8/24/94

# Bennett's Integral for Vector Quantizers

Sangsin Na[1]
David L. Neuhoff[2]

ABSTRACT

This paper extends Bennett's integral from scalar to vector quantizers, giving a simple formula that expresses the rth-power distortion of a many-point vector quantizer in terms of the number of points, point density function, inertial profile and the distribution of the source. The inertial profile specifies the normalized moment of inertia of quantization cells as a function of location. The extension is formulated in terms of a sequence of quantizers whose point density and inertial profile approach known functions as the number of points increases. Precise conditions are given for the convergence of distortion (suitably normalized) to Bennett's integral. Previous extensions did not include the inertial profile and, consequently, provided only bounds or applied only to quantizers with congruent cells, such as lattice and optimal quantizers.

The new version of Bennett's integral provides a framework for the analysis of suboptimal structured vector quantizers. It is shown how the loss in performance of such quantizers, relative to optimal unstructured ones, can be decomposed into point density and cell shape losses. As examples, these losses are computed for product quantizers and used to gain further understanding of the performance of scalar quantizers applied to stationary, memoryless sources and of transform codes applied to Gaussian sources with memory. It is shown that the shortcoming of such quantizers is that they must compromise between point density and cell shapes.

INDEX TERMS: Asymptotic quantization theory, high-rate quantization theory, high-resolution quantization theory, moment of inertia, point density, inertial profile, product quantizer, transform code, structured vector quantizers.

[1]Department of Electrical Engineering, Ajou University, Suwon, Korea.
[2]Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI 48109.

Portions of this work were supported by NSF Grant NCR-9105647 and were presented at the IEEE International Symposium on Information Theory, San Diego, 1990.

## I. INTRODUCTION

A k-dimensional N-point vector quantizer (VQ) is principally characterized by a partition $S = \{S_1,...,S_N\}$ of k-dimensional Euclidean space $\Re^k$ into N *quantization cells* and a *codebook* $C = \{y_1, y_2,..., y_N\}$ consisting of N *quantization points* in k-dimensional Euclidean space $\Re^k$. A k-dimensional source vector $x = (x_1,...,x_k)$ is quantized into one of the $y_i$'s according to the *quantization rule* $Q(x) = y_i$ if $x \in S_i$. The quantizer will be denoted (S,C,Q) or simply (S,C) or Q. Its *encoding rate* is $(\log_2 N)/k$. When applied to a random vector $X = (X_1, X_2,..., X_k)$ with probability density $p(x)$, the *rth-power distortion* of (S,C) is

$$D(S,C) = \frac{1}{k} \int \|x - Q(x)\|^r p(x)\, dx = \frac{1}{k} \sum_{i=1}^{N} \int_{S_i} \|x-y_i\|^r p(x)\, dx , \tag{1.1}$$

where $\|x-y\| = \left(\sum_{i=1}^k (x_i-y_i)^2\right)^{1/2}$ denotes Euclidean distance and $p(x)$ is the (k-dimensional) probability density of X.

The pioneering work of Bennett [1] showed that the mean-squared error (r=2) of a scalar quantizer (k=1) with many small cells (N large) and with each $y_i$ in the center of its cell may be accurately approximated as

$$D(S,C) \cong \frac{1}{12N^2} \int \frac{1}{\lambda(x)^2} p(x)\, dx , \tag{1.2}$$

where $\lambda(x)$ is a function, called the *point density*, such that $\lambda(x)\Delta$ is, approximately, the fraction of quantization points in a small interval of width $\Delta$ surrounding x, and where an integral without limits denotes an integral over the entire space. The right-hand side of (1.2) is known as Bennett's integral. Although originally derived for companders (quantizers consisting of a compressor mapping, uniform quantizer and expander mapping) with $\lambda$ equal to the derivative of the compressor function, it was recognized by others [2-5] that (1.2) applies more generally. Besides giving a simple approximate formula for distortion, Bennett's integral shows how the distortion depends on the key characteristics of the quantizer, namely, the number of points N and the point density $\lambda$. Its utility is exemplified by the fact that one may use it to show that the best quantizers have $\lambda(x) \cong p(x)^{1/3} / \int p(x)^{1/3}dx$ and to evaluate the merits of other quantizers.

In this paper we extend Bennett's integral to vector quantizers. The goal is to give a simple approximate formula for distortion that again shows the influence of key characteristics and that can be used to analyze the performance of suboptimal structured vector quantizers. Previous extensions of Bennett's integral and related work are described in the next paragraph. Subsequently, an informal derivation of the principal result is given. Section II describes the application of the new Bennett integral to the analysis of optimal and suboptimal quantizers -- particularly, scalar, product and transform quantizers. Section III presents a rigorous formulation of Bennett's integral, and Section IV presents the proof, with certain details left to the Appendix.

For scalar quantizers, Bennett's integral was extended to rth-power distortion by Algazi [6]. It was extended to vector quantizers with congruent cells (as in a lattice quantizer) by Gersho [5]. Yamada et. al. [7] gave a Bennett-like lower bound to distortion that applies to all vector quantizers and difference distortion measures. Bucklew [8,9] extended Bennett's integral to companders in higher dimensions. All these results were heuristically formulated and derived. Subsequently, Bucklew and Wise [10] gave a rigorous formulation and proof of (1.2) for scalar quantizers. Although they stated their result for companders, when translated to arbitrary quantizers, their approach was to show that if a sequence of scalar quantizers $\{(S_N,C_N)\}_{N=1}^{\infty}$ has increasing numbers of points ($(S_N,C_N)$ has $N$ points) and has point densities approaching a given $\lambda(x)$, then under certain technical conditions on $\lambda$ and the source density $p$,

$$N^r D(S_N,C_N) \rightarrow \frac{1}{12} \int \frac{1}{\lambda(x)^r} p(x)\, dx \quad \text{as } N \rightarrow \infty \,. \tag{1.3}$$

A similar result was given by Cambanis and Gerr [11]. For vector quantizers, Bucklew [12] showed that under certain technical conditions, a form of Bennett's integral holds, whenever the sequence of quantizers is asymptotically optimum for some other source density. Over the years, other work has focused on deriving the asymptotically best performance of quantizers without explicit use of Bennett's integral [13-19]. A thorough summary of early work is contained in [19].

A complete extension of Bennett's integral to vector quantizers must take into account the shape of the quantization cells and, indeed, the possibility that in two or more dimensions the cells may have many different shapes.

In the following we give a heuristic derivation of our extension of Bennett's integral to vector quantizers. A rigorous formulation and proof in terms of sequences of quantizers is postponed to Sections III and IV.

Suppose a quantizer $(S,C)$ in $k$ dimensions has many points, suppose the quantization cells are small and suppose the source density is smooth. Then (1.1) may be approximated as

$$D(S,C) \cong \frac{1}{k} \sum_{i=1}^{N} p(y_i) \int_{S_i} \|x-y_i\|^r\, dx = \sum_{i=1}^{N} p(y_i)\, M(S_i)\, v(S_i)^{1+r/k} \,, \tag{1.4}$$

where $v(S_i)$ denotes the volume of the cell $S_i$ and $M(S_i)$ denotes the *normalized moment of inertia* of the cell $S_i$ about the point $y_i$ with respect to rth-power distortion, namely,

$$M(S_i) = \frac{\int_{S_i} \|x-y_i\|^r \, dx}{k \, v(S_i)^{1+r/k}} . \tag{1.5}$$

This normalization makes $M(S_i)$ invariant to a scaling of $S_i$, and for $r=2$, it makes $M(\text{cube}) = 1/12$ for any $k$.

Suppose further that there is a function $\lambda(x)$ (usually smooth) such that for any $x$, the fraction of quantization points contained in a small region $R$ containing $x$ is approximately $\lambda(x)v(R)$. Such a function is called the *point density* of $(S,C)$ and satisfies

$$\lambda(x) \cong \frac{1}{Nv(S_i)} \quad \text{for } x \in S_i . \tag{1.6}$$

The existence of a point density presumes that neighboring cells have similar volumes, as is often the case.

Finally, suppose there is a function $m(x)$ (usually smooth) such that for any $x$, the cells in the vicinity of $x$ have normalized moment of inertia (about their respective quantization points) approximately equal to $m(x)$. Such a function will be called the *inertial profile* of $(S,C)$. Its existence presumes that neighboring cells have similar normalized moments of inertia, as is the case for many quantizers.

Substituting the point density and inertial profile into (1.4) gives

$$D(S,C) \cong \frac{1}{N^{r/k}} \sum_{i=1}^{N} p(y_i) \frac{m(y_i)}{\lambda(y_i)^{r/k}} \, v(S_i) ,$$

which we recognize as an approximation to a Riemann integral. Therefore,

$$D(S,C) \cong \frac{1}{N^{r/k}} \int p(x) \frac{m(x)}{\lambda(x)^{r/k}} \, dx , \tag{1.7}$$

which is the new version of Bennett's integral for vector quantizers. Note that it is similar to the formula derived in [5] except that instead having a single moment of inertia that applies to all cells, it contains the inertial profile $m(x)$.

Equation (1.7) shows that distortion decreases as $1/N^{r/k}$, with proportionality constant depending on the source density, point density and inertial profile. This identifies the point density and inertial profile as key performance determining characteristics, which one should focus upon when considering the performance of a vector quantizer.

Since the normalized moment of inertia and, consequently $m(x)$, can never be less than that of a k-dimensional sphere, $V_k^{r/k}/(k+r)$, where $V_k$ is the volume of a k-dimensional sphere with radius 1,

$$D(S,C) \gtrapprox \frac{V_k^{r/k}}{N^{r/k}(k+r)} \int p(x) \frac{1}{\lambda(x)^{r/k}} \, dx ,$$

which also follows from the results of [5].

For scalar quantizers, the cells are almost always taken to be intervals, and the quantization points are usually taken to be the midpoints. In this case, $m(x) = 1/12$, assuming $r = 2$, and (1.7) reduces to Bennett's integral for scalar quantizers (1.2). If, on the other hand, the quantization points were placed at the left endpoints of their respective cells, then $m(x) = 1/3$.

## II. Applications of Bennett's Integral

### *A. Optimal and Suboptimal Quantizers*

Two important questions are: What is the least distortion among quantizers with a given dimension and size? And, what are the best point density and inertial profile for a given dimension?

According to Gersho's widely believed conjecture [5], when $N$ is large, most cells of the best k-dimensional quantizers are congruent, approximately, to the tessellating polytope $H_{r,k}^*$ with minimum normalized moment of inertia. Assuming such, the best inertial profile is the constant:

$$m_{r,k}^*(x) \triangleq M(H_{r,k}^*) \triangleq M_{r,k}^* , \tag{2.1}$$

and either calculus of variations or Hölder's inequality can be used to show that (1.7) is minimized by the point density [5]

$$\lambda_{r,k}^*(x) \triangleq \frac{p(x)^{k/(k+r)}}{\int p(x')^{k/(k+r)} \, dx'} . \tag{2.2}$$

Substituting (2.1) and (2.2) into (1.7), one finds that the least distortion of any k-dimensional quantizer with $N$ points is

$$D_{r,k,N}^* \cong M_{r,k}^* \frac{1}{N^{r/k}} \left( \int p(x)^{k/(k+r)} \, dx \right)^{(k+r)/k}, \tag{2.3}$$

which is the well known Zador-Gersho formula. Zador [15,16] derived the form of (2.3); Gersho [5] recognized that the multiplicative constant is $M^*_{r,k}$. See also [10].

Although as just shown, the vector version of Bennett's integral enables a direct derivation of the distortion of the best quantizers (assuming Gersho's conjecture), its real forte is that it provides a framework for understanding the limitations of various suboptimal quantizers. Specifically, since vector quantizers can be quite complex, "structure" is often introduced to simplify the encoding and/or decoding -- for example, scalar, product, lattice, transform, multi-stage, tree, trellis, and pyramid structures. Generally speaking, such structure causes a quantizer to have more distortion than an optimal quantizer with the same size and dimension. Nevertheless, structured quantizers are used when their complexity reduction outweighs their suboptimality. Indeed, their lower complexity may result in less distortion, when it permits a larger dimension than an optimal quantizer with the same rate. Although such structured quantizers are widely used, in many cases their performance is difficult to predict and, consequently, has only been determined empirically. This makes it difficult to understand the true source of their suboptimality, to predict which structures are appropriate for a given situation, and to find better techniques.

It is our claim that the new version of Bennett's integral provides a framework for qualitative and quantitative analyses of the increase in distortion caused by structures such as those mentioned above. Specifically, it suggests that one should focus on the effect of structure on the point density and inertial profile. Indeed, a qualitative comparison of these functions with the ideals given in (2.1) and (2.2) reveals much about the limitations of a specific form of structure. Moreover, if one can explicitly compute (or at least bound or estimate) the point density and inertial profile induced by a particular structure, then one can use Bennett's integral to quantify the resulting loss of performance. Examples of qualitative and quantitative loss analyses include those of tree-structured quantization [20], two-stage quantization [21,22], scalar quantization (Subsection C below) and transform coding (Subsection D below). In the following we show how to separately identify the losses due to the point density and the inertial profile.

The ratio of the distortion of a quantizer (S,C) to that of the best quantizer with the same number of points will be called its *loss* $L = L(S,C)$. Bennett's integral (1.7) and the Zador-Gersho formula (2.3) yield

$$L \triangleq \frac{D(S,C)}{D^*_{r,k,N}} \cong \frac{\dfrac{1}{N^{r/k}} \displaystyle\int \frac{m(x)}{\lambda(x)^{r/k}}\, p(x)\, dx}{\dfrac{1}{N^{r/k}} \displaystyle\int \frac{m^*_{r,k}(x)}{\lambda^*_{r,k}(x)^{r/k}}\, p(x)\, dx} = \frac{\displaystyle\int \frac{m(x)}{\lambda(x)^{r/k}}\, p(x)\, dx}{M^*_{r,k} \left( \displaystyle\int p(x)^{k/(k+r)}\, dx \right)^{(k+r)/k}} .$$

To separate the effects of a suboptimal point density from those of a suboptimal inertial profile, we factor the loss into $L = L_{pt} L_{ce}$, where $L_{pt}$ is the *point density loss*

$$L_{pt} = \frac{\int \frac{1}{\lambda(x)^{r/k}} \, p(x) \, dx}{\int \frac{1}{\lambda^*_{r,k}(x)^{r/k}} \, p(x) \, dx} , \tag{2.4}$$

which is the ratio of the distortion of a hypothetical quantizer with the given point density and a constant (e.g. optimal) inertial profile to that of a quantizer with the optimum point density and the same inertial profile, and $L_{ce}$ is the *cell shape loss*

$$L_{ce} = \frac{\int \frac{m(x)}{\lambda(x)^{r/k}} \, p(x) \, dx}{\int \frac{M^*_{r,k}}{\lambda(x)^{r/k}} \, p(x) \, dx} , \tag{2.5}$$

which is the ratio of the distortion of the quantizer to that of a hypothetical quantizer with the same point density, but an optimum inertial profile. One may also write

$$L_{ce} = \int \frac{m(x)}{M^*_{r,k}} \, q(x) \, dx ,$$

where

$$q(x) = \frac{p(x)\lambda(x)^{-r/k}}{\int p(x')\lambda(x')^{-r/k} dx'} ,$$

which shows that $L_{ce}$ is an average of "local" cell shape loss.

Both point density and cell shape losses are greater than or equal to one -- the former by the Hölder inequality used in the derivation of $\lambda^*_{r,k}$, the latter by Gersho's conjecture. They serve to focus attention on the effects of a suboptimal point density and inertial profile. When they can be computed, they show the precise quantitative effect of structure on distortion. One interesting case where they can be explicitly computed is the case of product quantizers. Doing so leads to new insight into scalar and transform quantizers, even though they have been extensively studied in the past. This is discussed next.

*B. Product Quantizers*

The *product* of $k$ scalar quantizers $(S_i,C_i,Q_i)$, $i = 1, \dots, k$, is the k-dimensional vector quantizer $(S^{pr},C^{pr},Q^{pr})$ formed by using each scalar quantizer in succession.[1] That is, if $(S_i,C_i,Q_i)$ has $N_i$ points, then the product VQ has product quantization rule $Q^{pr}(x_1,\dots,x_k) = (Q_1(x_1),\dots, Q_k(x_k))$, product codebook $C^{pr} = C_1 \times \dots \times C_k = \{ (u_1,\dots,u_k) : u_i \in C_i, i = 1, \dots, k \}$ with $N^{pr} = N_1 \cdots N_k$ points, and product partition $S^{pr} = S_1 \times \cdots \times S_k$, where the cell corresponding to the point $(u_1,\dots,u_k)$ is $S_{1,u_1} \times \dots \times S_{k,u_k}$. The product of $k$ identical copies of a scalar quantizer will be called its k*th power*.

As an example, a scalar quantizer and its second power are shown in Figure 1. The former is the optimal (with respect to squared error) 16-point scalar quantizer for a Gaussian density with zero mean and unit variance (cf. [24, p. 134]).

From now through the rest of Section 2, we assume the $N_i$'s are large enough that the scalar quantizers have well-defined point densities $\lambda_1(x_1), \dots, \lambda_k(x_k)$, and that Bennett's integral is accurate. We also assume squared error distortion, i.e. $r = 2$, and we omit the subscript $r$ in $m^*_{r,k}(x)$, $\lambda^*_{r,k}(x)$, $M^*_{r,k}$, $H^*_{r,k}$, etc.. The cell of the product quantizer containing $x = (x_1,\dots,x_k)$ is a rectangle with side lengths approximately $(N_1\lambda_1(x_1))^{-1}, \dots, (N_k\lambda_k(x_k))^{-1}$. It follows from this and (1.6) that the product quantizer has product point density

$$\lambda^{pr}(x) = \lambda_1(x_1) \cdots \lambda_k(x_k) . \tag{2.6}$$

Moreover, assuming the scalar quantization points are centered in their respective cells, the product quantizer has inertial profile

$$m^{pr}(x) = \frac{1}{12} \frac{\frac{1}{k} \sum_{i=1}^{k} \frac{1}{N_i^2 \lambda_i(x_i)^2}}{\left( \prod_{i=1}^{k} \frac{1}{N_i^2 \lambda_i(x_i)^2} \right)^{1/k}} = \frac{1}{12} \frac{1}{k} \sum_{i=1}^{k} \frac{1}{N_i^2 \lambda_i(x_i)^2} \left(N^{pr} \lambda^{pr}(x)\right)^{2/k} , \tag{2.7}$$

where the first equality uses the fact that the normalized moment of inertia of a rectangle (about its center) is $1/12$ times the ratio of the arithmetic and geometric means of the squares of the side lengths, and where the second inequality uses (2.6).[2] Notice that both $\lambda^{pr}$ and $m^{pr}$ are affected by the scalar point densities, but only the latter is affected by the sizes.

---

[1]There are more general notions of product quantizers than that considered here (cf. [23], p. 430 ff.).

[2]In the terminology of the next section, (2.6) and (2.7) hold for both the "specific" and "model" versions of point density and inertial profile.

Suppose that the product quantizer is applied to a random vector $X = (X_1,\ldots,X_k)$ with joint density $p(x)$. Substituting (2.6) and (2.7) into the k-dimensional version of Bennett's integral (1.7) and simplifying shows that product quantizer has distortion

$$D^{pr} = \frac{1}{k} \sum_{i=1}^{k} \frac{1}{12\,N_i^2} \int \frac{1}{\lambda_i(x_i)^2}\, p_i(x_i)\, dx_i\, ,$$

where $p_i(x_i)$ is the marginal density of $X_i$. The above is just the average of the 1-dimensional version of Bennett's integral (1.7) applied to each scalar quantizer. Obviously, it also follows directly from the fact that the distortion of the product quantizer is just the average of the distortions of the scalar quantizers. The real value of computing the point density (2.6) and the inertial profile (2.7) is to get better qualitative and quantitative understanding of the performance limitations of systems whose performance can be related to those of product quantizers.

*C. Scalar Quantizers*

We now address the frequently asked question: Why do vector quantizers outperform scalar quantizers for stationary, memoryless sources? The best answer to date is given by Lookabough and Gray [25]. By examining the point density and cell shape losses of kth power quantizers, we are able to add further insight.

To do so, consider a stationary, memoryless source $\{X_k\}$ with first-order density $p_1(x_1)$ and k-dimensional joint density $p(x) = p_1(x_1)\cdots p_1(x_k)$. We follow the approach of [25] and compare scalar quantization to vector quantization by comparing the features of kth-power quantizers to those of the best k-dimensional quantizers. (A kth-power quantizer is just an alternate description of a scalar quantizer.) Let $N_1$ denote the size of a scalar quantizer, and let $\lambda_1$ denote its point density. Its kth-power has $N_1^k$ points, point density

$$\lambda^{pw}(x) = \lambda_1(x_1)\cdots\lambda_1(x_k)\, ,$$

and inertial profile

$$m^{pw}(x) = \frac{1}{12} \frac{\frac{1}{k}\sum_{i=1}^{k} \frac{1}{\lambda_1(x_i)^2}}{\left(\prod_{i=1}^{k} \frac{1}{\lambda_1(x_i)^2}\right)^{1/k}} = \frac{1}{12}\,\frac{1}{k}\sum_{i=1}^{k} \frac{1}{\lambda_1(x_i)^2}\, \lambda^{pw}(x)\, ,$$

neither of which depends on $N_1$. The cells of the kth-power quantizer are cubic along the diagonal and wherever else $\lambda_1(x_1) = \lambda_1(x_2) = \ldots = \lambda_1(x_k)$. They are entirely cubes (except, possibly, for infinite volume cells on the periphery) if and only if the quantizer is uniform. Hence,

$$m^{pw}(x) \;\geq\; \frac{1}{12} \;, \tag{2.8}$$

with equality for all $x$ if and only if $\lambda_1$ (equivalently, the scalar quantizer) is uniform.

In comparison, an optimal k-dimensional quantizer has point density

$$\lambda_k^*(x) \;=\; c\, p(x)^{k/(k+2)} \;=\; c\, p_1(x_1)^{k/(k+2)} \cdots p_1(x_k)^{k/(k+2)} \,,$$

where in the above and the following, $c$ denotes whatever normalization constant makes the density integrate to one[3]. Its inertial profile is

$$m_k^*(x) \;=\; M_k^* \,.$$

As a result, the kth-power quantizer has point density loss

$$L_{pt} \;=\; \frac{\displaystyle\int \prod_{i=1}^{k} \frac{1}{\lambda_1(x_i)^{2/k}} \; p(x)\, dx}{\displaystyle\int \frac{1}{\lambda_k^*(x)^{2/k}} \; p(x)\; dx} \;=\; \frac{\left(\displaystyle\int \frac{1}{\lambda_1(x_1)^{2/k}} \; p_1(x_1)\, dx_1\right)^k}{\left(\displaystyle\int p_1(x_1)^{k/(k+2)}\, dx_1\right)^{k+2}} \,,$$

and cell shape loss

$$L_{ce} \;=\; \frac{1/12}{M_k^*} \; \frac{\displaystyle\int \frac{1}{\lambda_1(x_1)^2} \; p_1(x_1)\, dx_1}{\left(\displaystyle\int \frac{1}{\lambda_1(x_1)^{2/k}} \; p_1(x_1)\, dx_1\right)^k} \,.$$

It is useful to factor $L_{ce}$ into $L_{cu}\, L_{ob}$, whose first term is the *cubic loss*

$$L_{cu} \;\triangleq\; \frac{M_1^*}{M_k^*} \;=\; \frac{M(\text{cube})}{M(H_k^*)} \;=\; \frac{1/12}{M_k^*} \,,$$

and whose second term is the *oblongitis loss*

$$L_{ob} \;\triangleq\; \frac{\displaystyle\int \frac{1}{\lambda_1(x_1)^2} \; p_1(x_1)\, dx_1}{\left(\displaystyle\int \frac{1}{\lambda_1(x_1)^{2/k}} \; p_1(x_1)\, dx_1\right)^k} \;=\; \frac{\displaystyle\int \frac{m^{pw}(x)}{\lambda^{pw}(x)^{2/k}} \; p(x)\, dx}{\displaystyle\int \frac{1/12}{\lambda^{pw}(x)^{2/k}} \; p(x)\, dx} \;=\; \int \frac{m^{pw}(x)}{1/12} \; q(x)\, dx \,. \tag{2.9}$$

where

[3]The "c" in one expression is not necessarily the same as the "c" in another.

$$q(x) = \frac{p(x)\,\lambda^{pw}(x)^{-2/k}}{\int p(x')\lambda^{pw}(x')^{-2/k} dx'} .$$

The cubic loss is the cell shape loss that would occur if the cells were entirely cubes; i.e., it is the smallest possible cell shape loss for kth-power quantizers. Lookabough and Gray [25] called it the *space filling* loss[4]. It has been conjectured that a cube has the smallest n.m.i. of any polytope with $2k$ faces (or, at least, of any such tessellating polytope) [26]. In this light, one may view $L_{cu}$ to be a consequence of the fact that the cells of a kth-power quantizer have fewer faces than those of an optimum quantizer. The oblongitis loss is the ratio of the distortion of the kth-power quantizer to that of a hypothetical quantizer with the same point density, but cubic cells. In other words, it is the loss due to the rectangular (as opposed to cubic) nature of the cells. From (2.8), we have $L_{ob} \geq 1$, with equality if and only if $\lambda_1$ is uniform on the support of $p_1$.

To optimize the kth-power quantizer (equivalently, the scalar quantizer), the scalar point density $\lambda_1$, which affects both its point density and inertial profile, must be chosen to minimize the product $L_{pt}\, L_{ob}$. ($L_{ob}$ is the part of $L_{ce}$ that depends on $\lambda_1$.) On the one hand, choosing $\lambda_1$ to be uniform minimizes $L_{ob}$.[5] On the other hand, choosing $\lambda_1(x_1) = c\, p_1(x_1)^{k/(k+2)}$ minimizes $L_{pt}$. In this case, $\lambda^{pw} = \lambda_k^*$, but there is so much "oblongitis" that $L_{ob} = \infty$, as one may easily check.[6] The best scalar point density, $\lambda_1^*(x_1) = c\, p_1(x_1)^{1/3}$, is a compromise. It is more uniform than the point density that minimizes $L_{pt}$, which reduces "oblongitis".

The fact that power quantizers can have the optimal k-dimensional point density is often overlooked. To see why, consider a scalar quantizer with a point density of the form $\lambda_1(x_1) = c\exp\{-bx_1^2\}$ for $X_1$ having a Gaussian distribution with zero mean and unit variance. Its kth power has a spherically symmetric, bell-shaped point density of the form $c\exp\{-b\|x\|^2\}$ a fact that is disguised by the "squarish" appearance of Figure 1. If one chooses $b = k/2(k+2)$, then the kth-power point density is the optimal k-dimensional point density, which as mentioned previously causes $L_{ob} = \infty$. The optimal scalar point density, which has $b = 1/6$, generates a broader, lower "bell".

For the optimal scalar point density, the point density and oblongitis losses are

---

[4]Actually, it was called an "advantage" due to vector quantization, as opposed to a "loss" due to scalar quantization.
[5]If the support of $p_1$ is finite, then choosing $\lambda_1$ to be uniform on it and zero elsewhere makes $L_{ob} = 1$, the minimum possible value. If the support of $p_1$ is infinite, then choosing $\lambda_1$ to be uniform on a sufficiently large subset and zero elsewhere makes $L_{ob}$ arbitrarily close to 1.
[6]For this point density, the distortion does not decrease as rapidly as $1/N^2$.

$$L^*_{pt} \triangleq \frac{\left(\int p_1(x_1)^{(3k-2)/3k}\, dx_1\right)^k \left(\int p_1(x_1)^{1/3}\, dx_1\right)^2}{\left(\int p_1(x)^{k/(k+2)}\, dx_1\right)^{k+2}} \quad . \tag{2.10}$$

$$L^*_{ob} \triangleq \frac{\int p_1(x_1)^{1/3}\, dx_1}{\left(\int p_1(x_1)^{(3k-2)/3k}\, dx_1\right)^k} \tag{2.11}$$

The product of these has been called the *shape loss*[4] [25], because it is due to the "shape" of $p_1$ -- for a more uniformly shaped $p_1$, the optimal scalar point density $c\, p_1(x_1)^{1/3}$ is more uniform and, consequently, less of a compromise from the scalar point density $c\, p_1(x_1)^{k/(k+2)}$ that generates the best k-dimensional point density. Consequently, both $L^*_{ob}$ and $L^*_{pt}$ are smaller. Indeed, for a uniform density, both are 1.

For a Gaussian density, straightforward computations shows

$$L^*_{pt} = 3\left(\frac{3k}{3k-2}\right)^{k/2}\left(\frac{k}{k+2}\right)^{(k+2)/2} \rightarrow 3\ e^{-2/3} \quad \text{as } k \rightarrow \infty\, , \tag{2.12}$$

$$L^*_{ob} = \sqrt{3}\left(\frac{3k-2}{3k}\right)^{k/2} \rightarrow \sqrt{3}\, e^{-1/3} \quad \text{as } k \rightarrow \infty\, , \tag{2.13}$$

which are computed in Table 1, along with $L_{cu}$, for a number of values of k. For a Laplacian density, the point density and cubic losses, also given in Table 1, work out to be the squares of those for the Gaussian density. They are larger than for the Gaussian density because the sharper peak at the origin and the heavier tail means that a good scalar quantizer must be more nonuniform. This causes more oblongitis, which in turn causes more compromising of the optimal point density in order to reduce the oblongitis.

In summary, for memoryless sources, the shortcoming of scalar quantizers (equivalently, kth-power quantizers) is not, as is sometimes believed, that they are incapable of putting points in the best places (i.e., forming the optimum point density). Rather their structure is such that they are incapable of producing cells with shapes better than cubes and such that better point densities come at the expense of lesser cell shapes. Consequently, the best scalar point density is a compromise, with neither the best point k-dimensional density nor the best k-dimensional cell shapes. The decomposition into point density and cell shape losses shows quantitatively the effects of this compromise. For sources with memory, there is also a loss due to the point density not reflecting the correlation. This is well expressed by the *memory loss* quantified in [25].

*D. Transform Codes*

Transform codes are a common form of structured vector quantizer where the source vector $X = (X_1,\ldots,X_k)$ is multiplied by a $k \times k$ orthogonal matrix $T$ (the "transform") before quantization by a product quantizer.[7] (The inverse transform is applied at the decoder.) Although the performance of such codes is well understood in the high-rate case [28], we obtain additional insight by reexamining them from the point density, cell shape viewpoint. As usual in high-rate analyses of transform coding, we restrict attention to stationary Gaussian sources, and for the purpose of illustration, we presume the source to have zero mean, unit variance, and correlation coefficient $\rho = E(X_1X_2)$ close to 1, so that its k-dimensional density is an elliptically shaped hill, with principal axis on the diagonal.

The motivation for transform codes is that one would like to have vector quantizers with the simplicity of scalar quantizers and the performance of optimal ones. We now describe from a point density, cell shape viewpoint how transform coding approaches these goals. Viewed in k-dimensions, a scalar quantizer with a point density of the form $c\exp\{-bx_1^2\}$ forms a power quantizer whose point density is a spherical hill of the form $c\exp\{-b\|x\|^2\|\}$, which does not match the diagonally oriented source density. By themselves, product quantizers offer no improvement, because the best product quantizer for a stationary source is a power quantizer. However, we note that product quantizers can have elliptically shaped point densities of the form $c\ \exp\{-\sum_{i=1}^k b_i\, x_i^2\}$. Although the axes of the elliptical contours of such a point density are aligned with the Cartesian axes, a rotation of the product quantizer could bring the principal axis into alignment with the diagonal, which would then match the source density. (Secondary axes could be similarly aligned with their counterparts in the source density.) Such a rotation can be accomplished by rotating the source vector with an orthogonal transformation prior to quantization with an "elliptical" product quantizer -- in other words, by transform coding. Although this yields a good point density, it is not immediately clear what happens to the cell shapes. Note, for example, that a power quantizer has cubic cells along the diagonal where the source density is concentrated; this is good. Neither rotation nor the use of distinct scalar quantizers (as in product codes) are advantageous from the standpoint of cell shapes. On the other hand, from

[7] A product of vector quantizers may also be used, but for simplicity we restrict attention to the product of scalar quantizers. Generalizations are straightforward.

(2.6) and (2.7), we see that the "size allocations" $N_1,\ldots,N_k$ affect the cell shapes, but not the point density of the product quantizer. (The overall size, $N^{tr} = N_1 \cdots N_k$, is assumed to be fixed.) Thus, it might be possible to mitigate the bad effects on cell shape of the rotation and the use of distinct scalar quantizers. We now examine the point density and cell shape losses in detail and discuss the optimization of transform codes from this viewpoint.

Consider a k-dimensional transform code with orthogonal transformation $T$ and with product quantizer characterized by $\lambda_1, \ldots, \lambda_k$, sizes $N_1,\ldots,N_k$ and points centered in their respective cells. Let $Y = (Y_1,\ldots,Y_k) = TX$, with $X$ and $Y$ considered to be column vectors. Its point density and inertial profile are

$$\lambda^{tr}(x) \;=\; \lambda^{pr}(Tx) \quad \text{and} \quad m^{tr}(x) \;=\; m^{pr}(Tx), \tag{2.14}$$

where $\lambda^{pr}$ and $m^{pr}$ are given by (2.6) and (2.7), respectively. Notice that the latter depends on the scalar sizes and point densities, whereas the former depends only on the scalar point densities. One may easily show that for the transform code applied to $X$, the various losses are identical to those of the product quantizer applied to $Y$; i.e.

$$L^{tr}(X) \;=\; L^{pr}(Y), \quad L^{tr}_{pt}(X) \;=\; L^{pr}_{pt}(Y), \quad L^{tr}_{ce}(X) \;=\; L^{pr}_{ce}(Y)\,, \tag{2.15}$$

where we have added "X" and "Y" to emphasize the vector to which the quantizer is applied.

Let us first consider how to choose the transform and scalar point densities to minimize point density loss. Using (2.4), (2.6) and (2.15), the point density loss is

$$L^{tr}_{pt}(X) \;=\; L^{pr}_{pt}(Y) \;=\; \frac{\displaystyle\int \prod_{i=1}^{k} \frac{1}{\lambda_i(y_i)^{2/k}}\; p_Y(y)\,dy}{\displaystyle\int \frac{1}{\lambda^{*}_{Y,k}(y)^{2/k}}\; p_Y(y)\;dy} \;=\; \frac{\displaystyle\int \prod_{i=1}^{k} \frac{1}{\lambda_i(y_i)^{2/k}}\; p_Y(y)\,dy}{\left(\displaystyle\int p_Y(y)^{k/(k+2)}\,dy\right)^{(k+2)/k}}\;.$$

Notice that if one chooses $T$ so that the components of $Y$ are independent, then choosing $\lambda_i(y_i) = c\,p_i(y_i)^{k/(k+2)}$, $i = 1,\ldots,k$, makes the point density loss equal to 1. That is, the product and transform quantizers have the point density of an optimum k-dimensional vector quantizer for $Y$ and $X$, respectively, and point density loss is minimized. Such a $T$ is the familiar Karhunen-Loeve transform.

Next, let us consider how to choose the scalar point densities and sizes to minimize cell shape loss. Using (2.5), (2.7) and (2.15), we have

$$L^{tr}_{ce} \;=\; L^{tr}_{cu}\,L^{tr}_{ob} \;=\; \frac{1/12}{M^{*}_{k}}\; L^{tr}_{ob}\;,$$

where $L_{cu}^{tr}$ is the cubic loss and $L_{ob}^{tr}$ is the oblongitis loss:

$$L_{ob}^{tr} = L_{ob}^{pr} = \frac{\frac{1}{k}\sum_{i=1}^{k} \frac{1}{N_i^2} \int \frac{1}{\lambda_i(y_i)^2}\, p_i(y_i)\, dy_i}{\prod_{i=1}^{k} \frac{1}{N_i^{2/k}} \int \prod_{i=1}^{k} \frac{1}{\lambda_i(y_i)^{2/k}}\, p_Y(y)\, dy} .$$

where $p_i(y_i)$ denotes the density of $Y_i$. This can be minimized by choosing each scalar quantizer to be uniform with the same step size. Specifically, $L_{ob}^{pr} \cong 1$ when the $\lambda_i$'s and $N_i$'s are chosen so that $\lambda_i$ is uniform on a sufficiently large interval $[a_i,b_i]$ and zero elsewhere, and $N_1\lambda_1(y_1) = N_2\lambda_2(y_2) = \ldots = N_k\lambda_k(y_k)$, when $a_i \le y_i \le b_i$, $i = 1, \ldots, k$. In contrast, the choice that minimizes point density loss is easily shown to cause $L_{ob} = \infty$, just as with power quantizers.

The best transform code, as found in [28], is a compromise, which we now summarize. Given an arbitrary orthogonal matrix $T$, the scalar quantizers should be chosen so that the resulting product quantizer is optimum for $Y$. Specifically, for $i = 1,\ldots,k$, $\lambda_i(y_i) = c\, p_i(y_i)^{1/3}$ and

$$N_i = N^{1/k} \|p_i\|_{1/3}^{1/2} \left(\prod_{j=1}^{k} \|p_j\|_{1/3}^{1/2}\right)^{-1/k} = N^{1/k} \sigma_i \left(\prod_{j=1}^{k} \sigma_i\right)^{-1/k} , \tag{2.16}$$

where $\|p_i\|_a = \left(\int p_i(y_i)^a\, dy_i\right)^{1/a}$, where $\sigma_i^2$ is the variance of $Y_i$, and where the second equality uses the fact that the $p_i$'s differ only by a scaling. The resulting distortion is proportional to $\left(\prod_{j=1}^{k} \sigma_j^2\right)^{1/k}$; the transform that minimizes this is the Karhunen-Loeve; and the resulting $\sigma_i^2$'s are the eigenvalues of the covariance matrix of $X$. Thus, relative to the transform code with optimum point density, one need not compromise the transformation $T$, but one must compromise the point density in the same sort of way as for kth-power quantizers and memoryless sources.

Let us now examine the cell shapes of the product quantizer in the best transform code. (The latter's cells are just a rotation of these.) Such cells are rectangular. Moreover, they are cubic where and only where $N_1\lambda_1(y_1) = N_2\lambda_2(y_2) = \ldots = N_k\lambda_k(y_k)$, which works out to be at all points of the form $y = c\,(\pm\sigma_1, \pm\sigma_2, \ldots, \pm\sigma_k)$. (Recall that the $\sigma_i^2$'s are the eigenvalues of the covariance matrix of $X$.) These points form rays from the origin into each of the $2^k$ "quadrants". Assuming, without loss of generality, that $\sigma_1 \ge \sigma_2 \ge \ldots \ge \sigma_k$, the density of $Y$ concentrates along the first axis, which is not where the cubes are. However, when $\rho \cong 1$, $\sigma_1$ is large compared to the other $\sigma_i$'s, so the rays of cubes are close to the first axis, which is good.

For the optimum transform code, the resulting point density and oblongitis losses work out to be the geometric means of those of the individual scalar quantizers applied to IID sources. Since all components are Gaussian, these are just the losses of a single scalar quantizer for an IID source. Thus, the losses are exactly as given by (2.10)-(2.13) and Table 1.

On the one hand, it seems fitting that the optimum transform code should have the same losses as scalar quantization, since it is based on scalar quantization. But on the other hand, the point density and cell shapes of the transform code are not simply rotated versions of those of the optimal power quantizer for an IID source, so it is not immediately clear why this should be the case. To see concretely why these losses are identical, let us compare the optimum product quantizer with $N^k$ points for $Y$, denoted $Q_Y^{pr}$, to the optimum power quantizer, denoted $Q_Z^{pw}$, with the same number of points for an IID Gaussian random vector $Z$ whose $k$ components have zero mean and unit variance. To begin, observe that $Y$ has the same probability distribution as $\sigma \circ Z$, where $\sigma \triangleq (\sigma_1,\ldots,\sigma_k)$ is a vector consisting of the variances of the $Y_i$'s (which average to 1), and "$x \circ y$" denotes the vector formed by multiplying corresponding terms of vectors $x$ and $y$. Consequently, $p_Y(y) = p_Z(y/\sigma)/\Gamma$, where $\Gamma \triangleq \Pi_{i=1}^k \sigma_i$ and "$x/y$" denotes the vector formed by dividing corresponding terms of $x$ and $y$. Next, for any $z$, the cell containing $z$ in the optimum power quantizer for $Z$ is a rectangle of the form

$$S_Z(z) \;=\; \frac{1}{N\lambda^*_{Z_1}(z_1)} \times \frac{1}{N\lambda^*_{Z_1}(z_2)} \times \ldots \times \frac{1}{N\lambda^*_{Z_1}(z_k)} \;,$$

where $\lambda^*_{Z_1}(z_1)$ is the optimum scalar point density for $Z_1$. On the other hand, the cell containing the point $y = \sigma \circ z$ in the optimum product quantizer for $Y$ is a rectangle of the form

$$S_Y(y) \;=\; \frac{1}{N_1\lambda^*_{Y_1}(y_1)} \times \frac{1}{N_2\lambda^*_{Y_2}(y_2)} \times \ldots \times \frac{1}{N_k\lambda^*_{Y_k}(y_k)} \;. \tag{2.17}$$

where $\lambda^*_{Y_i}(y_i)$ is the optimum scalar point density for $Y_i$, and the $N_i$'s are the optimum size allocations given by (2.16). Now, since each $Y_i$ has the same distribution as $\sigma_i Z_i$, we have $\lambda^*_{Y_i}(y_i) = \lambda^*_{Z_1}(y_i/\sigma_i)/\sigma_i$. Substituting this relation and the optimum $N_i$'s into (2.17), yields

$$S_Y(y) \;=\; \Gamma^{1/k} S_Z(z) \;,$$

That is, the cell of $Q_Y^{pr}$ for $\sigma \circ z$ is just a scaled (but not distorted) version of the cell of $Q_Z^{pw}$ for $z$. Thus, there is a correspondence between the cells of $Q_Y^{pr}$ and those of $Q_Z^{pw}$. Besides having the same shape, corresponding cells have the same probability:

$$\Pr(Y\in S_Y(y)) \;\cong\; p_Y(y)\, v(S_Y(y)) \;=\; p_Z(y/\sigma)\,\frac{1}{\Gamma}\,\Gamma\, v(S_Z(z)) \;\cong\; \Pr(Z\in S_Z(y)) \,.$$

This correspondence between identically shaped cells with the same probability is what accounts for the sameness of the losses. Specifically, it implies

$$\lambda_Y^{pr}(y) \;=\; \tfrac{1}{\Gamma}\,\lambda_Z^{pw}(y/\sigma) \tag{2.18}$$

$$m_Y^{pr}(y) \;=\; m_Z^{pw}(y/\sigma)\,, \tag{2.19}$$

Equation (2.18) and the fact that $p_Y(y) = p_Z(y/\sigma)/\Gamma$, imply that the point density losses are the same. (Make substitutions into $L_{pt}(Y)$ and change variables to obtain $L_{pt}(Z)$). Equation (2.19), (2.9) and the easily verified fact that $q_Y(y) = q_Z(y/\sigma)/\Gamma$, imply that the oblongitis losses are the same. (Again make substitutions and change variables.)

Finally, we note that the scale factor $\Gamma$ is less than or equal to 1, because $\Gamma^2 = \left(\Pi_{i=1}^k \sigma_i^2\right)^{1/k} \le \frac{1}{k}\Sigma_{i=1}^k \sigma_i^2 = 1$, with equality if and only if the $\sigma_i$'s are all the same. Thus the cells of $Q_Y^{pr}$ are generally smaller (by the factor $\Gamma$) than the corresponding cells of $Q_Z^{pw}$. This is due to the fact that the distribution of Y (and X) is more compact than that of Z. And it accounts for why the product code for Y (equivalently, transform code for X) has distortion that is smaller than the distortion of the scalar quantizer for Z by the factor $\Gamma^{2/k}$. To see this, we argue that the contribution to distortion of the cell containing $y = \sigma \circ z$ of $Q_Y^{pr}$ is smaller than that for the cell containing z of $Q_Z^{pw}$ by the factor $\Gamma^{2/k}$

$$\begin{aligned}
\frac{1}{k}\int_{S_Y(y)} \|y'-Q_Y^{pr}(y')\|^2\, p_Y(y')\, dy' \;&\cong\; p_Y(y)\, M(S_Y)\, v(S_Y(y))^{1+2/k} \\
&\cong\; \Pr(Y\in S_Y(y))\, M(S_Y)\, v(S_Y(y))^{2/k} \\
&\cong\; \Pr(X\in S_Z(z))\, M(S_Z)\, v(S_Z(z))^{2/k}\, \Gamma^{2/k} \\
&\cong\; \Gamma^{2/k}\, \frac{1}{k}\int_{S_Z(z)} \|z'-Q_Z^{pw}(z')\|^2\, p_Z(z')\, dz' \;.
\end{aligned}$$

In summary, transform codes for a Gaussian source with memory suffer precisely the same point density, cubic and oblongitis losses as scalar quantizers for stationary, memoryless Gaussian sources.

## III. Formal Statement of Bennett's Integral

We begin by introducing and motivating the "sequence approach" to Bennett's integral. The result we wish to derive is that if a quantizer has, approximately, point density $\lambda$ and inertial profile $m$, then (1.7) holds. To formalize such a result, we need to identify two notions of point density and two notions of inertial profile.

The *specific point density* of a quantizer $(S,C)$ is the function

$$\lambda_S(x) \;=\; \frac{1}{N v(S_i)} \;, \quad \text{for } x \in S_i \,,$$

that characterizes the density of quantization points for this "specific" quantizer. The function $\lambda_S$ is piecewise constant, non-negative, and zero only where $S_i$ has infinite volume. We will assume that quantization cells have nonzero volumes, so $\lambda_S$ is always finite. Its integral over the union of $n$ bounded quantization cells is $n/N$; its integral over a convex set that is large relative to the quantization cells it intersects is approximately the number of quantization cells it contains divided by $N$; and its integral over all space is the fraction of cells having finite volume. As the subscript $S$ suggests, $\lambda_S$ actually depends on the cells in $S$, rather than the points in $C$. However, since ordinarily each point is contained within its corresponding cell, the term "point" density seems reasonable. On the other hand, a *model point density* is a (usually smooth) function $\lambda(x)$, independent of $N$, that specific point densities approximate or aspire to. For example, $\lambda^*_{r,k}$ defined in (2.2) is a model point density. It is not the specific point density of any specific quantizer, but the best quantizers have $\lambda_S \cong \lambda^*_{r,k}$. We will not try to identify what functions can be model point densities, except to say that they must be nonnegative and integrate to at most one.

The *specific inertial profile* of a quantizer $(S,C)$ is the function

$$m_{S,C}(x) \;=\; M(S_i) \,, \quad \text{for } x \in S_i \,,$$

that characterizes the normalized moment of inertia of the cells in $S$ about the corresponding points in $C$. The function $m_{S,C}$ is piecewise constant, non-negative, and bounded from below by the normalized moment of inertia of a k-dimensional sphere. If $x$ is contained in a cell $S_i$ having infinite volume, then (1.5) gives $M(S_i) = \infty/\infty$. We resolve this by adopting the convention that $m_{S,C}(x) = 0$ for $x$ in such $S_i$. A *model inertial profile* is a (usually smooth) function $m(x)$, independent of $N$, that specific inertial profiles approximate or aspire to. For example, $m^*_{r,k}$ defined in (2.1) is a model inertial profile. We will not try to identify what functions can be model inertial profiles, except to say that they must be bounded from below by the normalized moment of inertia of a k-dimensional sphere.

Returning now to Bennett's integral, one could formalize it by showing that if a quantizer (S,C) has $\lambda_S \cong \lambda$ and $m_{S,C} \cong m$, then (1.7) holds. The sense of these approximations could be specified with a metric. However, we choose instead to follow a convergence approach, somewhat like that in [10,12], in which one considers a sequence of quantizers $\{(S_N,C_N)\}_{N=1}^{\infty}$, where $(S_N,C_N)$ has N points and where the specific point densities $\lambda_{S_N}$ converge to $\lambda$ and the specific inertial profiles $m_{S_N,C_N}$ converge to m (for example, pointwise convergence). With this framework, one establishes the validity of Bennett's integral by showing that the distortion $D(S_N,C_N)$ approaches Bennett's integral as N tends to infinity. Actually, since ordinarily both $D(S_N,C_N)$ and Bennett's integral tend to zero, one shows the stronger result

$$N^{r/k} D(S_N,C_N) \rightarrow \int p(x) \frac{m(x)}{\lambda(x)^{r/k}} dx \ , \text{ as } N \rightarrow \infty \ , \tag{3.1}$$

which indicates the rate of the convergence of $D(S_N,C_N)$ to zero. Now recall that our original intent was to find an approximate formula for the distortion of just one specific quantizer (S,C) with many small cells. However, if we view (S,C) as a member of some sequence, then (3.1) indicates that Bennett's integral will give a good approximation to its distortion.

A few more definitions are needed. Given a quantizer (S,C), let $S_x$ denote the cell containing x, let $d_S(x)$ denote the diameter of the cell $S_x$; i.e.,

$$d_S(x) = \sup \{ \|u-v\| : u,v \in S_x \} ,$$

and let $e_{S,C}(x)$ denote the *normalized local distortion*

$$e_{S,C}(x) = \frac{1}{k} N^{r/k} \|x-Q(x)\|^r . \tag{3.2}$$

It follows that

$$e_{S,C}(x) \leq \frac{1}{k} N^{r/k} d_S(x)^r \leq \sup_{u \in S_x} 2^r e_{S,C}(u) . \tag{3.3}$$

Finally, let $\gamma_{S,C}(x)$ denote the *specific distortion profile*

$$\gamma_{S,C}(x) = \frac{\int_{S_x} e_{S,C}(y)\, dy}{v(S_x)} = \frac{m_{S,C}(x)}{\lambda_S(x)^{r/k}} , \quad \text{if } v(S_x) < \infty, \tag{3.4}$$

and $\gamma_{S,C}(x) = 0$ if $v(S_x) = \infty$. Note that $\gamma_{S,C}(x)$ is constant on the cells of S, that

$$N^{r/k}\, D(S,C) \;=\; \int e_{S,C}(x)\, p(x)\, dx \;, \tag{3.5}$$

and that for any small region F that is itself the union of several quantization cells

$$\int_F e_{S,C}(x)\, p(x)\, dx \;\cong\; \gamma_{S,C}(y)\, p(y)\, v(F) \;,$$

when p is continuous and y is an arbitrary point in F. When dealing with a sequence of quantizers $\{(S_N,C_N)\}$, we shall write $\lambda_N$, $m_N$, $e_N$, $d_N$ and $\gamma_N$ rather than $\lambda_{S_N}$, $m_{S_N,C_N}$, etc..

For the approximation arguments that follow, several definitions and results from analysis are needed. The reader interested only in the statement of the main result may skip to Theorem 2.

The term "measure" refers to Lebesgue measure on $\mathbb{R}^k$. All functions and sets will be assumed to be Lebesgue measurable; all integrals will be Lebesgue integrals. By the volume v(E) of a set E, we mean its Lebesgue measure.

A k-dimensional source density p(x) will be called *piecewise continuous* if there exists a finite, disjoint collection of measurable sets such that p(x) is continuous on each and the volume of points not contained in any of these sets is zero.

Finally, a sequence of functions $\{f_N(x)\}$ is said to be uniformly absolutely continuously integrable with respect to a density p (c.f. [29], p. 192, [30], p. 247) if for all $\varepsilon > 0$ there exists $\delta > 0$ such that

$$\left| \int_F f_N(x)\, p(x)\, dx \right| \;\le\; \varepsilon$$

for every N and every set F with probability $P(F) \;=\; \int_F p(x)\, dx \;\le\; \delta$. As an abbreviation, we will say $\{f_N\}$ is UACI-p. In a similar manner, a sequence $\{f_N(x)\}$ is said to be uniformly absolutely continuously integrable with respect to Lebesgue measure (UACI-L) if for all $\varepsilon > 0$ there exists $\delta > 0$ such that

$$\left| \int_F f_N(x)\; dx \right| \;\le\; \varepsilon$$

for every N and every set F with Lebesgue measure $\delta$ or less. Lemma A.1 shows that if $\{f_N\}$ is UACI-p, then $\{f_N\, p\}$ is UACI-L. Appendix A lists some simpler conditions implying UACI-p. The reason for paying attention to UACI-L is the following very general integral convergence theorem (c.f. [29], p. 192, [30], pp. 249, 324):

*Theorem 1:* Let $\{f_N\}_{N=1}^{\infty}$ be a sequence of functions such that $f_N(x)$ converges in (Lebesgue) measure to a function f(x) (i.e. for any $\varepsilon > 0$, $v(\{x: |f_N(x)-f(x)| > \varepsilon\}) \to 0$ as $N \to \infty$)

and such that the $f_N$'s are UACI-L. Then for any set $E$ with $v(E) < \infty$ and $\int_E |f_N(x)|\, dx < \infty$ for all $N$,

$$\lim_{N\to\infty} \int_E f_N(x)\, dx = \int_E f(x)\, dx . \qquad ❑$$

The above theorem will be used in the proof of the following, which is the main result of the paper.

*Theorem 2:* A sequence of k-dimensional vector quantizers $\{(S_N,C_N)\}_{N=1}^{\infty}$, with $(S_N,C_N)$ having $N$ points, is applied to a k-dimensional, absolutely continuous random vector $X$ with probability density $p$ that is piecewise continuous. Let $\lambda_N(x)$ and $m_N(x)$ denote, respectively, the specific point density and the specific inertial profile of $(S_N,C_N)$. Assume

(i) As $N$ tends to infinity, the specific point densities $\lambda_N$ converge in probability to a model point density $\lambda$; i.e., for any $\varepsilon > 0$,

$$P(\{x : |\lambda_N(x) - \lambda(x)| > \varepsilon\}) \to 0 \quad \text{as} \quad N \to \infty .$$

(ii) As $N$ tends to infinity, the specific inertial profiles $m_N$ converge in probability to a model inertial profile $m$.

(iii) As $N$ tends to infinity, the cell diameter functions $d_N$ converge in probability to zero.

(iv) The sequence of normalized local distortions $\{e_N(x)\}$ is UACI-p .

(v) $\displaystyle\int_E \frac{m(x)}{\lambda(x)^{r/k}}\, p(x)\, dx < \infty$ for every bounded set $E$.

Then

$$\lim_{N\to\infty} N^{r/k} D(S_N,C_N) = \int \frac{m(x)}{\lambda(x)^{r/k}}\, p(x)\, dx , \tag{3.6}$$

where both sides may be infinite. ❑

The proof of Theorem 2 is given in the next section. In the following remarks, we discuss its hypotheses and conclusions.

(1) Hypotheses (i) and (ii) insure that for large $N$ the quantizers are characterized, approximately, by the model point density $\lambda$ and the model inertial profile $m$, at least on the support of the density $p$. These "convergence in probability" hypotheses are, of course, weaker than convergence "pointwise", "almost everywhere" or "with probability one" hypotheses would

have been. They are, also, largely insensitive to the behavior of the quantizers outside the support of p. We say "largely" because if there is a cell $S_{N,i}$ that is partly inside and partly outside the support of p, then the values of $\lambda_N(x)$ and $m_N(x)$ for an x within both $S_{N,i}$ and the support are influenced by the portion of $S_{N,i}$ outside the support. For example, if $S_{N,i}$ has infinite volume, then $\lambda_N(x) = 0$.

(2) Although hypothesis (ii) is a reasonable way to specify that quantizers have, approximately, model inertial profile m, the following example shows there are other ways to specify such and have Bennett's integral be valid. Consider a sequence $\{(S_N,C_N)\}$ of scalar quantizers whose model point density is $\lambda(x)$, in the sense of (i), whose cells are intervals, and whose points are centered within their cells, except for overload cells. Assuming $r = 2$, this sequence has model inertial profile $m = 1/12$, in the sense of (ii). For each N, let $(S'_{2N},C'_{2N})$ denote a scalar quantizer with 2N cells, formed by dividing each cell of $S_N$ into two equal subcells. Let all "left" subcells have quantization points at their centers, and let all "right" subcells have quantization points at their right boundaries. Then because adjacent cells have very different moments of inertia (1/12 and 1/3, respectively), their specific inertial profiles do not converge in probability to a model inertial profile. Nevertheless, in a macroscopic sense, $(S'_{2N},C'_{2N})$ has inertial profile $m'(x) = (1/12 + 1/3)/2 = 5/24$, which is the average of the normalized moments of inertia of the left and right cells. And it is easy to see that Bennett's integral correctly predicts the distortion, assuming $\lambda(x)$ and $m'(x)$ are used. On the other hand, if each original cell were split into cells of different sizes with points in their middle, then even though the macroscopic point density is $\lambda(x)$, Bennett's integral would not correctly predict distortion.

(3) Hypothesis (iv), requiring UACI-p, may be viewed as demanding that the cells shrink rapidly over a large enough region of the support of p. From (3.3) and conditions (1) and (2) of Appendix A we see that hypotheses (iii) and (iv) can be replaced by the more restrictive but easier to check condition that the $e_N(x)$'s (or the functions $N^{1/k}d_N(x)$) are uniformly bounded on the support of the density p. Alternatively, hypothesis (iv) may be replaced by the condition that the $e_N(x)$'s (or the functions $N^{1/k}d_N(x)$) are uniformly bounded from above by some function whose integral with respect to p is finite.

(4) Much of the complexity of the proof derives from the fact that the source density is allowed to be unbounded and to have unbounded support. However, we emphasize that the UACI-p hypothesis is needed, even when source the density is bounded and has bounded support.

(5) Given that the UACI-p hypothesis (iv) demands that cells shrink rapidly, it may be surprising that hypothesis (iii) is needed. Moreover, hypothesis (i) implies that the cell volumes are shrinking as $1/N$. Unfortunately, we were unable to prove theorem without hypothesis (iii). Nor were we able to show that the other hypotheses imply (iii).

(6) Although our proof of the theorem relies on the piecewise continuity of the source density, there might be a proof that would not rely on this assumption. For example, it is shown in [31, Theorem 3.2.1], that if the support of the density is compact, then piecewise continuity is not required. One might even be able to prove the theorem without presuming the existence of a probability density. Notice also that no assumption about the continuity of $\lambda$ or $m$ has been made.

(7) Hypothesis (v) might not be essential. When it does not hold, Bennett's integral is infinite, and we conjecture that $N^{r/k}D(S_N,C_N)$ converges to infinity. (Example: If $\lambda(x)$ has finite support, but $p(x)$ has infinite support, then Bennett's integral is infinite.)

(8) It is interesting to compare Theorem 2 to Theorem 2 of [12]. To facilitate such, we state a simplified version of the latter: Suppose a sequence of quantizers $\{(S_N,C_N)\}$ is asymptotically optimum for a source with density $q(x)$, so that from (2.3)

$$\lim_{N\to\infty} N^{r/k} D(S_N,C_N,q) \;=\; M(H^*_{r,k}) \left( \int q(x)^{k/(k+r)}\, dx \right)^{(k+r)/k},$$

where the $q$ was added to $D$ to emphasize the latter's dependence on it. Then if $\{(S_N,C_N)\}$ is used on a source with density $p(x)$ such that

$$\int \frac{1}{\lambda_q(x)^{r/k}}\, p(x)\, dx \;<\; \infty\,, \tag{3.7}$$

where $\lambda_q(x) = q(x)^{k/(k+r)} / \int q(x')^{k/(k+r)}\, dx'$, and if the sequence $\{e_N\}$ is uniformly integrable with respect to $p$ (Condition 4 of Appendix A), then

$$\lim_{N\to\infty} N^{r/k} D(S_N,C_N,p) \;=\; \int \frac{M(H^*_{r,k})}{\lambda_q(x)^{r/k}}\, p(x)\, dx\,.$$

In essence, Theorem 2 of [12] says that if a sequence of quantizers is optimum for source density $q$, then Bennett's integral holds for source density $p$ with the model point density $\lambda_q$ that is the optimum for $q$ and model inertial profile $m(x) = M(H^*_{r,k})$, which is the optimum for any source density. Since any target point density $\lambda(x)$ is optimum for some source density (specifically, to the density proportional to $\lambda(x)^{(k+r)/k}$), this result is much like Theorem 2 except that the latter gives a direct statement and proof (there is no "if optimal for $q$"

hypothesis), and the latter holds for quantizers with arbitrary inertial profiles. Also, hypothesis (v) of Theorem 2 is a little weaker than (3.7).

## IV. PROOF OF THEOREM 2

We prove the theorem assuming $p$ is continuous on a set $B$ with nonzero volume, and zero elsewhere. (It could be discontinuous on the boundary of $B$.) The proof for a piecewise continuous density follows by direct application of this. The details will be omitted.

Let $\gamma(x) = m(x)/\lambda(x)^{r/k}$ denote the *model distortion profile*. We begin by considering the case where Bennett's integral is finite; i.e., $\int \gamma(x)\, p(x)\, dx < \infty$.

Fix an $\varepsilon > 0$. Since Bennett's integral is finite and since $\{e_N\}$ is UACI-p, we may choose $\delta$ so small that for any set $E$ with $P(E) \le \delta$

$$\int_E \gamma(x)\, p(x)\, dx \le \varepsilon . \tag{4.1}$$

$$\int_E e_N(x)\, p(x)\, dx \le \varepsilon , \text{ for all } N . \tag{4.2}$$

Let $A_o$ be a bounded, open subset of $B$ such that $p(x) > 0$ for $x \in A_o$ and such that

$$P(A_o) = \int_{A_o} p(x)\, dx \ge 1 - \delta . \tag{4.3}$$

The bulk of the remainder of the proof is devoted to showing

$$\lim_{N\to\infty} \int_{A_o} e_N(x)\, p(x)\, dx = \int_{A_o} \gamma(x)\, p(x)\, dx . \tag{4.4}$$

It will then follow from (3.5) and (4.1)-(4.4) that

$$\lim_{N\to\infty} \left| N^{r/k} D(S_N,C_N) - \int \frac{m(x)}{\lambda(x)^{r/k}}\, p(x)\, dx \right|$$

$$= \lim_{N\to\infty} \left| \int e_N(x) p(x) dx - \int \gamma(x) p(x)\, dx \right|$$

$$\le \lim_{N\to\infty} \left( \left| \int_{A_o} e_N(x) p(x) dx - \int_{A_o} \gamma(x) p(x)\, dx \right| + \int_{A_o^c} e_N(x) p(x) dx + \int_{A_o^c} \gamma(x) p(x)\, dx \right)$$

$$\le 2\varepsilon .$$

The final result (3.6) derives from the fact that the above holds for every $\varepsilon > 0$.

Let $I_N$ denote the set of indices $i$ such that $S_{N,i} \subset A_o$, and let

$$G_N = \bigcup_{i \in I_N} S_{N,i} .$$

Then

$$\int_{A_o} e_N(x)\, p(x)\, dx = \int_{G_N} e_N(x)\, p(x)\, dx + \int_{A_o - G_N} e_N(x)\, p(x)\, dx . \tag{4.5}$$

Let

$$p_N(x) \triangleq \begin{cases} \dfrac{\displaystyle\int_{S_{N,x}} e_N(y)\, p(y)\, dy}{\displaystyle\int_{S_{N,x}} e_N(y)\, dy} , & v(S_{N,x}) < \infty \\ 0 , & v(S_{N,x}) = \infty \end{cases} , \tag{4.6}$$

so that $p_N(x)$ is a piecewise constant approximation to $p(x)$ with the properties that if $v(S_{N,x}) < \infty$, then

$$\inf_{y \in S_{N,x}} p(y) \le p_N(x) \le \sup_{y \in S_{N,x}} p(y) , \quad \text{if } v(S_{N,x}) < \infty , \tag{4.7}$$

and

$$\int_{S_{N,x}} e_N(y)\, p(y)\, dy = p_N(x) \int_{S_{N,x}} e_N(y)\, dy = p_N(x)\, \gamma_N(x)\, v(S_x)$$

$$= \int_{S_{N,x}} \gamma_N(y)\, p_N(y)\, dy , \tag{4.8}$$

where the three equalities derive, respectively, from the definition of $p_N(x)$, the definition of $\gamma_N(x)$, and their piecewise constant natures. Using (4.8) in the first term on the right side of (4.5) gives

$$\int_{G_N} e_N(x)\, p(x)\, dx = \sum_{i \in I_N} \int_{S_{N,i}} e_N(x)\, p(x)\, dx .$$

$$= \sum_{i \in I_N} \int_{S_{N,i}} \gamma_N(x)\, p_N(x)\, dx$$

$$= \int_{G_N} \gamma_N(x)\, p_N(x)\, dx$$

$$= \int_{A_o} \gamma_N(x)\, p_N(x)\, dx \quad - \int_{A_o - G_N} e_N(x)\, p_N(x)\, dx \ .$$

Substituting the above into (4.5) gives

$$\int_{A_o} e_N(x)\, p(x) dx = \int_{A_o} \gamma_N(x)\, p_N(x)\, dx - \int_{A_o - G_N} \gamma_N(x)\, p_N(x)\, dx + \int_{A_o - G_N} e_N(x)\, p(x)\, dx \ . \quad (4.9)$$

From here, we take the following steps: Lemmas A.2 and A.3 of Appendix B show that the hypotheses about the convergence of $m_N$ and $\lambda_N$ imply that $\gamma_N$ converges in (Lebesgue) measure on $A_o$ to $\gamma$. Lemmas A.2 and A.4 show that the continuity of $p$ and the convergence of $d_N$ to zero imply that $p_N$ converges in (Lebesgue) measure to $p$ on $A_o$. Therefore, the product $\gamma_N p_N$ converges in measure to $\gamma p$ on $A_o$ (c.f. [29], p. 63). Lemma A.1 shows that the fact that $\{e_N\}$ is UACI-p implies $\{e_N p\}$ is UACI-L, and Lemma A.5 uses this to show that $\{\gamma_N p_N\}$ is UACI-L. It then follows from Theorem 1 that

$$\lim_{N\to\infty} \int_{A_o} \gamma_N(x)\, p_N(x)\, dx = \int_{A_o} \gamma(x)\, p(x)\, dx \ . \quad (4.10)$$

Next Lemma A.6 uses the hypothesized convergence of $d_N$ and the fact that $A_o$ is an open set with finite measure to show that $v(A_o\text{-}G_N)$ converges to zero. It follows from the UACI-L properties of $\{e_N p\}$ and $\{\gamma_N p_N\}$ that the last two terms of (4.9) converge to zero. Using this and (4.10) in (4.9) yields (4.4), and finishes the proof of Theorem 2 in the case that Bennett's integral is finite.

If Bennett's integral is infinite, i.e. $\int \gamma(x)\, p(x)\, dx = \infty$, then for any positive number $B$ one may choose a bounded, open set $A_o$ such that (4.3) holds and, additionally,

$$\int_{A_o} \gamma(x)\, p(x)\, dx \geq B \ .$$

Hypothesis (v) implies that the left-hand side of the above is finite. Therefore, (4.4) holds and shows that for all sufficiently large $N$,

$$\int e_N(x)\, p(x)\, dx \geq \int_{A_o} e_N(x)\, p(x)\, dx \geq B/2 \ .$$

Since $B$ can be chosen arbitrarily large, it must be that $\lim\limits_{N\to\infty} \int e_N(x)\, p(x)\, dx = \infty$.

## ACKNOWLEDGEMENT

The authors are indebted to Dennis Hui, Peter Moo and Neri Merhav for their helpful comments and suggestions.

## APPENDIX A. UNIFORM ABSOLUTE CONTINUITY

*Conditions implying UACI-p*

The following conditions imply $\{f_N\}$ is UACI-p:

1. The $f_N$'s are uniformly bounded on the support of $p$; i.e., there exists $K < \infty$ such that $|f_N(x)| \leq K$ for all $N$ and all $x$.

2. There exists a function $g$ such that $\int g(x)\, p(x)\, dx < \infty$, and $|f_N(x)| \leq g(x)$ for all $N, x$.

3. The integrals $\int f_N(x)^2 p(x)\, dx$ are uniformly bounded (c.f. [32], vol. 1, p. 159).

4. The $f_N$'s are uniform integrable (c.f. [33], p. 139); i.e., for any $\varepsilon > 0$ there exists $\alpha > 0$ such that

$$\int_{|f_N(x)|>\alpha} f_N(x)\, p(x)\, dx \;\leq\; \varepsilon \quad \text{for all } N.$$

Actually, uniform integrability is equivalent to UACI-p plus uniform boundedness of the integrals $\int f_N(x)\, p(x)\, dx$. However, if the sequence of functions converges in probability, then uniform integrability and UACI-p are equivalent.

*Lemma A.1:* If $\{f_N\}$ is UACI-p, then $\{f_N\, p\}$ is UACI-L. ❑

*Proof:* Fix $\varepsilon > 0$. By the UACI-p hypothesis there exists $\delta > 0$ such that $\int_E f_N(x)\, p(x)\, dx \leq \varepsilon$ for every $N$ and every set $E$ with $P(E) \leq \delta$. Moreover, by the integrability of $p$ there exists $\beta > 0$ such that if $v(E) \leq \beta$, then $P(E) = \int_E p(x)\, dx \leq \delta$. Putting these together gives $\int_E f_N(x)\, p(x)\, dx < \varepsilon$ for every $N$ and every set $E$ with $v(E) \leq \beta$. Thus $\{f_N\, p\}$ is UACI-L.

## APPENDIX B. DETAILS OF THE PROOF OF THEOREM 2

For the proof of Theorem 2, we need to show that $\{\gamma_N\}$ and $\{p_N\}$ converge, respectively, to $\gamma$ and $p$, in measure on $A_o$. The following lemma demonstrates that since $v(A_o) < \infty$ and $p(x)$ is strictly positive on $A_o$, it is sufficient to show convergence in probability.

*Lemma A.2:* Let $X$ be a k-dimensional, absolutely continuous random vector with density $p$. If $f_N(X)$ converges in probability to $f(X)$, then $f_N(X)$ also converges in (Lebesgue) measure to $f(X)$ on any set $B$ such that $v(B) < \infty$ and $p(x) > 0$ for all $x \in B$. ❑

*Proof:* Fix an $\varepsilon > 0$ and a set $B$ such that $v(B) < \infty$ and $p(x) > 0$ for all $x \in B$. Let $E_N = \{ x \in B : |f_N(x)-f(x)| > \varepsilon \}$. To prove the lemma we show that $v(E_N) \to 0$ as $N \to \infty$. For $\beta$ to be chosen later, let $C_\beta = \{ x \in B : p(x) > \beta \}$. Then

$$P(E_N) \geq \int_{E_N \cap C_\beta} p(x)\, dx > \beta\, v(E_N \cap C_\beta) .$$

Hence

$$v(E_N \cap C_\beta) < \frac{P(E_N)}{\beta}$$

and

$$v(E_N) = v(E_N \cap C_\beta) + v(E_N - C_\beta) < \frac{P(E_N)}{\beta} + v(B-C_\beta) .$$

Since $p$ is strictly positive on $B$, the set $B-C_\beta = \{ x \in B : p(x) \leq \beta \}$ approaches the empty set as $\beta$ tends to zero. Because of this and the fact that $v(B-C_\beta) \leq v(B) < \infty$ for all $\beta$, the "continuity of measure" property (of Lebesgue measure) implies that $v(B-C_\beta) \to 0$ as $\beta \to 0$ (c.f. [33], p. 11). Substituting $\beta = \sqrt{P(E_N)}$ into the above and using the fact that $P(E_N)$ tends to zero (by the assumed convergence in probability of $f_N$ to $f$ ), we find that $v(E_N)$ tends to zero. Since this is true for every $\varepsilon > 0$, $f_N$ converges to $f$ in measure on $B$.

The next two lemmas demonstrate convergence in probability, which by the previous lemma insures convergence in measure (Lebesgue) on $A_o$ .

*Lemma A.3:* Let $X$ be a k-dimensional, absolutely continuous random vector with density $p$. If Bennett's integral is finite on $A_o$, $\lambda_N(X)$ converges in probability to $\lambda(X)$ and $m_N(X)$ converges in probability to $m(X)$, then $\gamma_N(X) = m_N(X)/\lambda_N^{r/k}(X)$ converges in probability to $\gamma(X) = m(X)/\lambda^{r/k}(X)$ on $A_o$. ❑

*Proof:* Since $\lambda_N$ converges in probability to $\lambda$ and since $(\cdot)^{r/k}$ is a continuous function, it follows (c.f. [33], p. 90) that $\lambda_N^{r/k}$ converges in probability to $\lambda^{r/k}$. Furthermore, the assumption that $m_N$ converges in probability to $m$ will imply that $m_N/\lambda_N^{r/k}$ converges in probability to $m/\lambda^{r/k}$ on $A_o$, provided it can be shown that $P(\{\lambda^{r/k} = 0\} \cap A_o) = 0$ (c.f. [29], p. 63). Since Bennett's integral has been assumed to be finite on $A_o$, it must be that $P(\{m/\lambda^{r/k} = \infty\} \cap A_o) = 0$. Since each $m_N$ is bounded from below by the normalized moment of inertia of a sphere, $m$ must be similarly bounded, with probability one. Thus, $P(\{\lambda^{r/k} = 0\} \cap A_o) = 0$, and we conclude that $\gamma_N = m_N/\lambda_N^{r/k}$ converges in probability to $m/\lambda^{r/k}$ on $A_o$.

*Lemma A.4:* Let $X$ be a k-dimensional, absolutely continuous random vector with a continuous density $p$. For $N = 1, 2, \ldots$, let $S_N$ be a partition of $\Re^k$ into $N$ cells such that such that $d_N(X)$ converges in probability to zero; let $p_N(x)$ be such that

$$\inf_{y \in S_{N,x}} p(y) \le p_N(x) \le \sup_{y \in S_{N,x}} p(y) \tag{A.1}$$

for all $x$ such that $v(S_{N,x}) < \infty$, and let $p_N(x)$ be defined arbitrarily for all $x$ such that $v(S_{N,x}) = \infty$. Then $p_N(X)$ converges to $p(X)$ in probability. ❑

*Proof:* Fix some $\varepsilon > 0$, and let

$$\delta(x,\varepsilon) = \sup\{ \delta : |p(y)-p(x)| \le \varepsilon \text{ for all } y \text{ such that } \|y-x\| \le \delta \} .$$

Thus, if $\|y-x\| < \delta(x,\varepsilon)$, then $|p(y)-p(x)| \le \varepsilon$. The continuity of $p$ implies that $\delta(x,\varepsilon) > 0$ for all $x$, and that

$$\{x : \delta(x,\varepsilon) < \beta\} = \{ x : |p(y)-p(x)| \ge \varepsilon \text{ for some } y \text{ such that } \|y-x\| < \beta \}$$

tends to the empty set as $\beta$ tends to $0$. Thus for any $\varepsilon_1$, one may choose $\beta$ so small that

$$P(\{x: \delta(x,\varepsilon) < \beta\}) < \varepsilon_1 .$$

Next observe that if $d_N(x) < \delta(x,\varepsilon)$, then because of (A.1), $|p_N(x)-p(x)| \le \varepsilon$. The contrapositive implies

$$\begin{aligned} P( |p_N(X)-p(X)| > \varepsilon ) &\le P( d_N(X) \ge \delta(X,\varepsilon) ) \\ &\le P( d_N(X) \ge \beta \text{ or } \beta > \delta(X,\varepsilon) ) \\ &\le P( d_N(X) \ge \beta ) + P( \beta > \delta(X,\varepsilon) ) \end{aligned}$$

$$\leq \; P(\, d_N(X) \geq \beta \,) \; + \; \varepsilon_1 \, .$$

By the hypothesis that $d_N(X)$ converges to zero in probability, the first term on the right side of the above converges to zero. Thus $\limsup_{N\to\infty} P(\, |p_N(X)-p(X)| > \varepsilon \,) \leq \varepsilon_1$. But since this holds for any $\varepsilon_1 > 0$, we have $\lim_{N\to\infty} P(\, |p_N(X)-p(X)| > \varepsilon \,) = 0$; i.e., $p_N(X)$ converges in probability to $p(X)$.

*Lemma A.5:* Let $X$ be a k-dimensional, absolutely continuous random vector with density $p(x)$. For $N = 1, 2, \ldots,$ let $(S_N,C_N)$ be an N-point quantizer, and let $e_N(x)$, $\gamma_N(x)$ and $p_N(x)$ be as defined in (3.2), (3.4) and (4.6), respectively. If $\{e_N\, p\}$ is UACI-L, then $\{\gamma_N\, p_N\}$ is UACI-L. ❑

*Proof:* We will show that for any $N$ and any set $E$ with $v(E) < \infty$, there exists a set $F$ with $v(F) = v(E)$ and

$$\int_E \gamma_N(x)\, p_N(x)\; dx \;\; \leq \;\; \int_F e_N(x)\, p(x)\; dx \; . \tag{A.2}$$

Since $\{e_N\, p\}$ is UACI-L, for any $\varepsilon > 0$ there exists $\delta > 0$ such that $\int_F e_N(x)\, p(x)\; dx \leq \varepsilon$, whenever $v(F) \leq \delta$. It then follows from this and (A.2) that if $v(E) \leq \delta$, then $\int_E \gamma_N(x)\, p_N(x)\; dx \leq \varepsilon$, so that $\{\gamma_N\, p_N\}$ is UACI-L.

Equation (A.2) holds trivially when $v(E) = 0$; so we now assume $v(E) > 0$. We let $F = \cup_{i=1}^{N} F_{N,i}$, where if $v(S_{N,i}) = \infty$, then $F_{N,i} = E \cap S_{N,i}$, and if $v(S_{N,i}) < \infty$, then $F_{N,i}$ is a subset of $S_{N,i}$ with measure $v(F_{N,i}) = v(E\cap S_{N,i})$ that is chosen to contain the points $x$ in $S_{N,i}$ for which $e_N(x)p(x)$ is largest. Clearly, $v(F) = v(E)$. Now, if $v(S_{N,i}) = \infty$, then

$$\int_{E\cap S_{N,i}} \gamma_N(x)\, p_N(x)\; dx \;\; = \;\; 0 \;\; \leq \;\; \int_{F_{N,i}} e_N(x)\, p(x)\; dx \; , \tag{A.3}$$

while if $v(S_{N,i}) < \infty$, then

$$\int_{E\cap S_{N,i}} \gamma_N(x)\, p_N(x)\; dx \;\; = \;\; v(E\cap S_{N,i})\, \frac{\int_{S_{N,i}} e_N(x)\, dx}{v(S_{N,i})} \; \frac{\int_{S_{N,i}} e_N(x)\, p(x)\; dx}{\int_{S_{N,i}} e_N(x)\, dx}$$

$$= \;\; v(E\cap S_{N,i}) \;\; \frac{1}{v(S_{N,i})} \int_{S_{N,i}} e_N(x)\, p(x)\; dx$$

$$\leq \; v(E \cap S_{N,i}) \; \frac{1}{v(F_{N,i})} \int_{F_{N,i}} e_N(x)\, p(x) \; dx$$

$$= \int_{F_{N,i}} e_N(x)\, p(x) \; dx \, , \qquad \text{(A.4)}$$

where the inequality and last equality follow from the way $F_{N,i}$ was chosen. Finally, (A.3) and (A.4) imply (A.2). This completes the proof of Lemma A.5 .

*Lemma A.6:* Let $X$ be a k-dimensional, absolutely continuous random vector with density $p(x)$. Let $A_o$ be an open set with finite volume such that $p(x) > 0$, $x \in A_o$. For $N = 1, 2, \ldots$, let $S_N$ be a partition of $\mathfrak{R}^k$ into $N$ cells such that $d_N(X)$ converges to zero in probability. Let $G_N = \{ x : S_{N,x} \subset A_o \}$. Then

$$v(A_o\text{-}G_N) \to 0 \;\text{ as }\; N \to \infty \, . \qquad ❑$$

*Proof:* Since $A_o$ is an open set, for each $x \in A_o$ there exists $\delta_x > 0$ such that every $y$ within distance $\delta_x$ of $x$ is also in $A_o$. It follows that if $x \in A_o$ and $d_N(x) \leq \delta_x$, then $S_{N,x} \subset A_o$ and so $x \in G_N$. The contrapositive shows that if $x \in A_o$ and $x \notin G_N$, then $d_N(x) > \delta_x$. That is,

$$A_o\text{-}G_N \;\subset\; \{ x \in A_o : \; d_N(x) > \delta_x \} \, .$$

For $\beta > 0$, let $B_\beta = \{ x \in A_o : \beta > \delta_x)$. The sets $B_\beta$ converge monotonically to the empty set as $\beta$ approaches 0. Since $v(B_\beta) \leq v(A_o) < \infty$, the "continuity of measure" property (of Lebesgue measure) implies $v(B_\beta) \to 0$ as $\beta \to 0$ (c.f. [33], p.11). Therefore, given $\varepsilon > 0$, one may choose $\beta$ so small that $v(B_\beta) \leq \varepsilon$ . We then have

$$\{ x \in A_o : d_N(x) > \delta_x \} \;\subset\; \{ x \in A_o : \; d_N(x) > \beta \} \cup \{ x \in A_o : \; \beta > \delta_x \} \, .$$

So,

$$v(A_o\text{-}G_N) \;\leq\; v(\{x \in A_o : d_N(x) > \beta\}) + \varepsilon \, .$$

The hypothesis that $d_N(X)$ converges to zero in probability combined with Lemma A.2 implies that $d_N(X)$ converges in measure (Lebesgue) on $A_o$. Hence, from the above $\lim_{N\to\infty} v(A_o\text{-}G_N) \leq \varepsilon$. Since this holds for every $\varepsilon > 0$, the limit must be zero, and the lemma is proved.

## REFERENCES


1. W.R. Bennett, "Spectra of quantized signals", BSTJ, vol. 27, pp. 446-472, July 1948.
2. S.P. Lloyd, "Least squares quantization in PCM," Bell Labs Memorandum, 1957; also published in *IEEE Trans. Inform. Theory*, vol. IT-28, p. 129-137, March 1982.
3. H. Gish and J.R. Pierce, "Asymptotically efficient quantizing," *IEEE Trans. Inform. Theory* , vol. IT-14, pp. 676-683, Sept. 1968.
4. T. Berger, *Rate Distortion Theory*, p. 173, Englewood Cliffs, N.J.: Prentice-Hall, 1971.
5. A. Gersho, "Asymptotically optimal block quantization," *IEEE Trans. Inform. Theory*, vol. IT-25, pp. 373-380, July 1979.
6. V.R. Algazi, "Useful approximations to optimum quantization," *IEEE Trans. Comm.*, vol. COM-14, pp. 297-301, June 1966.
7. 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.
8. J.A. Bucklew, "Companding and random quantization in several dimensions," *IEEE Trans. Inform. Theory*, vol. IT-27, pp. 207-211, March 1981.
9. J.A. Bucklew, "A note on optimal multidimensional companders," *IEEE Trans. Inform. Theory*, vol. IT-29, p. 279, March 1983.
10. 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, March 1982.
11. S. Cambanis and N. Gerr, "A simple class of asymptotically optimal quantizers," *IEEE Trans. Inform. Theory*, vol. IT-29, pp. 664-676, Sept. 1983.
12. J.A. Bucklew, "Two results on the asymptotic performance of quantizers," *IEEE Trans. Inform. Theory*, vol. IT-30, pp. 341-348, March 1984.
13. P.F. Panter and W. Dite, "Quantization distortion in pulse-count modulation with nonuniform spacing of levels," *Proc. IRE*, vol. 39, pp. 44-48, Jan. 1951.
14. M.P. Schutzenberger, "On the quantization of finite-dimensional messages," *Inform. and Control*, vol. 1, pp. 153-158, 1958.
15. P. Zador, "Development and evaluation of procedures for quantizing multivariate distributions", Ph.D. Thesis, Stanford, 1964.
16. P. Zador, "Asymptotic quantization error of continuous signals and quantization dimension," *IEEE Trans. Inform. Theory*, vol. IT-28, pp. 139-149, 1982.

17. G.M. Roe, "Quantizing for minimum distortion," *IEEE Trans. Inform. Theory*, vol. IT-10, pp. 384- 385, Oct. 1964.
18. R.C. Wood, "On optimum quantization," *IEEE Trans. Inform. Theory* , vol. IT-15, pp. 248-252, March 1969.
19. P. Elias, "Bounds on performance of optimum quantizers," *IEEE Trans. Inform. Theory*, vol. IT-16, pp. 172-184, March 1970.
20. D. Neuhoff and D.H. Lee, "On the Performance of tree-structured vector quantizers," *ICASSP*, pp. 2277-2280, Tornoto, May 1991.
21. D. H. Lee and D.L. Neuhoff, "Conditionally corrected two-stage vector quantization," *1990 Conf. Inform. Sciences and Systems*, pp. 802-806, Princeton, Mar. 1990.
22. D.H. Lee, D.L. Neuhoff and K.K. Paliwal, "Cell-conditioned two-stage vector quantization of speech," *ICASSP*, vol. 4, pp. 653-656, Toronto, May 1991.
23. A. Gersho and R.M. Gray, *Vector Quantization and Signal Compression*, Boston: Kluwer, 1991.
24. N.S. Jayant and P. Noll, *Digital Coding of Waveforms,* Englewood Cliffs, NJ: Prentice-Hall, 1984.
25. T. Lookabough and R.M. Gray, "High resolution quantization theory and the vector quantizer advantage," *IEEE Trans. Inform. Theory*, vol. IT-35, pp. 1020-1033, Sept. 1989.
26. D.L. Neuhoff, "On the performance of structured quantizers," *Proc. 1988 International Symp. on Communications and Controls,* pp. 122-127, Baton Rouge, LA, Oct. 1988.
27. J.H. Conway and N.J.A. Sloane, *Sphere Packings, Lattices and Groups*, New York: Springer-Verlag, 1988.
28. J.J.Y. Huang and P.M. Schultheiss, "Block quantization of correlated Gaussian variables," *IEEE Trans. Commun.*, vol. COM-11, pp. 289-296, Sept. 1963.
29. R. L. Wheeden and A. Zygmund, *Measure and Integral an Introduction to Real Analysis*, New York: Dekker, 1977.
30. T. H. Hildebrandt, *Introduction to the Theory of Integration*, New York: Academic Press, 1963.
31. S. Na, *Source Coding of Composite Sources with Segmental Fidelity Criteria*, Ph.D. Dissertation, University of Michigan, 1989.
32. I. P. Natanson, *Theory of Functions of a Real Variable*, translated by L. F. Boron, New York: Ungar, 1955.
33. D. L. Cohn, *Measure Theory*, Boston: Birkhauser, 1980.

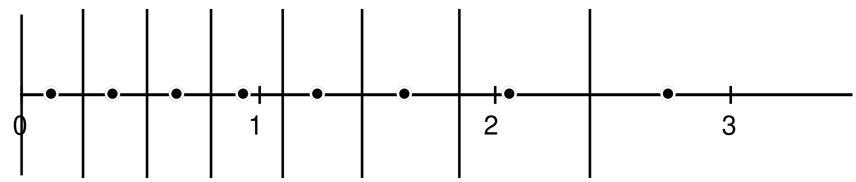
| k | cubic $L_{cu}$ | oblongitis $L^*_{ob}$ | point dens. $L^*_{pt}$ | shape $L^*_{ob}L^*_{pt}$ | total $L_{cu}L^*_{ob}L^*_{pt}$ |
|---|---|---|---|---|---|
| 2 | 0.1671 | 0.6247 | 0.5115 | 1.1362 | 1.3033 |
| 3 | 0.2571 | 0.7484 | 0.8622 | 1.6106 | 1.8677 |
| 4 | 0.3949' | 0.8020 | 1.0721 | 1.8741 | 2.2690' |
| 8 | 0.6572' | 0.8741 | 1.4373 | 2.3113 | 2.9686' |
| 12 | 0.8084' | 0.8962 | 1.5744 | 2.4705 | 3.2789' |
| 16 | 0.9094' | 0.9069 | 1.6462 | 2.5531 | 3.4625' |
| 24 | 1.0385' | 0.9175 | 1.7203 | 2.6377 | 3.6762' |
| $\infty$ | 1.5329 | 0.9380 | 1.8759 | 2.8139 | 4.3468 |

(a) Gaussian.

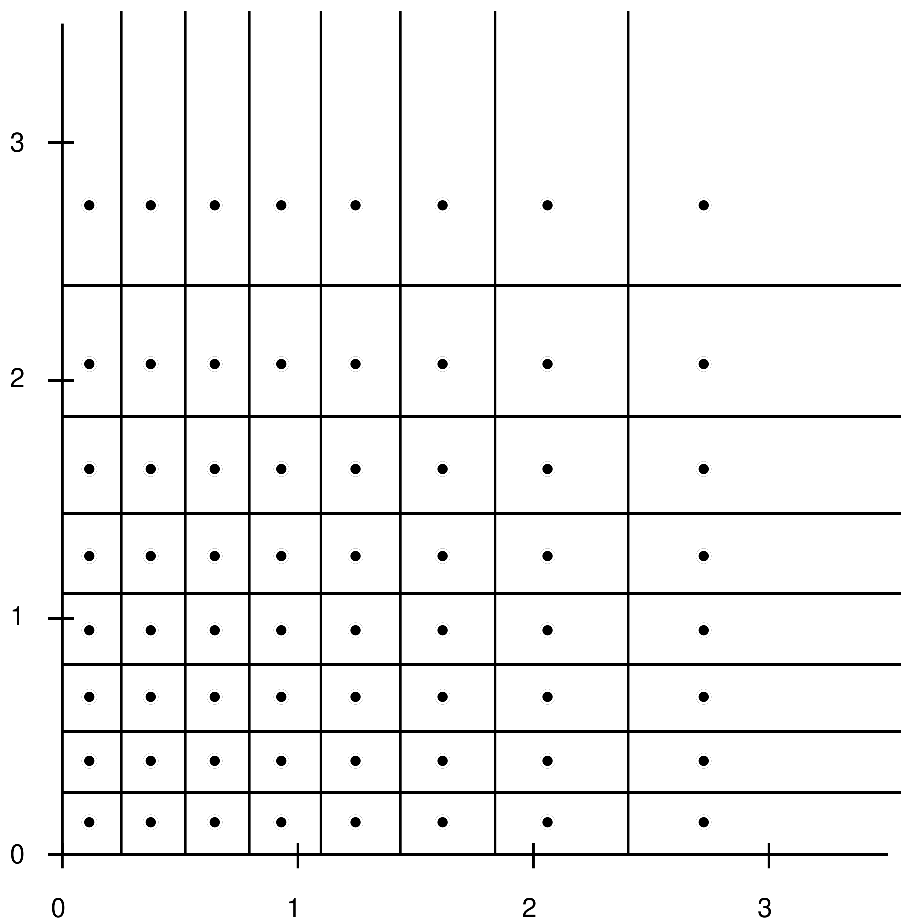
| k | cubic $L_{cu}$ | oblongitis $L^*_{ob}$ | point dens. $L^*_{pt}$ | shape $L^*_{ob}L^*_{pt}$ | total $L_{cu}L^*_{ob}L^*_{pt}$ |
|---|---|---|---|---|---|
| 2 | 0.1671 | 1.2494 | 1.0231 | 2.2724 | 2.4395 |
| 3 | 0.2571 | 1.4969 | 1.7243 | 3.2212 | 3.4783 |
| 4 | 0.3949' | 1.6040 | 2.1442 | 3.7482 | 4.1431' |
| 8 | 0.6572' | 1.7481 | 2.8745 | 4.6226 | 5.2799' |
| 12 | 0.8084' | 1.7924 | 3.1487 | 4.9411 | 5.7495' |
| 16 | 0.9094' | 1.8139 | 3.2923 | 5.1062 | 6.0155' |
| 24 | 1.0385' | 1.8349 | 3.4405 | 5.2755 | 6.3140' |
| $\infty$ | 1.5329 | 1.8759 | 3.7518 | 5.6277 | 7.1607 |

(b) Laplacian.

Table 1: Losses (in dB) for optimal kth-power quantizers for stationary, memoryless Gaussian and Laplacian sources. The "primed" losses are upper bounds based on Conway and Sloane's conjectured lower bound to $M^*_k$ [28, pp. 61,62].




(a)




(b)

Figure 1: (a) The optimal 16 point scalar quantizer for a unit variance Gaussian density, and (b) its second power. Only the positive quantization points are shown.

## FIGURES CAPTIONS

Figure 1: (a) The optimal 16 point scalar quantizer for a unit variance Gaussian density, and (b) its second power. Only the positive quantization points are shown.

## TABLE CAPTIONS

Table 1: Losses (in dB) for optimal kth-power quantizers for stationary, memoryless Gaussian and Laplacian sources. The "primed" losses are upper bounds based on Conway and Sloane's conjectured lower bound to $M_k^*$ [28, pp. 61,62].

## FOOTNOTES

[1]There are more general notions of product quantizers than that considered here (cf. [23], p. 430 ff.).

[2]In the terminology of the next section, (2.6) and (2.7) hold for both the "specific" and "model" versions of point density and inertial profile.

[3]The "c" in one expression is not necessarily the same as the "c" in another.

[4]Actually, it was called an "advantage" due to vector quantization, as opposed to a "loss" due to scalar quantization.

[5]If the support of $p_1$ is finite, then choosing $\lambda_1$ to be uniform on it and zero elsewhere makes $L_{ob} = 1$, the minimum possible value. If the support of $p_1$ is infinite, then choosing $\lambda_1$ to be uniform on a sufficiently large subset and zero elsewhere makes $L_{ob}$ arbitrarily close to 1.

[6]For this point density, the distortion does not decrease as rapidly as $1/N^2$.

[7]A product of vector quantizers may also be used, but for simplicity we restrict attention to the product of scalar quantizers. Generalizations are straightforward.