On Lattice Quantization Noise

Ram Zamir, *Member, IEEE*, and Meir Feder, *Senior Member, IEEE*

***Abstract*—We present several results regarding the properties of a random vector, uniformly distributed over a lattice cell. This random vector is the quantization noise of a lattice quantizer at high resolution, or the noise of a dithered lattice quantizer at all distortion levels. We find that for the optimal lattice quantizers this noise is wide-sense-stationary and *white*. Any desirable noise spectra may be realized by an appropriate linear transformation ("shaping") of a lattice quantizer. As the dimension increases, the normalized second moment of the optimal lattice quantizer goes to $1/2\pi e$, and consequently the quantization noise approaches a white Gaussian process in the divergence sense. In entropy-coded dithered quantization, which can be modeled accurately as passing the source through an additive noise channel, this limit behavior implies that for large lattice dimension both the error and the bit rate approach the error and the information rate of an Additive White Gaussian Noise (AWGN) channel.**

***Index Terms*—Lattice, quantization noise, shaping, normalized second moment, divergence from Gaussianity.**

I. Introduction

In high-resolution quantization theory, it is common to assume that the quantization error of a uniform or lattice quantizer has a uniform distribution over the basic cell of the quantizer [1], [8], [9]. This approximation is completely accurate for all resolution levels in (subtractive) dithered quantization, where a uniformly distributed noise, the dither, is added intentionally to the source before quantization and then subtracted from the quantizer output; see, e.g., [10], [11], [21], and [22]. In any case, the (additive) uniform quantization noise model provides a convenient tool in analyzing schemes incorporating uniform, lattice, or linear trellis quantizers.

In light of this model and the wide use of lattices in signal coding, it is interesting to characterize the statistical properties of a random vector which is uniformly distributed over the basic cell of a lattice; see, e.g., [12]. Thus we analyze in this paper the spectral properties and the divergence from Gaussianity of this random vector, referred to as *lattice quantization noise*. We mainly focus on optimal lattice quantizers, i.e., lattice quantizers that minimize the power of the quantization noise, and on their limit properties as the lattice dimension increases.

Manuscript received December 9, 1994; revised February 14, 1996. The research was performed while R. Zamir was at Tel-Aviv University, Tel-Aviv, Israel. The material in this paper was presented in part at the Data Compression Conference (DCC'94), Snowbird, UT, March 1994 and at the IEEE International Symposium on Information Theory, Trondheim, Norway, June 1994.

R. Zamir is with the Department of Electrical and Computer Engineering, University of California, Santa Barbara, CA 93106 USA.

M. Feder is with the Department of Electrical Engineering—Systems, Tel-Aviv University, Tel-Aviv, 69978 Israel.

Publisher Item Identifier S 0018-9448(96)04419-7.

To be precise let us begin with the definition of a *lattice quantizer*, which is somewhat broader than the usual definition given, e.g., in [4]. A quantizer is defined by a set of code points and a partition which is associated with it. The code points of a K-dimensional lattice quantizer form a K-dimensional lattice $\mathcal{L} = \{\underline{l}_i\}$, i.e.,

$$\underline{l}_i \in \mathcal{R}^K, \quad \underline{l}_0 = \underline{0}, \quad \underline{l}_i + \underline{l}_j \in \mathcal{L} \;\; \forall i, j\,. \tag{1}$$

No overload is assumed. Throughout the paper we assume that the lattice is *nondegenerate*, i.e., it is spanned by K linearly independent basis vectors (the rows of its generator matrix). The partition $\mathcal{P} = \{P_i\}$ associated with the lattice quantizer is a collection of disjoint regions (whose union covers $\mathcal{R}^K$) which satisfy

$$P_i = \underline{l}_i + P_0 = \{\underline{x} : \; \underline{x} - \underline{l}_i \in P_0\} \tag{2}$$

i.e., the ith cell is a shift of the basic cell P_0 by the ith point of the lattice. The lattice quantizer $Q_K = \{\mathcal{L}, \mathcal{P}\} \, : \, \mathcal{R}^K \to \mathcal{R}^K$ maps every vector $\underline{x}_K \in \mathcal{R}^K$ into the lattice point that is associated with the cell containing $\underline{x}_K$, i.e.,

$$Q_K(\underline{x}_K) = \underline{l}_i, \quad \text{if} \quad \underline{x}_K \in P_i\,. \tag{3}$$

The simplest example of a lattice quantizer is the uniform (scalar) quantizer, whose code points are $\{0, \pm\Delta, \pm 2\Delta, \cdots\}$, and its quantization function $Q_1 : \mathcal{R} \to \mathcal{R}$ is such that

$$Q_1(x) = i\Delta, \quad \text{for} \quad i\Delta - \Delta/2 \le x \le i\Delta + \Delta/2\,. \tag{4}$$

In general, however, there are many possible partitions with respect to a given lattice, all have the same cells volume $V = \mu(P_0)$ which equals the reciprocal of the lattice points density, or the determinant of the lattice's generator matrix [4]. When every source vector is mapped into the *nearest* code point (as in the uniform quantizer example above), we get the *Voronoi partition*, in which the ith cell (denoted now V_i) is given by

$$V_i = \{\underline{x} : \; ||\underline{x} - l_i|| \le ||\underline{x} - l_j||, \; \forall j \ne i\} \tag{5}$$

where $||\cdot||$ denotes some norm. The discussion in this paper will be limited to the Euclidean norm. Fig. 1 shows three possible partitions with respect to the hexagonal lattice, where partition A is a Voronoi partition. We note that while the Voronoi partition (5) is optimal in many cases, sometimes the more general definition of the partition and the mapping rule (3) is needed. One such example occurs when we incorporate pre- and post-filters [23] in the quantization process. Other examples are noisy source quantization, and entropy-constrained quantization.

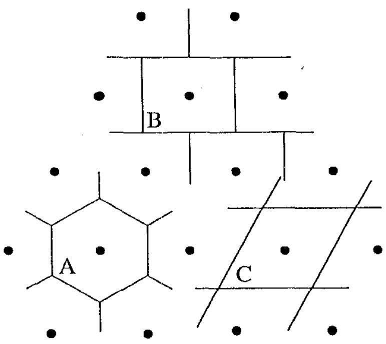

Fig. 1. The hexagonal lattice and three possible associated partitions: A—Voronoi partition (hexagons), B—rectangles, C—rhombuses generated via shaping of the $\mathcal{Z}$ lattice and its Voronoi partition.

An important structure figure of lattice quantizers which is used extensively in the literature as a measure for the quantizer efficiency is the *normalized second moment* [4]

$$G_K = G(Q_K) = G(\mathcal{L}, \mathcal{P}) = \frac{1}{K} \frac{\int_{P_0} \|\underline{x}\|^2 \, d\underline{x}}{V^{1+2/K}}. \tag{6}$$

The normalized second moment is invariant to scaling and rotation of the space. It relates the density of the lattice points $1/V$, with the Mean Squared quantization Error (MSE) per dimension

$$\epsilon = \frac{1}{K} \frac{\int_{P_0} \|\underline{x}\|^2 \, d\underline{x}}{V} = G_K \cdot V^{2/K}. \tag{7}$$

For the uniform scalar quantizer $G_1 = 1/12 \simeq 0.08333$, while $G_2 \simeq 0.080188$ for the hexagonal lattice quantizer with a Voronoi partition (see Fig. 1). It was recently shown by Poltyrev [19] (see also Section III in this paper) that as $K \to \infty$, the minimum value of $G_K \to 1/2\pi e \simeq 0.058550$.

The goal of this work is to analyze the quantization noise, given either by $Q_K(\underline{x}) - \underline{x}$, or by $Q_K(\underline{x}+\underline{z}) - \underline{x} - \underline{z}$, where $\underline{z}$ is the dither, for dithered quantization. As noted in the beginning of the paper, this noise is modeled as a random vector

$$\underline{Z} \sim \mathcal{U}(P_0) \tag{8}$$

uniformly distributed over the basic cell of Q_K. We investigate in this work the statistical properties of $\underline{Z}$. In the next section we show that for optimal lattice quantizers the correlation matrix of $\underline{Z}$ is proportional to the identity matrix, i.e., the components of $\underline{Z}$ are uncorrelated and have the same power. Then, we further analyze this correlation matrix when the lattice is linearly shaped. Section III determines the distance of this vector from a Gaussian vector measured by the (normalized) information divergence, and shows that this distance vanishes asymptotically for large (optimal) lattice dimension.

A slight generalization of the lattice quantizer is the *tessellating quantizer*, in which the basic cell P_0 may be rotated, and not only translated, to get the ith cell [8], [14]. For example, an equilateral triangle cell generates a tessellating quantizer which is not a lattice quantizer. Despite their slight generality, tessellating quantizers are not considered in this work since their resulting noise can not be modeled as additive.

II. Lattice Quantization Noise Spectra

In this section we consider

$$R_z \triangleq E\{\underline{Z}\,\underline{Z}^t\} = \frac{1}{V} \int_{P_0} \underline{z}\,\underline{z}^t \, d\underline{z} \tag{9}$$

the autocorrelation matrix of the quantization noise $\underline{Z} \sim \mathcal{U}(P_0)$. Note that since the lattice is nondegenerate, R_z is not singular. By (7) and (9), the MSE per dimension of the lattice quantizer may be rewritten as $\epsilon = \text{trace}\{R_z\}/K$.

Definition 1: A lattice quantizer is <u>white</u> if the samples of its quantization noise are uncorrelated and have the same power, i.e.

$$R_z = \epsilon \cdot I \tag{10}$$

where I is the identity matrix.

The normalized second moment of the lattice $G(Q_K)$ can be expressed in terms of R_z as

$$G(Q_K) = \frac{1}{K} \cdot \frac{E\|\underline{Z}\|^2}{V^{2/K}} = \frac{1}{V^{2/K}} \cdot \frac{\text{trace}\{R_z\}}{K}. \tag{11}$$

Furthermore, by the arithmetic–geometric means inequality (see, e.g., [5, Theorem 16.8.41]), $\text{trace}\{R_z\}/K \geq |R_z|^{1/K}$, and thus $G(Q_K)$ is lower-bounded by

$$G(Q_K) \geq \frac{1}{V^{2/K}} \cdot |R_z|^{1/K} \tag{12}$$

where $|\cdot|$ denotes determinant. Equality holds if and only if R_z is a diagonal matrix with identical elements on the diagonal, i.e., if and only if Q_K is a *white lattice quantizer.*

The *optimal* lattice quantizer in $\mathcal{R}^K$, denoted Q_K^{opt}, is the lattice quantizer with the minimal possible normalized second moment G_K^{opt}, where the minimization is taken over all lattices and their possible associated partitions [4]. Our main result in this section is given by the following theorem:

Theorem 1: Q_K^{opt} is white, and the autocorrelation of its quantization noise is

$$R_z = G_K^{\text{opt}} \cdot V^{2/K} \cdot I. \tag{13}$$

This theorem implies, for example, that the hexagonal lattice quantizer (A_2^*) and the body-centered cubic lattice quantizer (A_3^*), which are known to be optimal for $K = 2$ and 3, respectively, are *white*.

To prove this Theorem we need to introduce the notion of "shaping" a (lattice) quantizer:

Definition 2: The <u>shaping</u> of a quantizer Q_K by a $K \times K$ nonsingular matrix A, is the quantizer Q'_K for which

$$Q'_K(\underline{x}) = A \cdot Q_K(A^{-1}\underline{x}).$$

It is easy to verify that if Q_K in the above definition is a lattice quantizer, then so is the shaped quantizer $Q'_K = \{\mathcal{L}', \mathcal{P}'\}$, whose code points and quantization cells are

$$l'_i = A l_i \quad \text{and} \quad P'_i = \{\underline{x} : A^{-1}\underline{x} \in P_i\} \tag{14}$$

respectively. For example, Partition C (the rhombuses) in Fig. 1 was generated by shaping the $\mathcal{Z}$ lattice and its associated Voronoi regions (which are squares), using the transformation

$$A = \begin{pmatrix} 1 & 1/2 \\ 0 & \sqrt{3}/2 \end{pmatrix}.$$

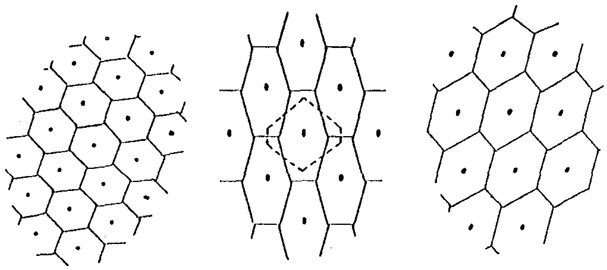

Fig. 2. Various shapings of the hexagonal-lattice / hexagonal-cells quantizer: (a) rotation, (b) scaling to north direction, (c) scaling to north–east direction.

We observe the interesting fact that the transformed partition $\mathcal{P}'$ *is not the Voronoi partition* associated with the transformed lattice $\mathcal{L}'$. The Voronoi regions associated with the lattice in the example of Fig. 1 are, of course, hexagons (Partition A). The reason is that, in general, $||\underline{x} - \underline{l}_i|| \leq ||\underline{x} - \underline{l}_j||$ does not imply $||A(\underline{x} - \underline{l}_i)|| \leq ||A(\underline{x} - \underline{l}_j)||$, unless A is proportional to some measure preserving transformation (i.e., if and only if $AA^t \propto I$) [9, sec. 8.58], and thus $A\underline{l}_i$ is not necessarily the nearest code point after the transformation. Fig. 2 illustrates various shapings of the hexagonal-lattice / hexagonal-cells quantizer. The dashed line in Fig. 2(b) represents the Voronoi cell of the transformed lattice.

Following well-known formulas from vector analysis, the volume of the shaped quantizer cells is

$$V' = \int_{P_0'} d\underline{x} = |A| \cdot \int_{P_0} d\underline{x} = |A| \cdot V$$

and the noise vector of the shaped quantizer, $\underline{Z}' \sim \mathcal{U}(P_0')$, equals in distribution to $A\underline{Z}$. This implies

$$R_z' = AR_zA^t \tag{15}$$

i.e., the transformation A *shapes* the spectrum of the quantization noise.

We can now proceed and prove Theorem 1.

Proof: Let R_z be the autocorrelation of the quantization noise associated with the optimal quantizer Q_K^{opt}. We show below that if R_z is nonwhite we can get a better quantizer. But we assume that Q_K^{opt} is optimal. Thus R_z must be white.

Specifically, let

$$A = \sqrt{\epsilon'} \cdot R_z^{-1/2} \tag{16}$$

where $R_z^{-1/2}$ is an inverse root of R_z and ϵ' is a positive scalar, i.e., $AA^t = \epsilon' \cdot R_z^{-1}$. As can be seen by substituting A into (15), shaping Q_K^{opt} by A results in a white lattice quantizer Q_K', for which the autocorrelation of the quantization noise is $R_z' = \epsilon' I$, and the cells volume is

$$V' = |A| \cdot V = \sqrt{(\epsilon')^K / |R_z|} \cdot V.$$

By substituting into (11), we get

$$G(Q_K') = \frac{1}{(V')^{2/K}} \cdot \frac{\text{trace}\{\epsilon' I\}}{K} = \frac{|R_z|^{1/K}}{V^{2/K}} \leq G(Q_K^{\text{opt}}) \tag{17}$$

where the inequality follows from (12), with equality if and only if Q_K^{opt} is white. □

Let us discuss some consequences of the result above. First, it follows from (16), (17), and the optimality of the Voronoi partition, that by iterating between *whitening* and *Voronoi partitioning* one monotonically reduces (i.e., improves) the normalized second moment of a given lattice quantizer. Actually, Theorem 1 adds another necessary condition to the well-known Lloyd conditions [13], [15] for the optimal quantizer.

Second, the shaping procedure implies that by an appropriate linear transformation of a (nondegenerate) lattice quantizer, any desired quantization noise spectra may be obtained. The optimal quantizer in the sense of minimizing the square error is white; however, a nonwhite lattice quantizer obtained by shaping Q_K^{opt} using a transformation A is optimal, for example, under a *weighted square* error criterion, i.e., minimizing the expected value of $(\hat{\underline{x}} - \underline{x})^t W (\hat{\underline{x}} - \underline{x})$, where $AA^t = W^{-1}$. Nonwhite lattice quantizers also occur when pre/post-filters are incorporated in the coding process (see, e.g., [23] or any recent article on the CELP technique for speech coding).

Finally, this theorem can be extended to infinite-dimensional lattice quantizers, i.e., to trellis-coded quantizers (TCQ) [3], [16] having a linear structure. We denote by Q_∞ the trellis quantizer, and we replace the matrix A by the discrete time invariant shaping filter $a = \{a_n\}$, $n = 0, \pm 1, \pm 2, \cdots$, whose frequency response is

$$A(f) = \sum_n a_n \cdot e^{-jn2\pi f}, \ -1/2 \leq f \leq 1/2.$$

The shaped infinite lattice quantizer is now

$$Q_\infty'(x) = a * Q_\infty(a^{(-1)} * x)$$

where x is a source sequence, $a^{(-1)} = \{a_n^{(-1)}\}$ are the coefficients of the inverse filter (i.e., the inverse Fourier transform of $1/A(f)$), and $*$ denotes the convolution operator. Spectral shaping will also occur when Q_∞ is in the feedback loop, resembling predictive-TCQ [16]. Now, let $Z = \cdots Z_{-1}, Z_0, Z_1, \cdots$, denote the (stationary) quantization noise process of Q_∞. As shown in the Appendix, in some sense Z is uniformly distributed over the infinite-dimensional basic cell of Q_∞. Clearly, if the quantization noise spectrum is $S_z(f)$, then after shaping it becomes $|A(f)|^2 S_z(f)$. As discussed in the Appendix, the normalized second moment of Q_∞ is defined as

$$G(Q_\infty) \triangleq (2\pi e)^{-1} \frac{\epsilon}{P(Z)}$$

where ϵ is the power of Z, and $P(Z)$ is the entropy rate power of Z. This definition coincides with (7) in the finite-dimensional case, since then $P(\underline{Z}) = (2\pi e)^{-1} \cdot V^{2/K}$. Thus if we shape a white trellis quantizer (i.e., a trellis quantizer whose noise spectrum is flat) we get $\epsilon' = \epsilon \cdot \int |A(f)|^2 df$, and

$$P(Z') = P(Z) \cdot \exp\left(\int \ln |A(f)|^2 \, df \right)$$

and thus

$$G(Q_\infty') = G(Q_\infty) \cdot \frac{\int |A(f)|^2 \, df}{\exp\left(\int \ln |A(f)|^2 \, df\right)}$$

which, by Jensen's inequality, is greater than $G(Q_\infty)$ with equality if and only if $A(f)$ is white.

III. The Divergence from Gaussianity of the Quantization Noise

In the preceding section we investigated the spectral properties of the lattice quantization noise. This section addresses the *distribution* of the lattice quantization noise, and specifically it focuses on its distance from a Gaussian distribution. A Gaussian quantization noise is desirable, at least for Gaussian sources and MSE criterion, since it resembles the form of the noise in the forward-channel realization of the rate-distortion function [2]. Furthermore, Gaussian quantization noise corresponds to an efficient covering of high-dimensional spaces.

In our context, the natural measure for the distance of the distribution of the quantization noise from a Gaussian distribution is the *information divergence*, also called "relative entropy" or "Kullback–Leibler distance"; see [5].

Definition 3: The divergence from Gaussianity of a vector $\underline{U} \in \mathcal{R}^n$ with a density $f_{\underline{U}}$, is

$$\mathcal{D}(\underline{U};\underline{U}^*) = \mathcal{D}(f_{\underline{U}}\|f_{\underline{U}^*}) = \int_{\mathcal{R}^n} f_{\underline{U}} \log \frac{f_{\underline{U}}}{f_{\underline{U}^*}} = h(\underline{U}^*) - h(\underline{U}) \tag{18}$$

where $h(\cdot)$ denotes differential entropy. This is the divergence between $\underline{U}$ and the Gaussian vector $\underline{U}^*$ having the same mean and covariance as $\underline{U}$. Similarly, the divergence of $\underline{U}$ from white Gaussianity is defined as

$$\mathcal{D}(\underline{U};\underline{W}) = h(\underline{W}) - h(\underline{U}) = \frac{n}{2}\log(2\pi e\epsilon) - h(\underline{U}) \tag{19}$$

where $\underline{W}$ is a white Gaussian vector with the same average power, ϵ, as $\underline{U}$.

Note that

$$\mathcal{D}(\underline{U};\underline{W}) = \mathcal{D}(\underline{U};\underline{U}^*) + \mathcal{D}(\underline{U}^*;\underline{W}) \geq \mathcal{D}(\underline{U};\underline{U}^*)$$

with equality if and only if the autocorrelation of $\underline{U}$ is white. Throughout the sequel the logarithm base is 2, and information quantities are measured in bits. The divergence in (18) can be written as a difference between entropies since

$$h(\underline{U}^*) \triangleq -\int f_{\underline{U}^*}\log f_{\underline{U}^*} = -\int f_{\underline{U}}\log f_{\underline{U}^*}$$

(see [5, p. 234.]), and a similar identity justifies (19).

We turn now to evaluate the divergence from Gaussianity of the lattice quantization noise. Since $\underline{Z}$ is uniformly distributed over a region whose volume is V, we may use (7) and write the entropy per dimension of $\underline{Z}$ as

$$\frac{1}{K}h(\underline{Z}) = \frac{1}{K}\log V = \frac{1}{2}\log(\epsilon/G(Q_K)). \tag{20}$$

Now, let $\underline{W} \sim \mathcal{N}(\underline{0}, \epsilon I)$ be a white Gaussian vector with power ϵ. By (19) and (20), the divergence of $\underline{Z}$ from white Gaussianity per dimension is given by

$$\begin{aligned}\frac{1}{K}\mathcal{D}(\underline{Z};\underline{W}) &= \frac{1}{2}\log(2\pi e\epsilon) - \frac{1}{2}\log(\epsilon/G(Q_K)) \\ &= \frac{1}{2}\log(2\pi e G(Q_K))\end{aligned} \tag{21}$$

bits per sample. Note that shaping the lattice quantizer does not affect its divergence from Gaussianity $\mathcal{D}(\underline{Z};\underline{Z}^*)$, since the divergence is invariant to invertible transformation of its arguments. But $\mathcal{D}(\underline{Z};\underline{W})$ does vary with shaping, and attains its minimum, $\mathcal{D}(\underline{Z};\underline{Z}^*)$, when Q_K is shaped by the whitening transformation A given in (16). The figure $\frac{1}{2}\log(2\pi e G(Q_K))$ is about 0.254 bits for uniform scalar quantizer (where $G(Q_1) = 1/12$), and is about 0.227 bits for the two-dimensional hexagonal quantizer. We show next that for the optimal lattice quantizers $G_K^{\rm opt} \to \frac{1}{2\pi e}$, as $K \to \infty$, and thus for these quantizers

$$\frac{1}{K}\mathcal{D}(\underline{Z};\underline{W}) \to 0, \quad \text{as } K \to \infty. \tag{22}$$

The interpretation of this result is that the distribution of the quantization noise converges to a white Gaussian distribution in the *divergence sense*.

A. The Asymptotic Normalized Second Moment of Optimal Lattice Quantizers

It is well known that a ball has the minimal moment of inertia among all shapes of equal volume. A simple application of this property to (6) provides the following lower bound on $G(Q_K)$:

$$\begin{aligned}G(Q_K) \geq G_K^* &\triangleq \frac{1}{K}\frac{\int_{S_K}\|\underline{x}\|^2\,d\underline{x}}{V^{1+2/K}} \\ &= \frac{1}{\pi(K+2)}\cdot\Gamma^{\frac{2}{K}}(K/2+1) \\ &= \frac{1}{(K+2)\cdot V_K^{2/K}}\end{aligned} \tag{23}$$

where S_K is a K-dimensional ball with volume V (i.e., with the same volume as P_0), G_K^* denotes the normalized second moment of S_K [4, p. 452.], $\Gamma(\cdot)$ is the Gamma function, and

$$V_K = \frac{\pi^{K/2}}{\Gamma(K/2+1)}$$

is the volume of a ball with a unit radius. From (23) we can compute $G_1^* = \frac{1}{12}$, and realize that as $K \to \infty$, G_K^* decreases monotonically to the limit $\frac{1}{2\pi e} \approx \frac{1}{17}$, at a rate

$$\log(2\pi e G_K^*) = O\left(\frac{\log K}{K}\right).$$

We claim that $1/2\pi e$ is also the limit of $G_K^{\rm opt}$:

Lemma 1 (Poltyrev):

$$G_K^{\rm opt} \to \frac{1}{2\pi e}, \quad \text{as } K \to \infty \tag{24}$$

at a rate

$$\log\left(2\pi e G_K^{opt}\right) = O\left(\frac{\log K}{K}\right). \tag{25}$$

Lemma 1 was originally inferred from the work of Zador [20] and a conjecture made by Gersho in [8]. The bounds obtained by Zador implied that the average normalized second moment of the cells of the optimal K-dimensional quantizer (for uniformly distributed data) goes to $\frac{1}{2\pi e}$ as $K \to \infty$.

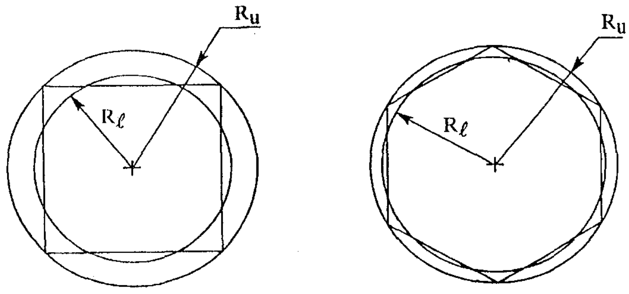

Fig. 3. The radii R_u and R_l for a square (a), and for a hexagon (b).

However, Zador did not give any characterization of these quantizers. Later on, Gersho claimed (without a proof) that this optimal K-dimensional quantizer is a tessellating quantizer, implying by Zador bounds that $\frac{1}{2\pi e}$ is the asymptotic normalized second moment of the best K-dimensional space-filling polytope. Lemma 1, whose proof below is due to Poltyrev [19], shows that this minimal possible G_K is actually asymptotically achievable by the basic cells of a sequence of lattice quantizers.

Proof: Let R_u denote the covering radius of P_0, and R_l denote the radius of a ball having the same volume as P_0, i.e., $V = V_K \cdot R_l^K$; see Fig. 3. For any lattice quantizer Q_K we can write

$$G(Q_K) \leq \frac{R_u^2}{K \cdot V^{2/K}} = \left(\frac{R_u}{R_l}\right)^2 \frac{1}{K \cdot V_K^{2/K}}. \tag{26}$$

To see this, note that from (11)

$$G(Q_K) = \frac{1}{K} E||\underline{Z}||^2 / V^{2/K}$$

where $\underline{Z}$ is uniformly distributed over P_0, and, since $||\underline{z}|| \leq R_u$ for every $\underline{z} \in P_0$, we have immediately $E||\underline{Z}||^2 \leq R_u^2$. Using (23) and (26), we can bound $G(Q_K)$ by

$$G_K^* \leq G(Q_K) \leq G_K^* \cdot \frac{K+2}{K} \cdot \Theta^{2/K} \tag{27}$$

where $\Theta = (R_u/R_l)^K$ is the *density of the covering* of $\mathcal{R}^K$ by the lattice quantizer [4, p. 312].

Now, following a result by Rogers [4, p. 392], there exist lattices whose covering density (with respect to their Voronoi partition) satisfies

$$\Theta \leq cK(\log K)^a \tag{28}$$

for some constants c and a. This implies that for these lattices $\Theta^{1/K} \rightarrow 1$ as $K \rightarrow \infty$ and

$$\frac{1}{K} \log \Theta = O\left(\frac{\log K}{K}\right).$$

Substituting into (27), and observing that $\log(K \cdot V_K^{2/K} / 2\pi e)$ goes to zero as $O\left(\frac{\log K}{K}\right)$, yields the desired result. □

B. Convergence of Noise Blocks to Gaussianity

The proof of Lemma 1 shows that in some sense the Voronoi regions of the optimal lattice quantizers become closer to *balls*. This observation provides a simple geometric explanation of the Gaussian limit behavior of the quantization noise, since, by Poincaré theorem (see, e.g., [6]), the projection on any finite-dimensional hyperspace of a uniform distribution over a ball, becomes, in the limit as the ball dimension increases, a distribution of an i.i.d. Gaussian vector.

By utilizing (22) we can provide a statistical characterization of this limit behavior. Let $\oplus$ denote a modulo-K addition, and let

$$Z_{\theta\oplus 1}^{\theta\oplus m} = (Z_{\theta\oplus 1}, \cdots, Z_{\theta\oplus m})$$

denote an m-block starting at $\theta \oplus 1$ of the vector $\underline{Z}$ associated with the quantizer Q_K. The density of $Z_{\theta\oplus 1}^{\theta\oplus m}$ is denoted $f_{Z_{\theta\oplus 1}^{\theta\oplus m}}$. Let

$$\phi(\underline{z}) = \frac{1}{(2\pi\epsilon)^{m/2}} e^{-||\underline{z}||^2/2\epsilon}$$

denote the density of a zero-mean m-dimensional Gaussian vector with an autocorrelation ϵI, where $\epsilon = \frac{1}{K} E\{||\underline{Z}||^2\}$ is the MSE per dimension of the lattice quantizer. Finally, let

$$||f - \phi||_1 \triangleq \int_{\mathcal{R}^m} |f(\underline{z}) - \phi(\underline{z})| \, d\underline{z}$$

denote the $\mathcal{L}_1$ distance between the m-variate densities f and ϕ. Now, if we could assume that the quantization noise samples $Z_1, \cdots, Z_K$ are circular-stationary, then (22) would imply that for *every* sub-block of size m starting at θ

$$||f_{Z_{\theta\oplus 1}^{\theta\oplus m}} - \phi||_1 \rightarrow 0, \quad \text{as } K \rightarrow \infty.$$

However, this assumption, in fact, is not true and so we show this claim only on the average over all starting points θ. This result will follow as an immediate corollary from Lemma 1 and the following theorem:

Theorem 2: Let $\underline{Z}$ be uniform over the basic cell of Q_K. Then, for any $1 \leq m \leq K$

$$\frac{1}{K} \sum_{\theta=0}^{K-1} ||f_{Z_{\theta\oplus 1}^{\theta\oplus m}} - \phi||_1^2 \leq m \cdot \ln(2\pi e G(Q_K)) \tag{29}$$

where $||\cdot||_1^2$ denotes square $\mathcal{L}_1$ distance, and $\ln(\cdot)$ denotes natural logarithm.

Corollary 1: For a fix m and the sequence of optimal lattice quantizers

$$\frac{1}{K} \sum_{\theta=0}^{K-1} ||f_{Z_{\theta\oplus 1}^{\theta\oplus m}} - \phi||_1^2 \rightarrow 0, \quad \text{as } K \rightarrow \infty \tag{30}$$

i.e., $||f_{Z_{\Theta\oplus 1}^{\Theta\oplus m}} - \phi||_1 \rightarrow 0$ in the Mean Square (MS) sense over $\Theta \sim \mathcal{U}(0 \cdots K-1)$.

Since convergence in the MS sense implies convergence in probability, most of the dither sub-blocks tend to have a white Gaussian density in the $\mathcal{L}_1$ sense in the limit of large lattice dimension, i.e.

$$\text{Prob}\{\Theta : ||f_{Z_{\Theta\oplus 1}^{\Theta\oplus m}} - \phi||_1 > \delta\} \rightarrow 0 \; \forall \delta > 0.$$

Proof: Let $\Theta \sim \mathcal{U}(0,\cdots,K-1)$. Let $W_i \sim \mathcal{N}(0,\epsilon)$, $i=1,\cdots,m$ be i.i.d. Gaussian variables, independent of Θ, whose variance ϵ is the MSE per dimension of Q_K. Note that $W_1^m \sim \phi$. We have the following chain of identities and an inequality:

$$\begin{aligned}\frac{1}{K}\sum_{\theta=0}^{K-1}\mathcal{D}(Z_{\theta\oplus 1}^{\theta\oplus m};W_1^m) &\triangleq \mathcal{D}(Z_{\Theta\oplus 1}^{\Theta\oplus m};W_1^m|\Theta)\\ &\overset{(a)}{=} \frac{m}{2}\log(2\pi e\epsilon) - h(Z_{\Theta\oplus 1}^{\Theta\oplus m}|\Theta)\\ &= \frac{m}{2}\log(2\pi e\epsilon) - \frac{1}{K}\sum_{\theta=0}^{K-1}h(Z_{\theta\oplus 1}^{\theta\oplus m})\\ &\overset{(b)}{\leq} \frac{m}{2}\log(2\pi e\epsilon) - \frac{m}{K}h(Z_1^K)\\ &\overset{(c)}{=} \frac{m}{2}\log(2\pi e G(Q_K))\end{aligned} \tag{31}$$

where $\mathcal{D}(\cdot;\cdot|\Theta)$ and $h(\cdot|\Theta)$ denote conditional divergence and conditional entropy given Θ [5, secs. 2.5, 9.4], and where (a) follows from a conditional version of the decomposition in (19), since

$$E\{\|Z_{\Theta\oplus 1}^{\Theta\oplus m}\|^2\} = \frac{m}{K}\sum_{\theta=1}^{K}E\{Z_\theta^2\} = m\epsilon$$

where the expectation is taken over Θ and Z (note that for an *optimal* lattice quantizer it follows from Theorem 1 that $E\{Z_i^2\}=\epsilon$ for each i); (b) follows since, similarly to a theorem by Han [5, Theorem 16.5.1.]

$$\frac{1}{K}h(Z_1^K) \leq \frac{1}{K}\sum_{\theta=0}^{K-1}\frac{1}{m}h(Z_{\theta\oplus 1}^{\theta\oplus m}) \tag{32}$$

(the normalized joint entropy of a set is upper-bounded by the normalized sum of the entropies of circular shifted subsets); and (c) follows from (20).

The desired result (29) now follows from [5, Lemma 16.3.1], which states that the divergence upper-bounds the square of the $\mathcal{L}_1$ distance, i.e.

$$\mathcal{D}(Z_{\theta\oplus 1}^{\theta\oplus m};W_1^m) \geq \frac{1}{2\ln 2}\|f_{Z_{\theta\oplus 1}^{\theta\oplus m}} - \phi\|_1^2. \tag{33}$$

□

C. The Equivalence in Information of the AWGN Channel and the ECDQ

We recall from [21] and [22] that the coding rate of Entropy Coded Dithered Quantization (ECDQ) is given by the mutual information in an additive noise channel, whose input is the quantized source, and where the noise is independent of the source and is uniformly distributed over the mirror image of the lattice quantizer basic cell. Using Lemma 1 and Corollary 1, it is shown below that the coding rate of ECDQ tends, in the limit of large (optimal) lattice dimension, to the mutual information in an Additive White Gaussian Noise (AWGN) channel. This limit behavior is utilized in [22] to show a simple tradeoff between the sampling rate and the quantization resolution in ECDQ of band-limited sources, and in [23] to achieve the rate-distortion function in pre/post-filtered ECDQ of Gaussian sources.

Let $X = X_1, X_2, \cdots$ be a discrete-time stationary source. Let $Z^{(K)} = Z_1^{(K)}, Z_2^{(K)}, \cdots$ be a process obtained by concatenating i.i.d. blocks $\underline{Z} \sim \mathcal{U}(P_0)$. Namely, $Z^{(K)}$ is the quantization noise process associated with successive quantizations by Q_K. Finally, let $W = W_1, W_2, \cdots$ be a white Gaussian process with power ϵ. We denote by

$$\bar{I}(X;Y) \triangleq \lim_{n\to\infty}\frac{1}{n}I(X_1\cdots X_n; Y_1\cdots Y_n) \tag{34}$$

the *mutual information rate* per sample between the jointly stationary processes X and Y, assuming the above limit exists. Note that $\bar{I}(X;X+Z^{(K)})$ always exists for a stationary X independent of $Z^{(K)}$, since the entropy rate of $Z^{(K)}$ is finite [22].

Theorem 3: Assume that $Z^{(K)}$ is the noise process associated with the optimal K-dimensional lattice quantizer. Then, if X is a stationary Gaussian source

$$\lim_{K\to\infty}\bar{I}(X;X+Z^{(K)}) = \bar{I}(X;X+W). \tag{35}$$

Furthermore, there exist a sub-sequence of lattice dimensions $K_1, K_2, \cdots$, such that for <u>any</u> stationary source X

$$\limsup_{n\to\infty}\bar{I}(X;X+Z^{(K_n)}) \leq \bar{I}(X;X+W). \tag{36}$$

In this theorem, the result for Gaussian sources (35) is slightly stronger than the result we were able to prove for general sources (36). In addition, the proof in the Gaussian case is much simpler and follows directly from Lemma 1. It is conjectured, though, that (35) holds for non-Gaussian sources as well [18].

Proof: We start with the proof of the Gaussian source case, which is straightforward. As was shown in Section I, the optimal lattice quantizers are white, i.e., $E\{Z_i\}=0$, and $E\{Z_iZ_j\} = \epsilon\cdot\delta_{ij}$, for all i and j. Thus $\left(Z^{(K)}\right)^* = W$ for all K, where, as in (18), $(\cdot)^*$ denotes a Gaussian process having the same mean and correlation function. Let

$$\bar{h}(X) \triangleq \lim_{n\to\infty} 1/n\cdot h(X_1\cdots X_n) \tag{37}$$

denote entropy rate of a stationary process, and recall that the mutual information rate in the additive noise channel $X \to X+N$ is given by

$$\bar{I}(X;X+N) = \bar{h}(X+N) - \bar{h}(N). \tag{38}$$

Utilizing known bounds for the mutual information in an additive noise channel fed with a Gaussian source [5, p. 2631], and the relation

$$\bar{h}(W) - \bar{h}(Z^{(K)}) = \frac{1}{K}\mathcal{D}(Z_1^K;W_1^K)$$

we may write

$$\begin{aligned}\bar{I}(X;X+W) &\leq \bar{I}(X;X+Z^{(K)})\\ &\leq \bar{I}(X;X+W) + \frac{1}{K}\mathcal{D}(Z_1^K;W_1^K).\end{aligned} \tag{39}$$

By (22), the upper and lower bounds coincide asymptotically, and (35) is proved.

We now consider the general case. Writing the information rate $\overline{I}(X; X+Z^{(K)})$ as a difference between entropy rates as in (38), we immediately see from (20) and (24) that

$$\overline{h}(Z^{(K)}) \rightarrow \frac{1}{2}\log 2\pi e\epsilon = \overline{h}(W).$$

Thus we only need to show that the upper limit of $\overline{h}(X+Z^{(K)})$ is upper-bounded by $\overline{h}(X+W)$. For the ease of notation, let $\underline{Z}_\theta$ denote $Z_{\theta\oplus 1}^{\theta\oplus m}$. Since (30) implies convergence in probability over $\Theta \sim \mathcal{U}(0\cdots K-1)$, it follows from the Borel–Canteli Lemma [7, p. 403] that there exist some (monotonically increasing) sub-sequence of indices $K_1, K_2, \cdots$ for which the convergence in (30) is in probability 1, i.e., for any block length m

$$\lim_{n\rightarrow\infty} \|f_{\underline{Z}_\Theta}^{(K_n)} - \phi\|_1 = 0$$

for almost every realization of initializations $\Theta = \Theta(K)$. This implies that for any block of size m taken from X, the $\mathcal{L}_1$ distance between the density functions of $\underline{X} + \underline{Z}_\Theta^{(K_n)}$ and $\underline{X} + \underline{W}$ converges to zero *almost surely*, as $n \rightarrow \infty$. Note that all types of convergence in n which are used throughout this proof are with respect to the random initializations Θ. Thus by the semi-continuity of the divergence (see, e.g., [17])

$$\liminf_{n\rightarrow\infty} \mathcal{D}(\underline{X}+\underline{Z}_\Theta^{(K_n)}; \underline{X}^*+\underline{W}) \geq \mathcal{D}(\underline{X}+\underline{W}; \underline{X}^*+\underline{W}) \quad \text{a.s.} \tag{40}$$

where, of course, (40) holds also *in probability*. Since the divergence is nonnegative, (40) must hold also on the average over Θ, i.e.

$$\liminf_{n\rightarrow\infty} \mathcal{D}(\underline{X}+\underline{Z}_\Theta^{(K_n)}; \underline{X}^*+\underline{W}|\Theta) \geq \mathcal{D}(\underline{X}+\underline{W}; \underline{X}^*+\underline{W}).$$

Since $\left(Z^{(K)}\right)^* = W$ for all K, we utilize the decomposition (18) of the divergence from Gaussianity into entropies and get, for any sub-block of size m

$$\limsup_{n\rightarrow\infty} h(\underline{X}+\underline{Z}_\Theta^{(K_n)}|\Theta) \leq h(\underline{X}+\underline{W}) \tag{41}$$

where $h(\cdot|\Theta)$ denotes conditional entropy given Θ.

Now, since X is a stationary process, $Z^{(K)}$ is a block-stationary (cyclostationary) process with a period K, and Θ is uniformly distributed over the period of $Z^{(K)}$, it is easy to show using (32) that $\frac{1}{m}h(\underline{X}+\underline{Z}_\Theta^{(K)}|\Theta)$ is monotonically decreasing with the block size m to the limit $\overline{h}(X+Z^{(K)})$, that is, to the entropy rate of $X+Z^{(K)}$. Thus

$$\frac{1}{m}h(\underline{X}+\underline{Z}_\Theta^{(K)}|\Theta) \geq \overline{h}(X+Z^{(K)})$$

for every block size m, implying by (41)

$$\limsup_{n\rightarrow\infty} \overline{h}(X+Z^{(K_n)}) \leq \overline{h}(X+W) \tag{42}$$

which proves the theorem. □

Appendix
More on Q_∞

In this Appendix we discuss quantities associated with the quantization noise of a linear trellis quantizer, which is an infinite-dimensional lattice quantizer whose structure is time-invariant. Let

$$P_0^{(K)} = \{x_1\cdots x_K : Q_\infty(x) = \cdots 0,0,0\cdots, \text{ for some } x = \cdots x_{-1}, x_0, x_1\cdots x_K, x_{K+1}, \cdots\}$$

be the projection of the basic cell of Q_∞ on the first K coordinates, and let $V^{(K)} = \mu(P_0^{(K)})$ be its volume. Define the quantity

$$\overline{V} = \lim_{K\rightarrow\infty} \sqrt[K]{V^{(K)}}.$$

Since the quantizer is time-invariant, the sequence $\sqrt[K]{V^{(K)}}$ is monotonically decreasing[1] with K, implying that the limit above exists. $\overline{V}$ may be interpreted as the effective volume per dimension of $P_0^{(\infty)}$, the infinite-dimensional basic cell of Q_∞.

Consistently with our model of the quantization noise, the quantization noise process $Z = \cdots Z_{-1}, Z_0, Z_1\cdots$ is associated with trellis quantization of a uniformly distributed source. Particularly, $Z_i = Q_\infty(X)_i - X_i$, where

$$X_i \sim \mathcal{U}(-L/2, L/2), \quad i = 0, \pm 1, \pm 2, \cdots$$

are i.i.d., and where L is very large compared to the size of the quantizer's cells, so that edge effects may be neglected. Note that since the trellis structure is time-invariant, the quantization noise Z is a stationary process. Clearly, for each K, $(Z_1\cdots Z_K) \in P_0^{(K)}$, and for large K all blocks $(Z_1\cdots Z_K)$ are equally likely. This implies that the entropy rate of the quantization noise is

$$\overline{h}(Z) \triangleq \lim_{K\rightarrow\infty}\frac{1}{K}h(Z_1\cdots Z_K) = \lim_{K\rightarrow\infty}\frac{1}{K}\log V^{(K)} = \log\overline{V}$$

and hence its entropy rate power is [2]

$$P(Z) \triangleq \frac{1}{2\pi e}2^{2\overline{h}(Z)} = \frac{1}{2\pi e}\overline{V}^2 = \frac{1}{2\pi e}\frac{\epsilon}{G(Q_\infty)}.$$

When a K-dimensional lattice quantizer is shaped by a matrix A, its second moment is multiplied by trace$\{AA^t\}/K$, and the volume of its basic cell is multiplied by $|A| = |AA^t|^{1/2}$, or equivalently, the entropy power of a uniform distribution over its basic cell is multiplied by $|AA^t|^{1/K}$. In other words, the second moment and the entropy power are multiplied by the arithmetic and the geometric averages, respectively, of the eigenvalues of the matrix AA^t. Analogously to that, it follows from the Toeplitz Distribution Theorem [2, p. 112] that when Q_∞ is shaped by a filter $A(f)$, $-1/2 \leq f \leq 1/2$, the second moment and the entropy rate power of its quantization noise are multiplied by $\int |A(f)|^2\,df$, and $\exp\left(\int \ln|A(f)|^2\,df\right)$, respectively.

[1] This property is analogous to the behavior of the entropy rate of blocks of a stationary process, and it is shown easily using the notion of "conditional volume" defined in [24].

ACKNOWLEDGMENT

Lemma 1 is due to Gregory Poltyrev. The authors wish to thank S. Litsyn for discussions regarding its proof. The authors are also greatful to M. Pinsker for helpful discussions regrading Theorems 2 and 3.

REFERENCES

[1] W. R. Bennett, "Spectra of quantized signals," *Bell Syst. Tech. J.*, vol. 27, pp. 446–472, July, 1948.
[2] T. Berger, *Rate Distortion Theory: A Mathematical Basis for Data Compression.* Englewood Cliffs, NJ: Prentice-Hall, 1971.
[3] A. R. Calderbank, P. C. Fishburn, and A. Rabinovich, "Covering properties of convolutional codes and associated lattices," *IEEE Trans. Inform. Theory*, vol. 41, no. 3, pp. 732–746, May 1995.
[4] J. H. Conway and N. J. A. Sloane, *Sphere Packings, Lattices and Groups.* New York: Springer-Verlag, 1988.
[5] T. M. Cover and J. A. Thomas, *Elements of Information Theory.* New York: Wiley, 1991.
[6] P. Diaconis and D. Freedman, "A dozen de-Finetti-style results in search for a theory," *Ann. Inst. H. Poincaré, Prob. Stat.*, vol. 23, suppl. 2, pp. 397–423, 1987.
[7] R. Durrett, *Probability Teory and Examples.* Pacific Grove, CA: Wadsworth and Brooks, 1989.
[8] A. Gersho, "Asymptotically optimal block quantization," *IEEE Trans. Inform. Theory*, vol. IT-25, pp. 373–380, July 1979.
[9] A. Gersho and R. M. Gray. *Vector Quantization and Signal Compression.* Boston, MA: Kluwer, 1992.
[10] R. M. Gray and Jr. T. J. Stockham, "Dithered quantizers," *IEEE Trans. Inform. Theory*, vol. 39, pp. 805–812, May 1993.
[11] A. Kirac and P. P. Vaidyanathan, "Results on lattice quantization with dithering," *IEEE Trans. Circuits Syst.*, submitted.
[12] D. H. Lee and D. L. Neuhoff, "Asymptotic distribution of the errors in scalar and vector quantizers," *IEEE Trans. Inform. Theory*, vol. 42, no.2, pp. 446–460, March 1996.
[13] Y. Linde, A. Buzo, and R. M. Gray, "An algorithm for vector quantizer design," *IEEE Trans. Commun.*, vol. COM-28, pp. 84–95, Jan. 1980.
[14] T. Linder and K. Zeger, "Asymptotic entropy constrained performance of tesselating and universal randomized lattice quantization," *IEEE Trans. Information Theory*, vol 40, pp. 575–579, Mar. 1994.
[15] S. P. Lloyd, "Least squares quantization in PCM," *IEEE Trans. Inform. Theory*, vol. IT-28, pp. 129–137, Mar. 1982 (originally presented at the Inst. of Math. Stat. Meet., 1957).
[16] M. W. Marcellin and T. R. Fischer, "Trellis coded quantization of memoryless and Gauss–Markov sources," *IEEE Trans. Commun.*, vol. 38, pp. 82–93, Jan. 1990.
[17] M. S. Pinsker, *Information and Information Stability of Random Variables and Processes.* San Francisco. CA: Holden Day, 1964.
[18] ——, private communications.
[19] G. Poltyrev, private communications.
[20] P. Zador, "Asymptotic quantization error of continuous signals and their quantization dimension," *IEEE Trans. Inform. Theory*, vol IT-28, pp. 139–149, Mar. 1982 (previously an unpublished Bell Labs. memo, 1966).
[21] R. Zamir and M. Feder, "On universal quantization by randomized uniform / lattice quantizer," *IEEE Trans. Inform. Theory*, vol. 38, pp. 428–436, Mar. 1992.
[22] ——, "Rate distortion performance in coding band-limited sources by sampling and dithered quantization," *IEEE Trans. Inform. Theory*, vol. 41, pp. 141–154, Jan. 1995.
[23] ——, "Information rates of pre/post filtered dithered quantizers," *IEEE Trans. Inform. Theory*, to be published in vol. 42, no.5, Sept. 1996.
[24] ——, "A matrix form of the Brunn-Minkowski inequality," in *Proc. Int. Symp. on Information Theory (ISIT-95),*(Whistler, BC, Canada, Sept. 1995), p. 71.