R. Bennett [1] modeled the one-dimensional quantizer using a memoryless monotonically increasing nonlinearity $E(x)$ (called the compressor) followed by a uniform N point quantizer. This model is completely general since any finite partition of the real line into intervals can be obtained in this way using a suitable continuous compressor curve. He showed that the distortion could be approximated by the integral

$$D\cong\frac{1}{12N^2}\int_{L_1}^{L_2}\frac{p(x)}{[\lambda(x)]^2}\,dx \tag{1}$$

where $\lambda(x)=E'(x)/(L_2-L_1)$. The result assumes that the $N-2$ finite regions S cover the interval (L_1,L_2) and that L_1 and L_2 are appropriately chosen so that the contribution to the distortion due to the tail or "overload" regions can be neglected. The integral is based on some implicit regularity conditions on the density $p(x)$ and on the assumption that N is very large. Bennett's formula is a convenient analytical tool for optimization studies of one dimension quantization.

Several authors have pursued the problem of minimizing the distortion in one-dimensional quantization. Panter and Dite [2] found an expression for the minimum mean-square distortion ($r=2$) for large N. Lloyd [3] found optimality conditions and an algorithmic approach for finding the optimum quantizer valid for each N. Smith [4] was the first to use Bennett's formula to find the best compressor curve and the minimum distortion for optimum quantization for large N. Algazi [5] used the rth power distortion measure and showed that for large N the minimum distortion is

$$D_1(N)=\frac{1}{r+1}(2N)^{-r}\|p_1(x_1)\|_{1/(1+r)} \qquad (2)$$

where $p_1(x_1)$ is the (one-dimensional) density of the random variable X_1 and

$$\|p(x)\|_\alpha=\left[\int [p(x)]^\alpha\, dx\right]^{1/\alpha}. \qquad (3)$$

For the fixed entropy constraint, Gish and Pierce [6] solved the one-dimensional problem. Their result can be expressed as

$$D_2(H_Q)=\frac{1}{r+1}2^{-r}e^{-r[H_Q-H(p_1)]} \qquad (4)$$

where $H(p_1)$ is the differential entropy of X_1, $H(p_1)=-\int p_1(x_1)\log p_1(x_1)\,dx_1$. Both (2) and (4) are valid only as asymptotic results for high quality quantization.

Extensions to block quantization have been studied by Zador [7], Schutzenberger [8], and Elias [9]. Schutzenberger derived an inequality bounding the distortion D for given H_Q, namely,

$$D\geqslant Ke^{-H_Q r/k}$$

where K is an unspecified constant depending on k, r, and $p(x)$. Elias defined the quantizer distortion measure $D^*=\Sigma_i^N P(S_i)[V(S_i)]^{r/k}$ where $V(S_i)$ is the k-dimensional volume of the region S_i. He showed that

$$D^*\geqslant N^{-r/k}\|p(x)\|_{k/k+r} \qquad (5)$$

where $\|p(x)\|_\alpha$, the L_α norm of $p(x)$, is defined in (3) except that the integration is now k-dimensional. He also showed that for N sufficiently large there exists a quantizer with D^* arbitrarily close to this bound. Elias assumed the input vector x to be bounded so that each region S_i has finite volume. Zador, in a lengthy unpublished manuscript, found for the asymptotic case of high quality quantization that

$$D_1(N)=A(k,r)N^{-r/k}\|p(x)\|_{k/(k+r)} \qquad (6)$$

and that

$$D_2(H_Q)=B(k,r)e^{-r/k[H_Q-H(p)]} \qquad (7)$$

where $H(p)$ is the differential entropy of the random vector X. Zador did not obtain $A(k,r)$ or $B(k,r)$ explicitly, but he showed that

$$\frac{1}{1+\beta}V_k^{-\beta}\leqslant kB(k,r)\leqslant kA(k,r)\leqslant\Gamma(1+\beta)V_k^{-\beta} \qquad (8)$$

where $\beta=r/k$, $\Gamma(x)$ is the gamma function, and V_k is the volume of a unit sphere in k dimensions. A key feature of Zador's result is that $A(k,r)$ and $B(k,r)$ are independent of the density $p(x)$. Hence these functions can be studied in the simpler context of the uniform density on a unit cube in k dimensions. For $k=2$ and $r=2$, Fejes Toth [10] found the (asymptotically) minimum distortion having the form (6) with the explicit value $A(2,2)=5\sqrt{3}/108$.

Recently Gray and Gray [11] gave a simple derivation of the one-dimensional results (2) and (4) using Bennett's integral. Here we explore a similar approach but generalized to the k-dimensional case.

IV. Admissable and Optimal Polytopes

In the one-dimensional case, Bennett's integral is derived by separating the description of a quantizer into two aspects: a) a uniform quantizer that is optimal for the uniform density function and b) the compressor slope function which determines how the output points of the uniform quantizer must be redistributed to take into account the probability density function of the random variable to be quantized. Zador's expression for the minimal distortion also has the striking feature that the factor $A(k,r)$ is independent of the probability density of the random variable. Since $\|p(x)\|_\alpha=1$ if $p(x)$ is unity in a bounded region of unit volume and zero elsewhere, it follows that $A(k,r)$ is determined by the optimal quantizer for a *uniformly* distributed random variable. These observations suggest that Bennett's integral can be generalized by first considering the optimum quantizer for a uniformly distributed k-dimensional random variable and then considering the effect of a nonuniform distribution of output points on the distortion of a quantizer. We begin by exploring some relevant geometrical features of partitions in R_k.

For every finite (or countably infinite) set of points $y_1,y_2,\cdots,y_N$ in R_k a Dirichlet partition is defined with each point in S_i closer to y_i than to any other y_j for $j\neq i$. That is,

$$S_i=\{x:\ \|x-y_i\|\leqslant\|x-y_j\| \text{ for each } j\neq i\}.$$

An optimal N-point quantizer that minimizes distortion will clearly have a Dirichlet partition. An example of a Dirichlet partition in the plane is shown in Fig. 1. In general, each bounded Dirichlet region is a polytope (bounded by segments of $k-1$ dimensional hyperplanes) and is convex. For a quantizer an effective partition would have the property that the unbounded regions or "overload" regions would make a sufficiently small contribution to the distortion. This is always possible when $E\|X\|^r<\infty$.

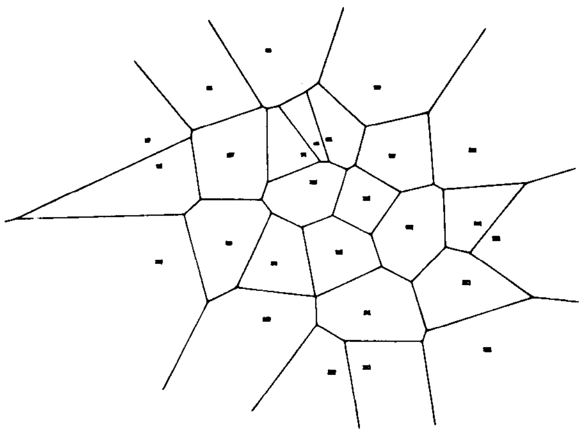

Fig. 1. Direchlet partition for Cambridge, Massachusetts, schools (from H. L. Loeb [17]).

The *centroid* $\hat{y}$ of a convex polytope H in R_k is the value of y that minimizes $\int_H \|x-y\|^r \, dx$. For $r=2$, $\hat{y}$ coincides with the usual definition for the centroid of a body with uniform mass distribution. It should be noted (see Fig. 1) that in general the points generating a Dirichlet partition are not necessarily the centroids of their respective regions. For a uniformly distributed random vector x, a quantizer will have a Dirichlet partition defined on the bounded set in R_k where $p(x)$ is positive. For the quantizer that minimizes distortion it is clearly necessary that each output point will be the centroid of the region in which it lies. The two necessary conditions for optimality, i.e., that the partition be a Dirichlet partition and that the output points be centroids, were noted for $k=1$ by Lloyd [3].

A convex polytope H is said to generate a tessellation if there exists a partition of R_k whose regions are all congruent to H. For example in the plane all triangles, quadrilaterals, and hexagons generate tessellations.

We now make the basic conjecture that for N sufficiently large the optimal (distortion-minimizing) quantizer for a random vector uniformly distributed on some convex set S will have a partition whose regions are all congruent to some polytope H, with the possible exception of regions touching the boundary of S. In other words, the optimal partition is essentially a tessellation of S. This conjecture plays a key role in the heuristic approach which follows.

We define H_k, the class of admissible polytopes in R_k as follows. A convex polytope H in R_k is in H_k if a)H generates a tessellation that is a Dirichlet partition with respect to the centroids of each region in the partition. For example, the equilateral triangle, the rectangle, and the regular hexagon are the admissible polygons in H_2. Fig. 2 illustrates a tessellation of the regular hexagon. Now we define the *normalized inertia* $I(H)$ of a polytope H as

$$I(H)=\int_H \|x-\hat{x}\|^r \, dx / [V(H)]^{1+r/k} \tag{9}$$

where $\hat{x}$ is the centroid of H and $V(H)$ is the k-dimensional volume of H. The normalization has the property

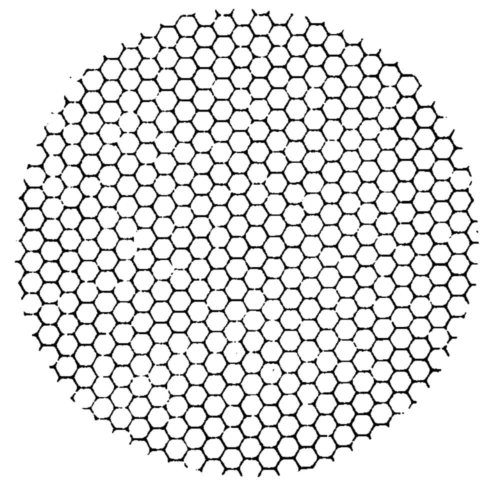

Fig. 2. Tesselation of regular hexagons.

that $I(\alpha H)=I(H)$ for $\alpha>0$ where the polytope $\alpha H=\{\alpha x: x\in H\}$. In other words, when the size of H is scaled, its normalized inertia remains unchanged.

We define the *coefficient of quantization*

$$C(k,r) \triangleq \frac{1}{k} \inf_{H\in H_k} I(H).$$

An *optimal polytope* H^* is an admissible polytope which attains the minimum inertia of all admissable polytopes with the same volume. Hence

$$I(H^*)=kC(k,r).$$

By calculating the normalized inertia of each admissible polygon, it can be shown that for $k=2$ and $r=2$, the optimal polytope is the regular hexagon. We conjecture that an optimal polytope exists for each k.

It is a classic isoperimetric result that every convex polytope has a greater moment of inertia with respect to its centroid than a k-dimensional sphere with the same volume. For the unit radius sphere B centered at the origin it is known that

$$\int_B \|x\|^r \, dx = \frac{k}{k+r} V_k$$

where V_k is the volume of B. Hence we have

$$I(B)=\frac{k}{k+r} V_k^{-r/k}$$

so that we have the lower bound

$$C(k,r) \geq \frac{1}{k+r} V_k^{-r/k}. \tag{10}$$

An upper bound on $kC(k,r)$ can be found by calculating the normalized inertia for any admissible polytope in H_k. One such choice is the k-dimensional cube (centered at the origin), which is clearly admissible. The cube has normalized inertia $k/[(r+1)2^r]$ so that

$$C(k,r) \leq \frac{1}{1+r} 2^{-r}. \tag{11}$$

Note that this bound is independent of the dimension k.

V. Heuristic Derivation of the Distortion Integral

Generalizing the concept of "asymptotic fractional density of quanta" introduced by Lloyd in a classic paper [3] on one-dimensional quantization, define the output point density function of a k-dimensional quantizer as

$$g_N(x)=\frac{1}{NV(S_1)}, \qquad \text{if } x\in S_i, \text{ for } i=1,2,\cdots,N.$$

where $V(S_1)$ denotes the volume of S_1. Note that $g_N(x)=0$ if x is in a region of the partition having infinite volume. In the asymptotic situation where N is very large, $g_N(x)$ can be expected to approximate closely a continuous density function $\lambda(x)$ having unit volume. Then $\lambda(x)\Delta V(x)$ may be taken as the fraction of output points located in an incremental volume element $\Delta V(x)$ containing x. Thus the volume of the quantizing region S_i associated with the output point y_i is given approximately by

$$V(S_i)\approx\frac{1}{N\lambda(y_i)} \tag{12}$$

for every bounded region S_i. Note that $N\lambda(y_i)$ is the number of points per unit volume in the neighborhood of y_i so that its reciprocal (12) is the volume per output point.

The distortion (1) can be expressed as

$$D=\frac{1}{k}\sum_{i=1}^{N}\int_{S_i}\|x-y_i\|^r p(x)\,dx. \tag{13}$$

For N large it is reasonable to assume that most of the regions S_i will be bounded sets, and the "overload" regions S_i will correspond to the tail region of the density $p(x)$. Assume the partition has been suitably chosen so that the overload distortion is negligible, treat N as the number of bounded regions, and for N large make the approximation

$$p(x)\backsim p(y_i), \qquad \text{for } x\in S_i.$$

Then we obtain

$$D=\frac{1}{k}\sum_{i=1}^{N}p(y_i)\int_{S_i}\|x-y_i\|^r\,dx. \tag{14}$$

As N becomes large the partition for any bounded region should look more and more like the partition for a uniform density, assuming $\lambda(x)$ is smoothly varying. Thus we approximate S_i by a suitably rotated, translated, and scaled optimal polytope H^*. Then

$$\int_{S_i}\|x-y_i\|^r\,dx=I(H^*)[V(S_i)]^{1+r/k} \tag{15}$$

using (9). We then have

$$D=\frac{1}{k}\sum_{r=1}^{N}p(y_i)I(H^*)[V(S_i)]^{1+r/k}, \tag{16}$$

and from (12) we obtain

$$D=N^{-\beta}C(k,r)\sum_{i=1}^{N}p(y_i)[\lambda(y_i)]^{-\beta}V(S_i). \tag{17}$$

The summation can be approximated by an integral yielding

$$D=N^{-\beta}C(k,r)\int\frac{p(y)}{[\lambda(y)]^{\beta}}\,dy. \tag{18}$$

The region of integration is actually the union of all bounded regions of the partition but may be taken to be the entire k-dimensional space since the contribution to the distortion of the overload regions will be negligible for any reasonable quantizer with sufficiently large N.

Equation (18) may be recognized as the k-dimensional version of Bennett's formula (1) for one-dimensional quantization with mean-square distortion.

VI. Minimization of the Distortion Integral

The distortion integral (18) allows the minimization of the distortion by optimizing the choice of $\lambda(x)$, the asymptotic output point density function. No reference is needed to the explicit quantizer characteristics (the output points and partition regions).

For problem a), D is to be minimized over all quantizers with N fixed. Hölder's inequality gives

$$\int p\lambda^{-\beta}\,dy\{\int\lambda\,dy\}^{\beta} \geqslant\{\int(p\lambda^{-\beta})1/(1+\beta)\lambda\beta/(1+\beta)\,dy\}^{1+\beta}.$$

Noting that $\int\lambda\,dy=1$, we obtain the result

$$\int p\lambda^{-\beta}dy\geqslant\|p\|_{1/(1+\beta)}$$

with equality attained only when λ is proportional to $p^{1/(1+\beta)}$. Hence the minimum value of D, referring to (18), is

$$D_1(N)=C(k,r)N^{-\beta}\|p(x)\|_{k/(k+r)}. \tag{19}$$

This is the desired result. Note that (19) coincides with Zador's result (6) when we take $A(k,r)=C(k,r)$. Furthermore using (10) we obtain a lower bound for $D_1(N)$ that coincides with Zador's lower bound.

A significant property of the optimum quantizer can now be demonstrated. Since the optimum point density λ is proportional to $p^{1/(1+\beta)}$, we observe that each term in the sum (16) reduces to a constant independent of the index i. *Therefore each region S_i of the partition makes an equal contribution to the distortion for an optimal quantizer.* This property was observed by Panter and Dite [2] for $k=1$ and by Fejes Toth [10] for $k=2$.

In problem b), D is to be minimized subject to a constraint on the quantizer output entropy H_Q. Since $p_i\backsim p(y_i)V(S_i)$ for each bounded set S_i and for large N,

$$H_Q=-\Sigma p(y_i)\frac{1}{N\lambda(y_i)}\log[p(y_i)/N\lambda(y_i)]$$

$$=-\Sigma p(y_i)\log p(y_i)\Delta V(y_i)-\Sigma p(y)\log\frac{1}{N\lambda(y_i)}\Delta V(y_i)$$

where $\Delta V(y_i)=1/N\lambda(y_i)$. As in the derivation of (18), the sums can be approximated by integrals, yielding

$$H_Q = H(p) - \int p(y) \log \frac{1}{N\lambda(y)}\, dy \tag{20}$$

where $H(p)$ is the differential entropy of the random vector $\boldsymbol{X}$. Equation (20) reduces for $k=1$ to the corresponding one-dimensional result given by Gish and Pierce [6]. From (18) we have

$$D = C(k,r) \int e^{-\beta \log [N\lambda(y)]} p(y)\, dy. \tag{21}$$

Now applying Jensen's inequality we get

$$D \geqslant C(k,r) e^{-\beta \int p(y) \log [N\lambda(y)]\, dy}. \tag{22}$$

Applying (20) we see that

$$D \geqslant C(k,r) e^{-\beta [H_Q - H(p)]}. \tag{23}$$

The application of Jensen's inequality yields an equality when $\lambda(y)$ is a constant corresponding to a *uniform* distribution of output points. Hence the solution to problem b) is

$$D_2(H_Q) = C(k,r) e^{-\beta [H_Q - H(p)]}. \tag{24}$$

Note that (24) coincides with Zador's result (7) when we take $B(k,r) = C(k,r)$. Furthermore applying the lower bound (10) to (24) gives a bound for $D_2(H_Q)$ which coincides with Zador's lower bound (8). *It is significant to observe that for large N the optimal quantizer for a constrained entropy is very nearly, the uniform quantizer.* For $k=1$ this was noted by Gish and Pierce [6].

As an additional illustration of the use of the function $\lambda(x)$ we give a heuristic derivation of Elias' result [9]. Since

$$P(S_i) \approx p(y_i) V(S_i),$$

$$D^* \approx \sum_{i=1}^{N} p(y_i) V(S_i)^{r/k} V(S_i).$$

Using (11) gives

$$D^* \approx \sum_{i=1}^{N} p(y_i) \left[\frac{1}{N\lambda(y_i)} \right]^{r/k} \Delta V(y_i),$$

and approximating the sum for N large by an integral yields

$$D^* = N^{-\beta} \int \frac{p(y)}{[\lambda(y)]^{\beta}}\, dy. \tag{25}$$

The minimization of this integral as shown above leads to the result that

$$D^* = N^{-\beta} \| p(x) \|_{k/k+r} \tag{26}$$

which is Elias' lower bound.

Finally, it should be noted that the formulas (2), (4), (6), (7), (18), (19), (24), and (26), which have been written as equalities, should more correctly be taken as lower bounds on attainable distortion for any finite N. Since the minimum distortion attainable is nonincreasing as N (or H_Q) increases, the actual distortion can only be greater than these asymptotic values for any quantizer with a finite number of quantizing regions N.

VII. Special Cases and Bounds

For $k=1$, a finite interval on the real line is the only admissible polytope. The interval is therefore the optimal polytope for $k=1$. Calculating its normalized inertia gives

$$C(1,r) = \frac{1}{r+1} 2^{-r}.$$

For $r=2$, we have $C(1,2) = 1/12$ and hence our generalized Bennett integral (18) reduces to the original Bennett integral (1). For $k=1$, the minimum distortion formula (19) coincides with the known result (2) as given by Algazi [5]. Also for $k=1$, our constrained-entropy minimum distortion formula (24) reduces to the known result (4) due to Gish and Pierce [6].

For $k=2$ we have already noted that the regular hexagon is the optimal polytope. This yields the coefficient of quantization

$$C(2,2) = \frac{5}{36\sqrt{3}}.$$

A theorem by Fejes Toth [12] shows in effect that for a uniformly distributed random variable the minimum distortion for each r is obtained by a tessellation of regular hexagons. Newman [13] independently found a proof of this result for $r=2$. Their results imply that Zador's coefficient $A(2,2)$ has the value $5/(36\sqrt{3})$. Hence the complete solution for nonuniform densities $p(x_1,x_2)$ is in fact given by

$$D \sim \frac{5}{36\sqrt{3}} N^{-1} \left[\int \int \sqrt{p(x_1,x_2)}\, dx_1\, dx_2 \right]^2$$

asymptotically as $N \to \infty$, when $k=2$ and $r=2$. Unknown to Newman and Zador, Fejes Toth [10] had given a complete proof of this result. Hence our minimum distortion formula reduces to this known result for $k=2$ and $r=2$.

Fejes Toth [10] noted that the optimal partition for a given probability density $p(x_1,x_2)$ in the plane consists of "approximately" regular hexagons with the centroids distributed with a nonuniform density over the plane. An example of a hexagonal partition whose centroids are distributed nonuniformly in the plane is shown in Fig. 3. These results for $k=2$ help to clarify the role of the output point density function $\lambda(x)$ in characterizing a quantizer as used in this paper.

For $k \geqslant 3$, the minimum distortion attainable for a quantizer is not known. However we can obtain upper bounds on the quantization coefficient $C(k,r)$ as noted in Section IV by calculating the normalized inertia for any admissible polytope. Any admissible polytope generates a tessellation that can be used for the quantization of a random vector that is uniformly distributed on a unit volume region. Hence neglecting the boundary regions when N is large, the normalized inertia I' of that polytope gives the attainable distortion

$$D = \left(\frac{1}{k} I' \right) N^{-r/k}.$$

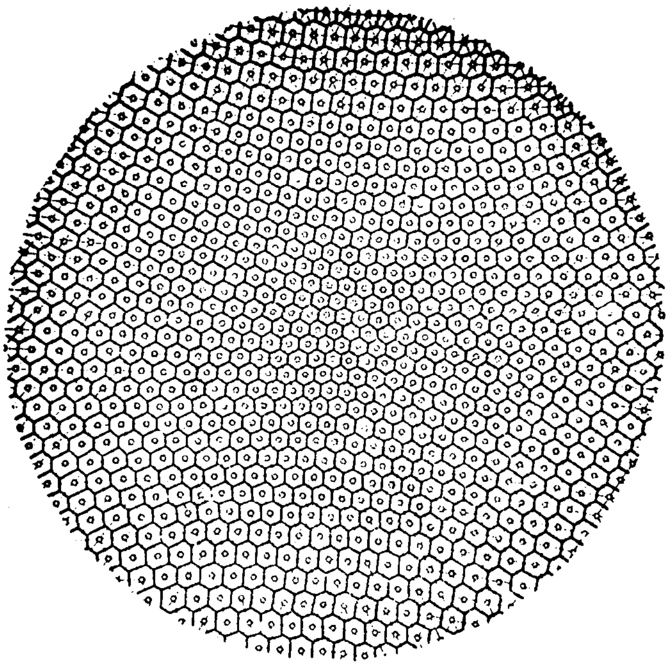

Fig. 3. Hexagonal partition for nonuniform density of points (from Fejes Toth [10]).

Hence

$$A(k,r) \leqslant \frac{1}{k} l'.$$

Therefore any upper bounds we obtain for $C(k,r)$ are in fact upper bounds for Zador's $A(k,r)$. Even though our derivation of (24) is not rigorous, these upper bounds are rigorously valid.

For $k=3$ the admissible polyhedra include the five principal parallelohedra: cube, hexagonal prism, rhombic dodecahedron, elongated dodecahedron, and the truncated octahedron. Of these five, the truncated octahedron shown in Fig. 4 and specified by the set

$$\{(x_1,x_2,x_3): |x_1|+|x_2|+|x_3|<1.5, |x_1|<1, |x_2|<1, |x_3|<1\}$$

has by direct calculation the smallest normalized inertia with

$$l=\frac{19}{64\sqrt[3]{3}}=0.23563\cdots$$

which gives the upper bound

$$C(3,2) \leqslant 0.0785445\cdots$$

which is surprisingly close to the sphere lower bound (0.07697) discussed in Section IV. The truncated octahedron is clearly the best parallelohedron and is very likely the optimal polyhedron.

For $k=4$, the analog of the truncated octahedron is the admissible polytope

$$\{x: |x_1|+|x_2|+|x_3|+|x_4| \leqslant 2; |x_i| \leqslant 1, i=1,2,3,4\}$$

which by a crude Monte Carlo integration gives the bound

$$C(4,2) \leqslant 0.0766\cdots.$$

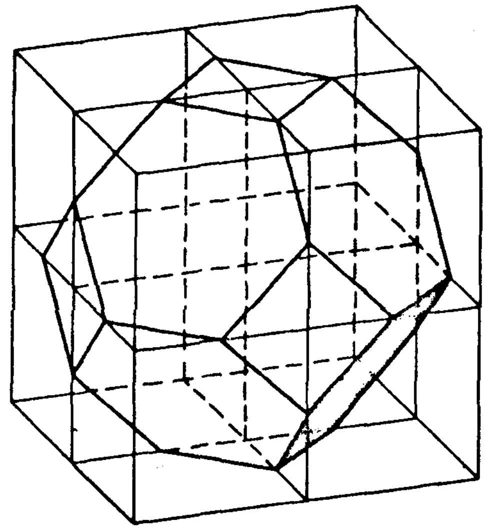

Fig. 4. Truncated octahedron imbedded in cube of side length 2, corresponding to analytical description given in text.

TABLE I
VALUES AND BOUNDS FOR QUANTIZATION COEFFICIENT $C(k,2)$

K	SPHERE LOWER BOUND	ACTUAL VALUE	POLYTOPE UPPER BOUND	CUBE UPPER BOUND	ZADOR UPPER BOUND
1	.08333	.08333(a)		.08333	.5
2	.07958	.08019(b)		.08333	.159
3	.07697		.07854(c)	.08333	.1157
4	.07503		.0766(d)	.08333	.09974
5	.07352		.0779(e)	.08333	.09133
100	.0574		.0766(f)	.08333	.05789

(a) interval
(b) regular hexagon
(c) truncated octahedron
(d) four dimensional analog of (c)
(e) cross product of truncated octahedron with an interval
(f) cross product of 25 type (d) polytopes

One technique for obtaining upper bounds is to select admissible polytopes in a higher dimensional space by forming "cross-products" or prisms using lower dimensional polytopes. For example, the cube in k-space is simply the kth cross-product of the interval, the hexagonal prism in $k=3$ is the cross product of the regular hexagon ($k=2$) with the interval ($k=1$), and the cross product of the regular hexagon with the truncated octahedron gives an admissible polytope for $k=5$. For such cross-product polytopes the normalized moment of inertia when $r=2$ is trivially obtained by summing the normalized moments of inertia of each lower dimensional polytope.

In Table I values of $C(k,2)$ are given when known together with the available lower and upper bounds.

VIII. INFINITE BLOCK LENGTH

All of the preceding results have been based on a fixed (but arbitrary) block length k with $1 \leqslant k < \infty$. The results can be compared with the known performance bounds of rate distortion theory by examining the limiting case for mean-square distortion ($r=2$) letting $k \rightarrow \infty$. Since Zador's upper and lower bounds coincide asymptotically as $k \rightarrow \infty$, we use the lower bound to study the limiting behavior of $D_1(N)$. As noted by Zador, Stirling's formula gives $V_k^{-\beta} \sim (2\pi e)^{-1}k$ as $k \rightarrow \infty$. Zador also conjectured that in general

$$\|p_k(x)\|_{k/k+2} \rightarrow e^{2\bar{H}} \quad \text{as } k \rightarrow \infty \tag{27}$$

where $\bar{H}$ is the differential entropy rate of the source

$$\bar{H} = -\lim_{n\rightarrow\infty} \frac{1}{n} \int p_n(x) \log p_n(x)\, dx$$

and a subscript has been added to the joint density $p(x)$ to identify its dimensionality. In the particular case of a stationary Gaussian process, Zador showed that (27) holds. The result (27) appears to be of some fundamental theoretical interest. A proof that it holds for any stationary ergodic process is given in the Appendix. Without assuming ergodicity, the weaker result

$$\lim_{k\rightarrow\infty} \|p_k\|_{k/k+2} \geqslant e^{2\bar{H}} \tag{28}$$

will now be shown

$$\begin{aligned} \log \|p_k\|_{k/k+2} &= \frac{k+2}{k} \log \int p_k(x) e^{-[2/(k+2)] \log p_k(x)}\, dx \\ &\geqslant \frac{k+2}{k} \int p_k(x) \log e^{-[2/(k+2)] \log p_k(x)}\, dx \\ &= -\frac{2}{k} \int p_k(x) \log p_k(x)\, dx, \end{aligned}$$

using Jensen's inequality. Thus

$$\|p_k\|_{k/k+r} \geqslant e^{2H(p_k)/k} \tag{29}$$

from which (28) follows.

Now using (19) and the lower bound (10) on $C(k,r)$ gives the result for very large block length k that

$$D \geqslant \frac{1}{2\pi e} e^{-2(R-\bar{H})} \tag{30}$$

where $R=(\log N)/k$ is the *rate* or average number of nats per component of X needed to identify (or transmit) the quantizer output approximation $Q(x)$. Inverting (30) gives a more familiar result

$$R \geqslant \bar{H} - \frac{1}{2} \log (2\pi e D), \tag{31}$$

which is the generalized Shannon lower bound on the rate-distortion function (see Berger [14]). Note that (31) as derived here is only valid for the asymptotic case of small distortion (corresponding to large N).

IX. Companding Revisited

For $k=1$, Bennett introduced the "companding" model of a quantizer as a monotonically increasing nonlinear mapping $E(x)$ the compressor, followed by a uniform quantizer, and by the inverse mapping, E^{-1}. Lloyd's point density function $\lambda(x)$ is then the slope of the mapping $E(x)$. For block quantization we have introduced the concept of a point density function and derived a generalization of Bennett's integral without any reference to a mapping. To complete the connection with the one-dimensional case we can define a continuously differentiable and invertible mapping f which maps a point x in R_k into another point $f(x)$ in R_k. Then a family of k dimensional quantizers can be modeled as shown in Fig. 5 as the cascade of such a mapping with an optimal N-point uniform quantizer, let us say on the unit cube in R_k followed by the inverse mapping f^{-1}. The overall quantizer is described by

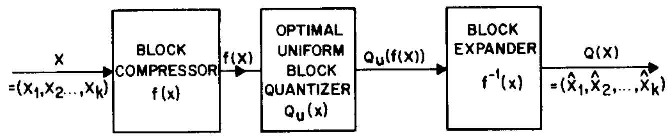

Fig. 5. Companding model of block quantization.

$$Q(x) = f^{-1} \circ Q_u \circ f(x)$$

where $Q_u(x)$ denotes the uniform quantizer. The point density function for the overall quantizer is then given by the Jacobian determinant of f. The Dirichlet partition for the uniform quantizer with regions S_i, will induce a new partition for the overall quantizer with regions $S_i^* = f^{-1}(S_i)$. In general the new partition will not be a Dirichlet partition.

The question then arises, for a given probability density $p(x)$, does there exist a mapping $f(x)$ which makes $Q(x)$ the optimum quantizer? In order to preserve the Dirichlet property it is necessary that the mapping be conformal since the line joining the centroids of two adjacent regions in a Dirichlet partition must always be perpendicular to the hyperplane separating the two regions. For $k=2$, Heppes and Szusz [15] have noted, in effect, that a necessary and sufficient condition for the existence of $f(x)$ is that the logarithm of the point density function be a harmonic function, i.e., that $\log \lambda(x)$ satisfy Laplace's equation. Since for the optimum quantizer we have shown that λ must be proportional to a power of $p(x)$, it follows that the condition is equivalent to having $\log p(x)$ satisfy Laplace's equation. This condition eliminates the joint normal density as well as any other density whose curves of constant density close. Hence there appears to be a fundamental limitation to the possibility of generalizing Bennett's companding approach to the multidimensional case.

X. Concluding Remarks

In this paper we have shown that the point density function of a quantizer, first conceptualized by Lloyd, can be generalized to the multidimensional case to provide a fruitful and intuitively satisfying way to develop the theory of optimal quantization for random vectors. With it we have heuristically generalized the classic integral of Bennett for the distortion of a quantizer which we hope will be useful for future studies of block quantization. In deriving the results of Zador, albeit heuristically, we have gone further toward a constructive theory of optimal quantization by introducing some of the salient geometrical features of the partition of space defined by a quantizer. Finally, we have pointed out the possibilities and limitations of the companding approach to modeling a quantizer for random vectors, an approach that has been of great practical importance in the one-dimensional case.

Acknowledgment

I would like to thank Thomas Liggett of the University of California, Los Angeles, who provided the proof given in the Appendix, Hans Witsenhausen, who introduced me to the work of Fejes Toth, and Neil Sloane, who introduced me to the polyhedra associated with sphere packings.

Appendix[1]

Theorem 1:

$$\lim_{k\to\infty} \|p_k(x)\|_{k/k+r} = e^{r\bar{H}}, \qquad 0<r<r_0$$

if $p_k(x)$ is the joint density of $X^k=(X_1,X_2,\cdots,X_k)$ where $\{X_k\}$ is a stationary ergodic process and $\|p_1(x_1)\|_{1/(1+r_0)}<\infty$.

Proof: Define the random variables

$$Z_k = -\frac{1}{k}\log p_k(x^k).$$

Then

$$\log \|p_k(x)\|_{k/k+r} = \frac{k+r}{k}\log Ee^{[kr/(k+r)]Z_k}. \tag{A.1}$$

From the Shannon-McMillan theorem for abstract alphabets [16], Z_k converges in probability to $\bar{H}$. Hence $e^{a_k Z_k}$ converges in distribution to $e^{r\bar{H}}$ where $a_k=rk/(k+r)$. If also

$$Ee^{dZ_k}<C<\infty \tag{A.2}$$

where C is independent of k, and $r<d$, then convergence in the mean would follow, namely,

$$Ee^{a_kZ_k}\to e^{r\bar{H}} \qquad \text{as } k\to\infty$$

which would complete the proof. To prove (A.2), define

$$\rho_k(X^k) \triangleq \frac{k+r}{r}\log Ee^{a_kZ_k}.$$

Then (A.2) will follow from the fact that $\rho_k(X^k)$ is subadditive, that is,

$$\rho_{k+l}(X^{k+l}) \leqslant \rho_k(X^k)+\rho_l(X^l), \tag{A.3}$$

which is analogous to the subadditivity of the entropy. Denote $s=k+l+r$, $u=(x_1,x_2,\cdots,x_k)$, $v=(x_{k+1},x_{k+2},\ldots,x_{k+l})$, and let $f=p_{k+l}(u,v)$, $p_k=p_k(u)$, $p_l=p_l(v)$. Then

$$\begin{aligned}\int\!\!\int f^{(k+l)/s}\,du\,dv &= \int\!\!\int \left[fp_k^{-r/(k+r)}\right]^{k/s}\left[fp_l^{-r/(l+r)}\right]^{l/s}\cdot\left[p_k^{k/(k+r)}p_l^{l/(l+r)}\right]^{r/s}du\,dv\\ &\leqslant \left[\int\!\!\int fp_k^{-r/(k+r)}\,du\,dv\right]^{k/s}\left[\int\!\!\int fp_l^{-r/(l+r)}\,du\,dv\right]^{l/s}\cdot\left[\int\!\!\int p_k^{k/(k+r)}p_l^{l/(l+r)}\,du\,dv\right]^{r/s}\\ &= \left[\int p_k^{k/(k+r)}\,du\right]^{(k+r)/s}\left[\int p_l^{l/(l+r)}\,dv\right]^{(l+r)/s}.\end{aligned}$$

using the triple Hölder inequality. Hence $k^{-1}\rho_k(X^k)$ is bounded by $\rho_1(X_1)$ so that

$$\frac{k+r}{k}\log\{Ee^{a_kZ_k}\}\leqslant K\log\|p_1(x_1)\|_{1/(1+r_0)}<\infty \tag{A.4}$$

which holds, in particular, for $r=r_0$. Given d, with $0<d<r_0$, for sufficiently large k,

$$d<\frac{kr_0}{k+r_0}.$$

Hence

$$\log Ee^{dZ_k}<\log Ee^{[kr_0/(k+r_0)]Z_k},$$

so that (A.4) implies (A.2), completing the proof of the theorem.

[1]This proof is due to Thomas Liggett.

References

[1] W. R. Bennett, "Spectra of quantized signals," *Bell Syst. Tech. J.*, vol. 27, pp. 446–472, July, 1948.
[2] P. F. Panter and W. Dite, "Quantization in pulse-count modulation with nonuniform spacing of levels," *Proc. IRE*, vol. 39, pp. 44–48, 1951.
[3] S. P. Lloyd, "Least squares quantization in PCM," unpublished memorandum, Bell Laboratories, 1957.
[4] B. Smith, "Instantaneous companding of quantized signals," *Bell Syst. Tech. J.*, vol. 52, pp. 1037–1076, Sept. 1973.
[5] V. R. Algazi, "Useful approximations to optimum quantization," *IEEE Trans. Commun. Tech.*, vol. COM-14, pp. 297–301, 1966.
[6] H. Gish and J. N. Pierce, "Asymptotically efficient quantizing," *IEEE Trans. Inform. Theory*, vol. IT-14, pp. 676–683, Sept. 1968.
[7] P. Zador, "Asymptotic quantization of continuous random variables," unpublished memorandum, Bell Laboratories, 1966.
[8] M. P. Schutzenberger, "On the quantization of finite dimensional messages," *Inform. Contr.*, vol. 1, pp. 153–158, 1958.
[9] P. Elias, "Bounds and asymptotes for the performance of multivariate quantizers," *Ann. Math. Stat.*, vol. 41, pp. 1249–1259, 1970.
[10] L. Fejes Toth, "Sur la representation d'une population infinie par un nombre fini d'éléments," *Acta Math. Acad. Scient. Hung.*, vol. 10, pp. 299–304, 1959.
[11] R. M. Gray and A. H. Gray, Jr., "Asymptotically optimal quantizers," *IEEE Trans. Inform. Theory*, vol. IT-23, pp. 143–144, Jan. 1977.
[12] L. Fejes Toth, *Lagerungen in der Ebene auf der Kugel und im Raum. Springer-Verlag* p. 81, 1953.
[13] D. J. Newman, "The hexagon theorem," unpublished memorandum, Bell Laboratories, 1964.
[14] T. Berger, *Rate-Distortion Theory.* Englewood Cliffs, N.J.: Prentice-Hall, 1971.
[15] A. Heppes and P. Szüsz, "Bemerkung zu einer Arbeit von L. Fejes Toth, " *El. Math.*, vol. 15, pp. 134–136, 1960.
[16] A. Perez, "On the theory of information in the case of an abstract alphabet," in *Trans. First Prague Conf. on Inform. Theory, Statistical Decisions Functions, and Random Processes,* (held at Prague, Nov. 28–30, 1956). Prague: Publishing House of the Czechoslovak Academy of Sciences, 1957, pp. 209–243.
[17] H. L. Loeb, *Space Structures*, Addison-Wesley, 1975.