Gaussian Source Coding with Spherical Codes*

Jon Hamkins and Kenneth Zeger

IEEE Transactions on Information Theory
Submitted: February 20, 2001
Revised: April 10, 2002
Final version: June 4, 2002

Abstract

A fixed rate shape-gain quantizer for the memoryless Gaussian source is proposed. The shape quantizer is constructed from wrapped spherical codes that map a sphere packing in $\mathbb{R}^{k-1}$ onto a sphere in $\mathbb{R}^k$, and the gain codebook is a globally optimal scalar quantizer. A wrapped Leech lattice shape quantizer is used to demonstrate a signal to quantization noise ratio within 1 dB of the distortion-rate function for rates above 1 bit per sample, and an improvement over existing techniques of similar complexity. An asymptotic analysis of the tradeoff between gain quantization and shape quantization is also given.

Index Terms: *vector quantization, lattice coding, Gaussian source, data compression*

*J. Hamkins (`hamkins@jpl.nasa.gov`) is with the Jet Propulsion Laboratory, 4800 Oak Grove Dr., Pasadena, CA 91109-8099. K. Zeger (`zeger@ucsd.edu`) is with the Department of Electrical and Computer Engineering, University of California at San Diego, La Jolla, CA 92093-0407. This work was supported in part by the National Science Foundation. This paper was presented in part at the IEEE International Symposiums on Information Theory, in Ulm, Germany, July 1997, and in Washington, D.C., June 2001.

1 Introduction

An important goal in source coding is to design quantizers that have both reasonable implementation complexity and performance close to the distortion-rate function of a source. Scalar quantizers have low implementation complexity, but their distortion performance is usually much worse than the distortion-rate function. Conversely, fixed blocklength constructive techniques for structured vector quantizers (VQ), such as the generalized Lloyd algorithm (GLA) [1] perform well, but their creation, storage, and encoding complexities each grow exponentially in both dimension and rate (also see [2, 3]).

A number of complexity constrained VQs have been proposed in an attempt to improve upon scalar quantization while retaining low implementation complexity (e.g., see [4]). This paper makes use of two of these methods, lattice quantization and shape-gain quantization, together with wrapped spherical codes for channel coding [5]. Our proposed fixed rate quantizer does not have exponential complexity; in fact, the operating complexity grows linearly with the rate.

The quantizer presented in this paper is designed for a memoryless Gaussian source. One reason for studying the memoryless Gaussian source is that it naturally arises in numerous applications. For example, the prediction error signal in a DPCM (differential pulse code modulation) coder for moving pictures is well-modeled as Gaussian [6]. Also, discrete Fourier transform coefficients and holographic data can often be considered to be the output of a Gaussian source [7] (although some other aspects of images and speech are better modeled as Laplacian distributions [8, 9]). Furthermore, a known filtering technique tends to make memoryless sources appear Gaussian, which makes the system insensitive to errors in modeling the input [10]. The Gaussian source is also easier mathematically to analyze compared to some other sources, because its distortion-rate function is known explicitly. In fact, the Gaussian is known to be the most difficult source to compress, in a rate vs. distortion sense [11]. Finally, the Gaussian source has provided a historical benchmark for measuring how close a practical quantizer can come to the theoretical performances predicted by Shannon [2, 3, 6, 7, 10, 12–28].

Section 2 gives properties of Gaussian source coding and Section 3 describes the construction and performance analysis of the proposed wrapped shape-gain vector quantizer. It is shown how a fixed rate lattice quantizer can be transformed into a shape-gain quantizer. An asymptotic analysis gives the optimal high resolution tradeoff between allocating rate to the gain and shape quantizers, and the indexing problem is discussed. Section 4 describes a specific implementation of the proposed Gaussian coder using the 24-dimensional Leech lattice for the shape codebook. The performance is compared against other known quantizers and the computational complexity and confidence intervals are determined. For a memoryless Gaussian source, this shape-gain quantizer performs better than other quantizers in the literature at rates of three bits per sample or higher. Some extensions are given in Section 5.

2 Preliminaries

Let $X \in \mathbb{R}^k$ be a random vector with independent components drawn from a $N(0, \sigma^2)$ memoryless Gaussian source. The probability density function (pdf) of X is $f_X(Y) = (2\pi\sigma^2)^{-k/2} \exp(-\|Y\|^2/(2\sigma^2))$ for $Y \in \mathbb{R}^k$. Let $\Omega_k = \{Y \in \mathbb{R}^k : \|Y\| = 1\}$ be the unit sphere in k-dimensions, and let $S_k = 2\pi^{k/2}/\Gamma(k/2)$ be the $(k-1)$-dimensional content ("surface area") of Ω_k, where $\Gamma(u) = \int_0^\infty t^{u-1}e^{-t}\,dt$ is the usual gamma function. Also denote the beta function by $\beta(u, v) = \Gamma(u)\Gamma(v)/\Gamma(u+v)$. The following lemma gives some properties of $g = \|X\|$.

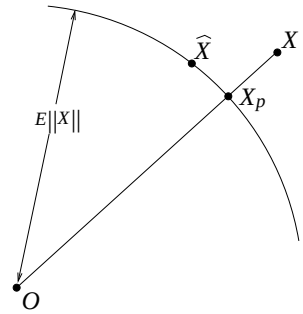


Figure 1: Encoding using Sakrison's quantizer.

Lemma 1.

$$\textit{pdf:} \qquad f_g(r) = f_{\|X\|}(r) = \frac{2r^{k-1}\exp\left(\frac{-r^2}{2\sigma^2}\right)}{\Gamma(k/2)(2\sigma^2)^{k/2}} \tag{1}$$

$$\textit{mean:} \qquad E[\|X\|] = \frac{\sqrt{2\sigma^2}\,\Gamma\left(\frac{k+1}{2}\right)}{\Gamma\left(\frac{k}{2}\right)} = \frac{\sqrt{2\pi\sigma^2}}{\beta\left(\frac{k}{2},\frac{1}{2}\right)} \tag{2}$$

$$\textit{second moment:} \qquad E[\|X\|^2] = k\sigma^2 \tag{3}$$

$$\textit{variance:} \qquad \textit{var}[\|X\|] = k\sigma^2 - \frac{2\pi\sigma^2}{\beta^2\left(\frac{k}{2},\frac{1}{2}\right)}. \tag{4}$$

Eq. (1) is the generalized Rayleigh law [29] and (2), (3), and (4) follow by direct computation.

A consequence of Lemma 1 is that the mean of g is approximately $\sigma\sqrt{k-1/2}$ for large k (by application of Stirling's formula), while the variance of g is bounded by $\sigma^2/2$ for all k [30]. Thus, as $k \to \infty$, the normalized quantity $g/\sqrt{k\sigma^2}$ has a mean which tends to one and variance which tends to zero. This is the so-called "sphere-hardening" effect [31], and implies that for large k, the random vector $X/\sqrt{k\sigma^2}$ is approximately uniformly distributed on Ω_k, which provides motivation for mapping lattices from $\mathbb{R}^{k-1}$ to Ω_k. The performance of lattice quantizers for a uniform source in a region of $\mathbb{R}^{k-1}$ (asymptotically optimal under Gersho's conjecture [32]) can then be transformed to the same performance for a uniform source in Ω_k.

A k-dimensional *vector quantizer* is a mapping $Q : \mathbb{R}^k \to \mathbb{R}^k$ whose range, called a *codebook*, is finite. The elements of a codebook are called *codevectors*. A *spherical vector quantizer (SVQ) with radius* r is a vector quantizer whose codevectors each have Euclidean norm r. A *nearest neighbor* quantizer Q is a quantizer such that for every $x \in \mathbb{R}^k$, no codevector is closer to x than $Q(x)$. The *rate* of the vector quantizer Q is defined as $R = (\log_2 N)/k$ bits, where N is the number of codevectors of Q. For notational convenience, $Q(x)$ is often replaced by $\hat{x}$.

A nearest neighbor spherical vector quantizer satisfies $Q(cX) = Q(X)$ for all $c > 0$. Sakrison [25] showed that if a nearest neighbor spherical vector quantizer with radius $E[\|X\|]$ is used to quantize a Gaussian random vector X, then the resulting MSE distortion per dimension can be decomposed into shape and gain distortions as

$$D = \frac{1}{k}E\left[\|X-\hat{X}\|^2\right] = \underbrace{\frac{1}{k}E\left[\left\|X_p-\hat{X}\right\|^2\right]}_{\text{shape distortion}} + \underbrace{\text{var}[\|X\|]/k}_{\text{gain distortion}} \tag{5}$$

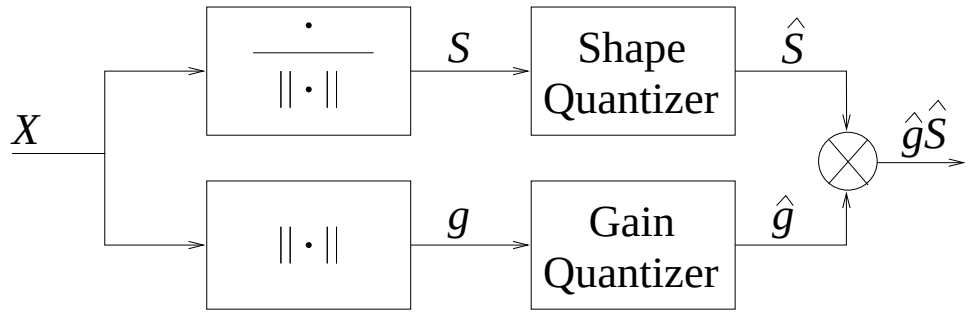


Figure 2: Block diagram of shape-gain quantizer encoder.

where $X_p = E[\|X\|] \cdot \frac{X}{\|X\|}$ (see Figure 1).

The gain distortion term of (5) becomes negligible as k increases and an effective quantizer for X is a spherical vector quantizer with radius $E[\|X\|]$ for a source uniformly distributed on Ω_k. Using a random coding argument Sakrison described such a quantizer and showed that it approaches the distortion-rate function, but the complexity of his quantizer grows linearly with the codebook size.

In the present paper, we describe a high performance Gaussian quantizer using shape-gain vector quantization. The shape quantizer is a wrapped spherical quantizer that can be effectively implemented and which also has excellent distortion performance. No assumption is made that k is asymptotically large, and hence it is not assumed that the gain distortion in (5) is negligible. For example, when $k = 25$ and $\sigma^2 = 1$, the gain distortion dominates the overall distortion performance at rates of three or higher.

A *shape-gain vector quantizer* decomposes a source vector X into a *gain* $g = \|X\|$ and *shape* $S = X/g$, which are quantized to $\hat{g}$ and $\hat{S}$, respectively, and the output is $\hat{X} = \hat{g}\hat{S}$ (see Figures 2 and 3). As is common practice we assume the quantized shape satisfies $\|\hat{S}\| = 1$. An advantage of shape-gain VQ is that the encoding and storage complexities grow with the *sum* of the gain codebook size and shape codebook size, while the effective codebook size is the *product* of these quantities. Necessary optimality conditions are known for optimal shape-gain quantization and these can be used to design locally optimal shape-gain vector quantizers [4, pg. 446]. However, such a design procedure yields unstructured shape codebooks, which can become too large in practice (we determine the optimal codebook sizes analytically for high rates in Section 3.3). In our example implementation, the gain codebook has 15 or fewer codevectors for rates under 4, and the shape codebook can be implicitly computed and thus does not need to be stored.

3 Shape-Gain Wrapped Spherical Vector Quantizer

The proposed shape-gain vector quantizer for Gaussian sources uses a wrapped spherical code for the shape quantizer codebook. We impose the constraint that the quantized gain $\hat{g}$ depends only on the true gain g and the quantized shape vector $\hat{S}$ depends only on the true shape vector S. This allows the gain and shape quantizers to operate in parallel and independently of each other, and it simplifies the analysis of the distortion. A small performance improvement can be realized by allowing $\hat{g}$ and $\hat{S}$ to each depend on both g and S, which is discussed in Section 5. The rates R_s and R_g of the shape and gain codebooks, respectively, are defined as the number of bits used to quantize the shape and gain per scalar component of $X \in \mathbb{R}^k$. Thus the number of bits used to quantize each $(k-1)$-dimensional shape vector is kR_s and the number of bits used to quantize each scalar gain is kR_g. The choice of rates R_s and R_g is discussed in Sections 3.2 and 3.3.

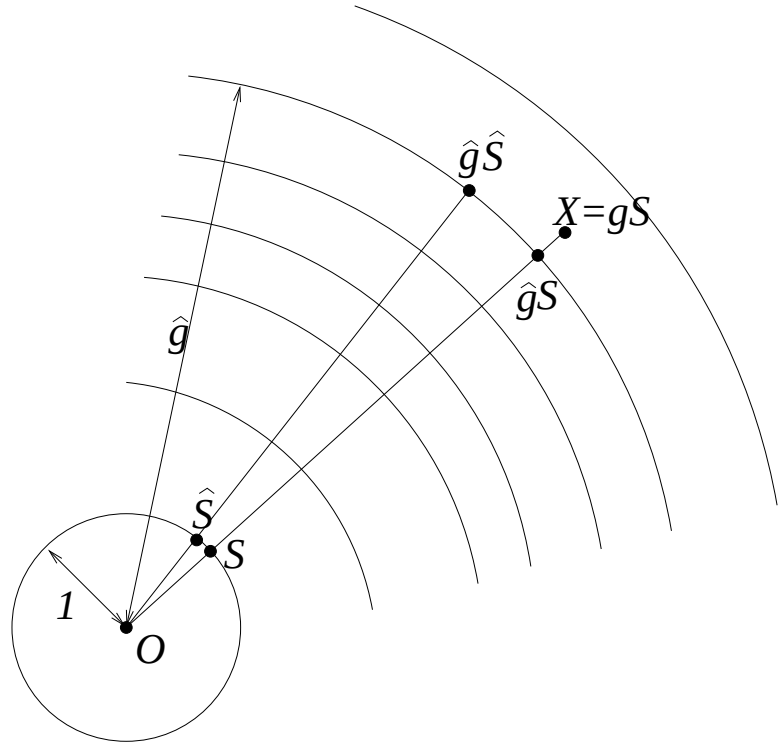


Figure 3: Encoding using the shape-gain quantizer.

We optimize the gain codebook with the Lloyd-Max algorithm [18, 33] using the gain pdf $f_g(r)$ from (1) (no training vectors are needed). Since $f_g(r)$ is strictly log-concave the Lloyd-Max algorithm converges to a *globally* optimum gain codebook [34, 35]. The centroid condition implies that $E[\hat{g}] = E[g]$ and the MSE is $E[g^2] - E[\hat{g}^2]$.

The shape codebook is generated by a wrapped spherical code whose construction is reviewed here (for more details see [5]). Let Λ denote a sphere packing in $\mathbb{R}^{k-1}$ which has minimum distance d_Λ and density Δ_Λ. The *latitude* of a point $X = (x_1, \ldots, x_k) \in \Omega_k$ is defined as $\sin^{-1}(x_k)$, i.e., the angle subtended from the "equator" to X. Let $-\pi/2 = \alpha_0 < \cdots < \alpha_N = \pi/2$ be a sequence of latitudes, where $N = \lceil \pi/\sqrt{d_\Lambda} \rceil$ and $\alpha_i = \pi(\frac{i}{N} - \frac{1}{2})$. The *ith annulus* is defined as the set

$$A_i = \{(x_1, \ldots, x_k) \in \Omega_k \ : \ \alpha_i \leq \sin^{-1} x_k < \alpha_{i+1}\}$$

i.e., the points between consecutive latitudes (see Figure 4). Let $(x)_+ \equiv \max(0, x)$, and for each $X = (x_1, \ldots, x_k) \in A_i$, let

$$X_L = \arg\min_Z \{\|X - Z\| : Z = (z_1, \ldots, z_{k-1}, \sin \alpha_i) \in \Omega_k\}.$$

i.e., the closest point to X that lies on the border between A_{i-1} and A_i (see Figure 5). Let prime notation denote the mapping from $\mathbb{R}^k$ to $\mathbb{R}^{k-1}$ obtained by deletion of the last coordinate, so that for example, $X' = (x_1, \ldots, x_{k-1})$. For each i, define a one-to-one mapping h_i from A_i to a subset of $\mathbb{R}^{k-1}$ by

$$h_i(X) \equiv \frac{X'}{\|X'\|} \cdot (\|(X_L)'\| - \|X_L - X\|)_+ . \tag{6}$$

The *wrapped spherical vector quantizer* codebook W_Λ with respect to a packing Λ is defined as

1. Given k source samples, form the vector $X \in \mathbb{R}^k$.
2. Compute $g = \|X\|$ and $S = X/g$.
3. Use the gain codebook to quantize g as $\hat{g}$.
4. Find i such that $\alpha_i \leq \sin^{-1} x_k < \alpha_{i+1}$, and compute $h_i(S)$.
5. Find the nearest neighbor $\hat{h}_i(S)$ to $h_i(S)$, in $\Lambda \backslash \{0\}$.
6. Compute $h_i^{-1}(\hat{h}_i(S))$ to identify the quantized shape $\hat{S}$.
7. Compute the index of $\hat{g}\hat{S}$ and transmit.

Table 1: Algorithmic description of the quantizer W_Λ .

$$\mathrm{W}_\Lambda \equiv \bigcup_i h_i^{-1}(\Lambda \setminus \{0\}). \tag{7}$$

An example of a wrapped spherical vector quantizer in $\mathbb{R}^3$ is shown in Figure 6, where the codevectors are the centers of the spherical caps. Table 1 describes the procedure for using W_Λ .

3.1 Decomposition into Shape and Gain Distortions

The distortion of the proposed Gaussian quantizer decomposes into gain and shape distortions in much the same way as for Sakrison's spherical vector quantizer in (5). The gain distortion can be evaluated using numerical integration. The shape distortion can be closely approximated and verified to be accurate by simulations.

The MSE per dimension of W_Λ can be decomposed as

$$\begin{aligned} D &= \frac{1}{k} E[\|X - \hat{g}\hat{S}\|^2] \\ &= \frac{1}{k} E[\|X - \hat{g}S\|^2] + \frac{2}{k} E[(X - \hat{g}S)^T(\hat{g}S - \hat{g}\hat{S})] + \frac{1}{k} E[\|\hat{g}S - \hat{g}\hat{S}\|^2] \\ &\equiv D_g + D_c + D_s \end{aligned} \tag{8}$$

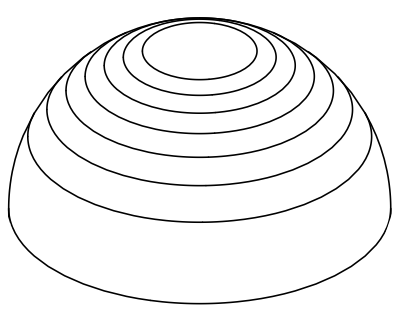
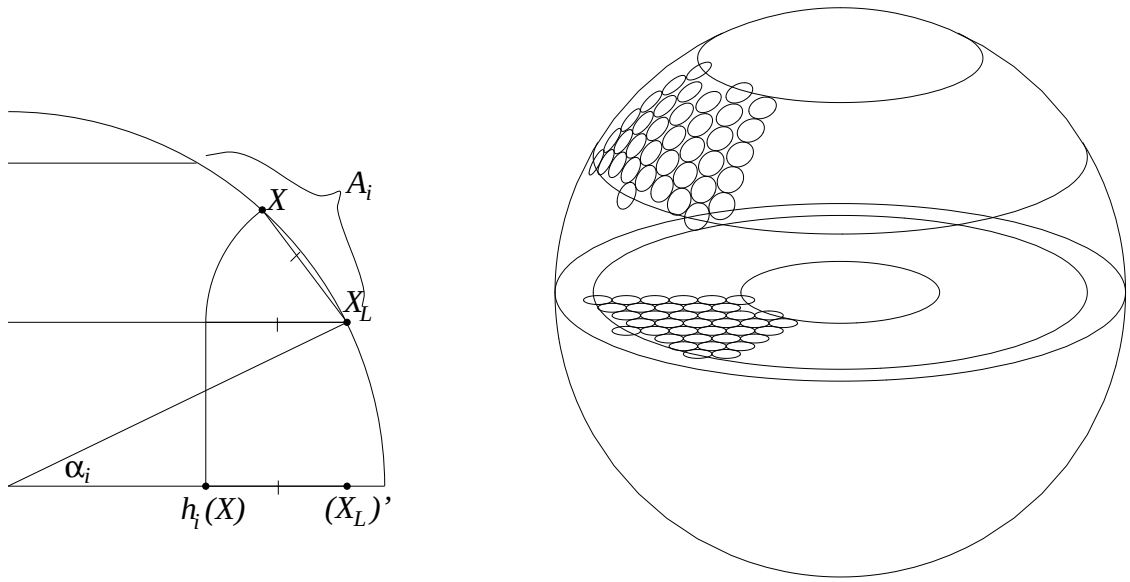


Figure 4: Ω_k is partitioned into annuli.



(a) (b)

Figure 5: (a) The mapping of a point X under h_i. (b) The mapping of many points under h_i. The lattice in the plane has similar structure on Ω_3.

where D_g, D_c, and D_s denote the first, second, and third terms of (8), respectively. Thus,

$$\begin{aligned} D_g &= \frac{1}{k} E\left[\left\| \left(1 - \frac{\hat{g}}{g}\right) X \right\|^2\right] \\ &= \frac{1}{k} E[(g - \hat{g})^2] \end{aligned} \tag{9}$$

which is the per-dimension distortion due to the gain quantizer. If $\hat{g}S$ is known, then $\hat{g}$, S, and $\hat{S}$ are each also known, so that

$$\begin{aligned} D_c &= \frac{2}{k} E[E[(X - \hat{g}S)^T(\hat{g}S - \hat{g}\hat{S})|\hat{g}S]] \\ &= \frac{2}{k} E[E[(X - \hat{g}S)^T|\hat{g}S](\hat{g}S - \hat{g}\hat{S})] \\ &= \frac{2}{k} E[E[(g - \hat{g})|\hat{g}]S^T(\hat{g}S - \hat{g}\hat{S})] \qquad (10) \\ &= 0 \qquad (11) \end{aligned}$$

where (10) follows by the independence of g and $\hat{g}$ from S, and (11) follows from the centroid condition

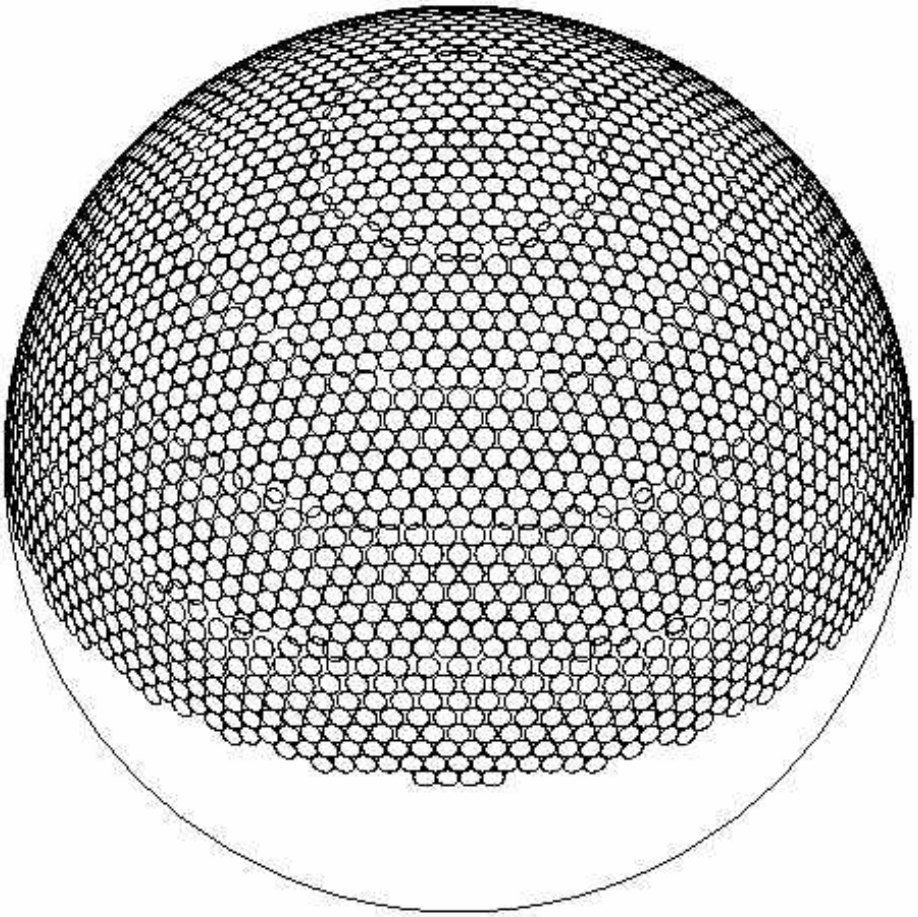


Figure 6: A wrapped spherical vector quantizer.

of the gain quantizer. Finally,

$$\begin{aligned} D_s &= \frac{1}{k}E[\hat{g}^2]E[\|S-\hat{S}\|^2] && (12)\\ &= \frac{1}{k}(E[g^2]-E[(g-\hat{g})^2])E[\|S-\hat{S}\|^2] && (13)\\ &\approx \frac{1}{k}E[g^2]E[\|S-\hat{S}\|^2] && (14)\\ &= \sigma^2 E[\|S-\hat{S}\|^2] && (15) \end{aligned}$$

where (12) follows from the independence of S and $\hat{S}$ from g, (13) follows from the centroid condition of $\hat{g}$, and (15) follows from (3). The approximation in (14) is accurate for high signal-to-noise ratios (SNR) for the gain quantizer, which we will assume. It can be made more exact by estimating the error term via high resolution analysis using Bennett's integral, but we will not need to do so here. Hence, D_s acts as a "shape distortion" (multiplied by the constant $E[g^2]$).

In summary, the distortion of W_Λ is

$$D \approx \frac{1}{k}E[(g-\hat{g})^2]+\sigma^2 E[\|S-\hat{S}\|^2] \tag{16}$$

which partitions the distortion of W_Λ into shape and gain components, as in [25]. The decomposition of D allows us to optimize W_Λ by separately optimizing the shape and gain components.

The gain distortion is given by

$$D_g = \frac{1}{k}E[(g-\hat{g})^2] = \frac{1}{k}\int_0^\infty (r-\hat{g}(r))^2 f_g(r)\,dr \tag{17}$$

where $\hat{g}(r)$ is the gain quantization of r and where $f_g(r)$ is given in (1). This integral can be numerically evaluated once the gain quantizer has been designed.

We estimate the shape distortion D_s, use it in the design algorithm, and validate its accuracy by the observed shape distortion in the simulations for W_Λ . In all cases reported, the approximate computations of distortion agree with the simulated results within 0.1dB.

It follows from [30, Lemma 4.2] that if $U, V \in A_i$ and $\|h_i(U)-h_i(V)\| = O(d_\Lambda)$, then

$$1-O(\sqrt{d_\Lambda}) \le \frac{\|h_i(U)-h_i(V)\|^2}{\|U-V\|^2} \le 1, \tag{18}$$

i.e., the mapping used in W_Λ nearly preserves distances. Thus, for asymptotically high R_s, the distortion, $E[\|S-\hat{S}\|^2]$, of W_Λ , for S uniformly distributed on Ω_k, is equal to the distortion of the underlying lattice quantizer with codebook Λ for a uniform source in $\mathbb{R}^{k-1}$. Let Π be a Voronoi region of a $(k-1)$-dimensional lattice Λ such that $0 \in \Pi$, and let $V(\Lambda)$ denote the volume of Π. The normalized second moment of Π (or of the lattice Λ) is

$$G(\Lambda) = \frac{\frac{1}{k-1}\int_\Pi \|t\|^2\,dt}{V(\Lambda)^{1+\frac{2}{k-1}}}$$

and the $(k-1)$-dimensional vector mean-squared error when Λ is used to quantize a uniform source[1],

[1]Moo and Neuhoff [36] showed that the minimum MSE for quantizing a non-uniform unbounded source using a lattice, decays to zero asymptotically as $2^{-2R+O(\log(R))}$ instead of the known $2^{-2R+O(1)}$ decay rate using asymptotically optimal quantizers.

1. For rate allocation (R_g, R_s), use (a)-(c) below to compute the distortion.
   (a) Estimate the minimum distance of Λ which will produce a rate R_s shape quantizer.
   (b) Use the Lloyd-Max algorithm to optimize the rate R_g gain scalar quantizer for $f_g(r)$.
   (c) Estimate the distortion of W_Λ using (17) and (19).
2. Identify the allocation (R_g, R_s) which minimizes the estimated distortion in step 1, using Brent's method [38].
3. Compute R_s exactly using the theta function of Λ.

Table 2: Optimization algorithm for construction of W_Λ at rate R.

neglecting overload distortion, is the mean-squared error in any Voronoi region, given by

$$\frac{1}{V(\Lambda)} \int_{\Pi} \|t\|^2 \, dt = (k-1) G(\Lambda) V(\Lambda)^{\frac{2}{k-1}}.$$

Thus, for finite R_s the shape distortion is approximated by

$$D_s \approx \sigma^2 E[\|S - \hat{S}\|^2] \approx (k-1)\sigma^2 G(\Lambda) V(\Lambda)^{\frac{2}{k-1}}. \tag{19}$$

For asymptotically large R_s and R_g, the first approximation in (19) becomes tight by (13) because $E[(g - \hat{g})^2] \to 0$, and the second becomes tight because $d_\Lambda \to 0$ in (18). Thus,

$$\lim_{R_s \to \infty} D_s V(\Lambda)^{\frac{-2}{k-1}} = \sigma^2 \lim_{R_s \to \infty} E[\|S - \hat{S}\|^2] V(\Lambda)^{\frac{-2}{k-1}} = (k-1)\sigma^2 G(\Lambda) \tag{20}$$

The values (or close approximations) of $G(\Lambda)$ are given for the best known lattices for the uniform source in [37, pg. 61]. The shape distortion is affected by scaling Λ. For example, doubling the minimum distance of Λ increases $V(\Lambda)$ by a factor of 2^{k-1}, while $G(\Lambda)$ is invariant to scaling, and the shape distortion therefore increases by a factor of four. The total distortion $D = D_g + D_s$ is estimated using (17) and (19).

3.2 Experimental Allocation of Shape and Gain Rates

Let R be the transmission rate of the shape-gain wrapped SVQ and let the shape code rate R_s and gain code rate R_g satisfy $R_s + R_g = R$. The rate R_s determined by (7) can be altered by rescaling Λ so that more or fewer points are contained in W_Λ . We numerically determine the allocation of rate R between R_s and R_g that minimizes the distortion of the wrapped SVQ, using the design algorithm given in Table 2. In the next section, we provide an analytical solution for large rates. Since the gain codebook size is an integer, the values of R_g are restricted to a finite set and the optimal value of R_g can be found exactly. (This is in contrast to optimizations over an infinite set, in which an iterative algorithm may not converge to precisely the optimal value in bounded time.)

For a given pair (R_g, R_s), the gain codebook is optimized using the Lloyd-Max algorithm with R_g bits. Since each Voronoi cell corresponds to one lattice point, the number of shape quantizer codevectors

is closely approximated by the $(k-1)$-dimensional content of the sphere Ω_k divided by the volume of one Voronoi cell (recall, $V(\Lambda) \approx V(h_i(\Lambda))$ [30]). That is, $2^{kR_s} \approx S_k/V(\Lambda)$ and it was shown in [30] that

$$\lim_{R_s \to \infty} V(\Lambda) 2^{kR_s} = S_k. \tag{21}$$

Thus, for a given shape rate R_s, we scale Λ before the shape codebook is constructed such that the volume of the Voronoi cell satisfies $V(\Lambda) = S_k 2^{-kR_s}$.

After optimization is complete, the actual number of codevectors is computed by evaluating the theta function. This more time-consuming step is avoided during the optimization step, which only uses estimates of the codebook sizes.

3.3 Theoretical Allocation of Shape and Gain Rates

Here we consider the theoretical tradeoff between allocating transmission rate to the gain quantizer and the shape quantizer. In order to facilitate analysis we use high resolution assumptions. For general shape-gain quantizers this is an unsolved problem. However, if: (i) the source is Gaussian; (ii) the shape codebook is based on a lattice; and (iii) the gain quantizer is independent of the shape quantizer, then it is possible to obtain a high resolution analytic solution. This may help to provide intuition about the more general case too.

Since the transmission rate R, the shape quantizer rate R_s, and the gain quantizer rate R_g are related by $R = R_s + R_g$ we can write the shape distortion and the high resolution gain distortion as

$$D_s \approx (k-1)\sigma^2 G(\Lambda) V(\Lambda)^{\frac{2}{k-1}} \approx C_s 2^{-2R_s\left(\frac{k}{k-1}\right)} \tag{22}$$

$$D_g \sim C_g 2^{-2R_g k} = C_g 2^{-2k(R-R_s)} \tag{23}$$

where (22) follows using (19) and where (23) holds for large R_g from Bennett's integral [4], with C_s and C_g constants that are independent of R_s and R_g. To determine the growth rate of R_s as a function of R that minimizes $D = D_s + D_g$ one can intuitively reason that the asymptotic expressions for D_s and D_g must decay at the same rate. Equating the exponents $2R_s\left(\frac{k}{k-1}\right) = 2k(R-R_s)$ gives $R_s = \left(\frac{k-1}{k}\right)R$, which gives an accurate first-order approximation of R_s. Indeed, this follows the intuition that the shape codebook based on a $(k-1)$-dimensional lattice and the gain codebook based on a scalar quantity should have a rate allocation of approximately $R(k-1)/k$ and R/k, respectively, for the rate R, k-dimensional vector quantizer. The exact optimal choice of R_s is given in the following theorem, where it is shown that $D \sim A2^{-2R}$, with the constant A identified and depending only on the vector dimension k, the source variance σ^2, and the the normalized second moment $G(\Lambda)$ of the lattice Λ.

Theorem 1. *Let $k > 1$, let $X \in \mathbb{R}^k$ be an uncorrelated Gaussian vector with zero mean and component variances $\sigma^2 < \infty$, and let Λ be a lattice in $\mathbb{R}^{k-1}$ with normalized second moment $G(\Lambda)$. Suppose X is quantized by a k-dimensional shape-gain vector quantizer at rate $R = R_s + R_g$ (where R_s and R_g are the shape and gain quantizer rates) with independent shape and gain encoders and whose shape codebook is a wrapped spherical code constructed from Λ. Then the asymptotic decay of the minimum mean squared quantization error D is given by*

$$\lim_{R \to \infty} D 2^{2R} = \frac{k}{(k-1)^{1-\frac{1}{k}}} \cdot C_g^{\frac{1}{k}} C_s^{1-\frac{1}{k}} \tag{24}$$

and is achieved by $R_s = R_s^*$ and $R_g = R_g^*$, where

$$R_s^* = \left(\frac{k-1}{k}\right)\left[R + \frac{1}{2k}\log_2\left(\frac{C_s}{C_g}\cdot\frac{1}{k-1}\right)\right], \tag{25}$$

$$R_g^* = \left(\frac{1}{k}\right)\left[R - \frac{k-1}{2k}\log_2\left(\frac{C_s}{C_g}\cdot\frac{1}{k-1}\right)\right], \tag{26}$$

$C_s = \sigma^2 \cdot (k-1)G(\Lambda)\left(\frac{2\pi^{k/2}}{\Gamma(k/2)}\right)^{\frac{2}{k-1}}$*, and* $C_g = \sigma^2 \cdot \frac{3^{k/2}\Gamma^3\left(\frac{k+2}{6}\right)}{8k\Gamma(k/2)}$.

Proof. Let D_s and D_g be the distortions of the shape and gain quantizers at rates R_s and R_g, respectively. From (11) we have

$$D = \inf_{\substack{R_s, R_g \\ R_s + R_g = R}} (D_s + D_g), \tag{27}$$

and therefore

$$D2^{2R} = s_R 2^{-2a_R\left(\frac{k}{k-1}\right)} + g_R 2^{2ka_R} \tag{28}$$

where

$$\begin{aligned} a_R &= R_s^* - \left(\frac{k-1}{k}\right)R \\ s_R &= D_s 2^{2R_s^*\left(\frac{k}{k-1}\right)} \\ g_R &= D_g 2^{2k(R-R_s^*)}. \end{aligned}$$

Also, $R_s^* \to \infty$ and $R - R_s^* = R_g^* \to \infty$ as $R \to \infty$, for otherwise either D_s or D_g (and hence D) would be bounded away from zero (i.e. not achieving the minimum MSE quantization). Define the quantity

$$\begin{aligned} C_g &= \lim_{R\to\infty} g_R \\ &= \frac{\|f_g\|_{1/3}}{12k} && (29) \\ &= \frac{1}{12k}\left(\int_0^\infty |f_g(r)|^{1/3} dr\right)^3 \\ &= \frac{1}{\Gamma(k/2)(2\sigma^2)^{k/2}6k}\left(\int_0^\infty r^{(k-1)/3} e^{-r^2/(6\sigma^2)} dr\right)^3 && (30) \\ &= \frac{1}{\Gamma(k/2)(2\sigma^2)^{k/2}6k}\left(2^{(k-4)/6}(3\sigma^2)^{(k+2)/6}\int_0^\infty t^{(k-4)/6} e^{-t} dt\right)^3 && (31) \\ &= \sigma^2 \cdot \frac{3^{\frac{k}{2}}\Gamma^3\left(\frac{k+2}{6}\right)}{8k\Gamma(k/2)} && (32) \end{aligned}$$

where (29) follows from Bennett's integral [4]; (30) follows using the density function f_g of the gain $g = \|X\|$ from (1); (31) follows by substituting $r^2 = 6\sigma^2 t$; and (32) follows from [39, pg. 342, eq. 662].

Define the quantity

$$\begin{aligned}
C_s &= \lim_{R\to\infty} s_R && (33)\\
&= \lim_{R\to\infty} \frac{1}{k}(E[g^2] - E[(g-\hat{g})^2])E[\|S-\hat{S}\|^2]\cdot 2^{2R_s^*\left(\frac{k}{k-1}\right)} && (34)\\
&= \sigma^2 \lim_{R\to\infty} E[\|S-\hat{S}\|^2]\cdot 2^{2R_s^*\left(\frac{k}{k-1}\right)} && (35)\\
&= \sigma^2(k-1)G(\Lambda) \lim_{R\to\infty} V(\Lambda)^{\frac{2}{k-1}}\cdot 2^{2R_s^*\left(\frac{k}{k-1}\right)} && (36)\\
&= \sigma^2(k-1)G(\Lambda)S_k^{\frac{2}{k-1}} && (37)\\
&= \sigma^2(k-1)G(\Lambda)\left(2\pi^{k/2}/\Gamma(k/2)\right)^{\frac{2}{k-1}} && (38)
\end{aligned}$$

where (34) follows from (13); (35) follows from (3) and $\lim_{R\to\infty} E[(g-\hat{g})^2] = 0$; (36) follows from (20); (37) follows from (21); and (38) follows from $S_k = 2\pi^{k/2}/\Gamma(k/2)$. Note that the limit in (33) exists by working backwards from (37).

Let $a = \left(\frac{k-1}{2k^2}\right)\log_2\left(\frac{C_s}{C_g}\cdot\frac{1}{k-1}\right)$ and notice that the unique minimum value of the function $f(x) = C_s 2^{-2x\left(\frac{k}{k-1}\right)} + C_g 2^{2kx}$ is achieved at $x = a$, since f is strictly convex and $f'(a) = 0$.

Suppose D_s and D_g are the distortions corresponding to $R_s = \left(\frac{k-1}{k}\right)R$. Then

$$\begin{aligned}
D2^{2R} \le (D_s + D_g)2^{2R} &= D_s 2^{2\left(\frac{k}{k-1}\right)R_s} + D_g 2^{2k(R-R_s)}\\
&\longrightarrow C_s + C_g
\end{aligned}$$

as $R\to\infty$. Thus $D2^{2R}$ is bounded as $R\to\infty$. In addition, s_R and g_R are bounded away from zero for sufficiently large R. Thus, a_R is bounded, from (28), and hence for $k \ge 2$,

$$\begin{aligned}
\left|D2^{2R} - C_s 2^{-2a_R\left(\frac{k}{k-1}\right)} - C_g 2^{2ka_R}\right| &\le |s_R - C_s|\cdot 2^{-2a_R\left(\frac{k}{k-1}\right)} + |g_R - C_g|\cdot 2^{2ka_R}\\
&\longrightarrow 0
\end{aligned}$$

as $R\to\infty$. Since $C_s 2^{-2a\left(\frac{k}{k-1}\right)} + C_g 2^{2ka} \le C_s 2^{-2a_R\left(\frac{k}{k-1}\right)} + C_g 2^{2ka_R}$ for all R, we have

$$D2^{2R} \ge D2^{2R} - \left(C_s 2^{-2a_R\left(\frac{k}{k-1}\right)} + C_g 2^{2ka_R}\right) + \left(C_s 2^{-2a\left(\frac{k}{k-1}\right)} + C_g 2^{2ka}\right).$$

So for any $\epsilon > 0$, we have $D2^{2R} \ge C_s 2^{-2a\left(\frac{k}{k-1}\right)} + C_g 2^{2ka} - \epsilon$ for sufficently large R. Thus

$$\liminf_{R\to\infty} D2^{2R} \ge C_s 2^{-2a\left(\frac{k}{k-1}\right)} + C_g 2^{2ka}. \quad (39)$$

On the other hand, suppose D_s and D_g are the distortions corresponding to $R_s = \left(\frac{k-1}{k}\right)R + a$. Then from (27),

$$\begin{aligned}
D2^{2R} &\le 2^{2R}(D_s + D_g) && (40)\\
&= D_s 2^{2R_s\left(\frac{k}{k-1}\right)} 2^{-2a\left(\frac{k}{k-1}\right)} + D_g 2^{2k(R-R_s)} 2^{2ak} && (41)\\
&\longrightarrow C_s 2^{-2a\left(\frac{k}{k-1}\right)} + C_g 2^{2ka} && (42)
\end{aligned}$$

as $R \to \infty$. Thus,

$$\limsup_{R\to\infty} D2^{2R} \quad \leq \quad C_s 2^{-2a\left(\frac{k}{k-1}\right)} + C_g 2^{2ka}. \tag{43}$$

Combining (39) and (43) gives

$$\lim_{R\to\infty} D2^{2R} = C_s 2^{-2a\left(\frac{k}{k-1}\right)} + C_g 2^{2ka}$$

which is achieved by $R_s^* = \left(\frac{k-1}{k}\right) R + a$. Substituting the definition of a into $C_s 2^{-2a\left(\frac{k}{k-1}\right)} + C_g 2^{2ka}$, $R_s^* = \left(\frac{k-1}{k}\right) R + a$, and $R_g^* = R - R_s^*$, gives (24), (25), and (26), respectively. □

Note that for large R, the optimal allocation of transmission rate between the shape quantizer and the gain quantizer is from (26) approximately $R_s^* \approx \left(1 - \frac{1}{k}\right) R$ and $R_g^* \approx \frac{1}{k} R$. This means that the shape codebook should have about $2^{(k-1)R}$ codevectors and the gain codebook should have about 2^R scalar codepoints, as intuition would indicate. This corresponds roughly to what was observed in the experimental rate allocation optimization. In simulations, we observed that the optimal gain codebook rate was within 8% of this figure when $R \geq 3$ and within 1% when $R \geq 5$.

3.4 Index Assignment

In order to implement the shape-gain spherical quantizer, the $M = 2^{kR_s}$ quantizer codevectors must be uniquely identified by binary strings of length kR which are transmitted across the channel. The assignment is accomplished in a similar manner as in [14] for the pyramid vector quantizer for the Laplacian source. First, the number of codevectors in each annulus of the shape codebook is counted using the theta function. We report on specific results using the Leech lattice Λ_{24}, for which the $\mathrm{W}_{\Lambda_{24}}$ codes need a one-time computation of the first few hundred coefficients of the theta function of the Leech lattice, which are stored and used as needed.

It is assumed that there is an efficient method for assigning indices to the underlying lattice. This is the case with many lattices, including the Leech lattice Λ_{24} (e.g., see [40]).

The codevectors of the wrapped spherical code are assigned to integers according to their quantized gain, annulus, and order within their annulus, as follows. Let N represent the number of annuli of the shape codebook. Let P_j be the number of points in the jth annulus of a shell, and let P be the total number of points in the shape codebook. Assuming all indices start at 0, the lth point within the jth annulus of the ith gain shell is assigned to the number

$$iP + \sum_{a=0}^{j-1} P_a + l. \tag{44}$$

Both the encoder and decoder must compute this summation. This can be made efficient by storing in memory the partial summations $\sum_{a=0}^{j-1} P_a$, for $j = 0, 1, \ldots, N-1$. The memory required for this is equal to the total number of annuli in the codebook, which is generally not large. For example, in the codebook $\mathrm{W}_{\Lambda_{24}}$of rate 4, there are 36 total annuli.

4 Simulations and Comparisons

4.1 Confidence Intervals of the Simulations

The codebook $\mathrm{W}_{\Lambda_{24}}$ was optimized according to Table 2 and its performance was evaluated with 500,000 i.i.d. Gaussian random samples blocked into 20,000 25-dimensional vectors and encoded as in Table 1. The lattice encoding used the Leech lattice nearest neighbor algorithm in [41]. The quality of the simulation results is expressed in terms of a 95% confidence interval. The simulation run of 20,000 vectors was broken down into 20 blocks of 1,000 vectors. For each block, the average distortion was determined. Applying the central limit theorem to each block distortion and using the students t-distribution, we calculated the 95% confidence interval. For each simulation, these intervals were found to be less than 0.03 dB.

4.2 Performance Comparisons

Table 3 demonstrates that shape-gain VQ using $\mathrm{W}_{\Lambda_{24}}$ performs within 1 dB of the distortion-rate function for rates in the range of 2-7 bits/sample. For this range, shape-gain VQ using $\mathrm{W}_{\Lambda_{24}}$ outperforms many of the best quantizers in the literature, including 256-state trellis coded quantization (TCQ) [20], two-dimensional four-state trellis coded vector quantization (TCVQ) [26], Fischer's spherical vector quantization [14], and Lloyd-Max scalar quantization. With a large number of trellis states, TCQ and TCVQ may perhaps outperform shape-gain VQ using $\mathrm{W}_{\Lambda_{24}}$; however, the reports of results in the literature have thus far been limited to trellises with 256 or fewer states because the design complexity of TCQ and TCVQ is somewhat prohibitive for larger trellises. Trellis-based scalar-vector quantization (TB-SVQ) [42] performs slightly better than shape-gain VQ using $\mathrm{W}_{\Lambda_{24}}$ at a rate of 2, but not at a rate of 3.

4.3 Computational Complexity

The arithmetic functions needed to implement the quantizer are addition, multiplication, division, trigonometric functions, square root, and comparison. To make a rough estimate of computational complexity (which of course is machine dependent) we count one operation for any arithmetic function.

In Table 1, Step 1 requires no computation, and Step 2 requires k multiplies, $k-1$ additions, and one square root to calculate the gain; and k divisions to calculate the shape. Step 3 requires one scalar quantization operation, which can be performed by a binary search with at most kR_g comparisons. Step 4 requires $\log_2 N \leq kR_s$ comparisons to identify i; one additional trigonometric function, one difference, one division, and one multiplication to compute $\|S_L - S\| = 2\sin((\sin^{-1} x_k) - \alpha_i)/2)$; one trigonometric function to compute $\|(S_L)'\| = \cos(\alpha_i)$; one difference to compute $\|(S_L)'\| - \|S_L - S\|$; one multiplication, one difference, and one square root to compute $\|S'\| = \sqrt{1 - x_k^2}$; and 1 division and $k-1$ multiplications to compute $h_i(S)$. Thus, Step 4 requires no more than $k + kR_s + 10$ operations. Step 5 requires the number of steps in a nearest neighbor algorithm for Λ. For the Leech lattice, the fastest known algorithm requires about 2955 operations on average [45]. Step 6 requires $k-1$ squarings, $k-2$ additions, and one square root to determine $\|\hat{h}_i(S)\|$; one difference and one division to determine $1/(\|(S_L)'\| - \|\hat{h}_i(S)\|)$; $k-1$ multiplications to determine the first $k-1$ coordinates of $h_i^{-1}(\hat{h}_i(S))$; and one square, one difference, and one square root to determine the last coordinate. Thus, Step 6 requires $3k+2$ operations. Step 7 requires one multiplication and two additions to determine the index. Altogether, this amounts to at most $k(R+7) + L + 15$ arithmetic operations, where k is the dimension, R is the rate, and L is the computational complexity of the nearest neighbor algorithm of Λ. Thus, per sample, the computational complexity is at

Method Rate:	1	2	3	4	5	6	7
Distortion-Rate function	6.02	12.04	18.06	24.08	30.10	36.12	42.14
Shape-gain VQ using $\mathrm{W}_{\Lambda_{24}}$	2.44	11.02	17.36	23.33	29.29	35.27	41.33
TB–SVQ (4 state) [42]	5.39	11.18	16.92				
TB–SVQ (32 state) [42]	5.49	11.28	17.05				
Wilson (128 state) [20, 43]	5.47	10.87	16.78				
TCQ (256 state) [20]	5.56	11.04	16.64				
TCVQ (2D, 16 state) [26]	5.29	10.84	16.62	22.63			
Entropy coded scalar quantizer [20, 22]	4.64	10.55	16.56	22.55	28.57	34.59	40.61
SVQ (estimated) [14]	4.49	10.51	16.53	22.55	28.57	34.59	40.61
GLA (kR=8) (simulation)		10.65		20.98			
Z^{16} lattice [16]		10.07	15.52	21.00	26.16	32.07	37.68
Unrestricted polar Quantizer [44]	4.40	9.63					
Lloyd-Max Scalar [18, 22]	4.40	9.30	14.62	20.22	26.02	31.89	37.81
Uniform scalar [17]	4.40	9.25	14.27	19.38	24.57	29.83	35.13

Table 3: Comparison of various quantization schemes for a memoryless Gaussian source. Values are listed as SNR in decibels. Blank entries indicate that referenced work does not contain a result. Shape-gain VQ using $\mathrm{W}_{\Lambda_{24}}$ is the proposed scheme using the Leech lattice as a shape codebook.

Method	Computations
Shape-gain VQ using $\mathrm{W}_{\Lambda_{24}}$	$R + 126$
TCQ (doubled alphabet) [20]	$3S + 4R + 4$
TCQ (quadrupled alphabet) [20]	$3S + 8R + 8$
Generalized Lloyd algorithm (GLA) [4]	2^{kR+1}
Wilson's stochastic trellis [44]	$S \cdot 2^{R+1}$
Pearlman's stochastic trellis [3]	$(S+2)2^R$

Table 4: Comparison of computational complexities of quantization schemes for a memoryless Gaussian source. Data for other methods are taken from [20, Table XII]. k = dimension, R = rate, S = number of trellis states.

most $R + 7 + (L + 15)/k$. For the shape-gain VQ using $\mathrm{W}_{\Lambda_{24}}$, the parameters are $k = 25$ and $L = 2955$, and the computational complexity is upper bounded by $R + 119$.

Thus, the computational complexity of W_Λ grows linearly with rate, and is comparable to that of trellis-coded quantization (TCQ). Table 4.3 summarizes these complexities.

5 Generalizations of the Shape-Gain Coder

5.1 Non-Gaussian Sources

Inherent in the treatment thus far is that the source has a Gaussian distribution, for if the source is not Gaussian then the high probability region may not be a sphere, but some other shape [46], and the wrapped SVQ cannot be effectively used. This section presents a method to obtain the performance above for any memoryless source. The method consists of transform coding the source. Typically, transform coding is done to remove dependencies between consecutive samples of the source; here, it is used to change the distribution of the source, which may or may not already be i.i.d., to be roughly Gaussian and i.i.d., so that wrapped SVQ may still be used. This same intuition was used in [10] to quantize an arbitrary source and obtain distortion performance that approximates that of a scalar quantizer for a Gaussian source. Unlike the approach in [10], in this section the source is transformed in blocks, instead of using FIR filters.

Let $Q(\cdot)$ be the output of any k-dimensional vector quantizer. Let

$$X \equiv \begin{pmatrix} X_1 & X_{m+1} & & X_{(k-1)m+1} \\ \vdots & \vdots & \cdots & \vdots \\ X_m & X_{2m} & & X_{km} \end{pmatrix}$$

where $X_i \in \mathbb{R}$ has an arbitrary distribution. Let $\mathcal{H}_m$ be a Hadamard matrix of order m, i.e., an $m \times m$ matrix with $+1$ and -1 entries only, such that $\mathcal{H}_m^T \mathcal{H}_m = mI$. Such matrices are known to exist when the

order is any power of 2, and for many other orders as well.[2] Let $H_m \equiv (1/\sqrt{m})\mathcal{H}_m$. Given X, the vector quantizer output is: $H_m^T Q(H_m X)$. If $Y \equiv H_m X$, $\hat{Y} \equiv Q(Y)$, and $e \equiv \hat{Y} - Y$, then it follows that

$$\begin{aligned}
\hat{X} &= H_m^T \hat{Y} \\
&= H_m^T (Y + e) \\
&= H_m^T (H_m X + e) \\
&= H_m^T H_m X + H_m^T e \\
&= X + H_m^T e.
\end{aligned}$$

The end-to-end distortion of this system is

$$\begin{aligned}
E[\|\hat{X} - X\|^2] &= E[(\hat{X} - X)^T (\hat{X} - X)] \\
&= E[(H_m^T e)^T (H_m^T e)] \\
&= E[e^T H_m H_m^T e] \\
&= E[e^T e] \\
&= E[(\hat{Y} - Y)^T (\hat{Y} - Y)] \\
&= E[\|\hat{Y} - Y\|^2]
\end{aligned}$$

Thus, the end-to-end distortion of the system is equal to the distortion due to the quantization of the intermediary signal Y alone. Most importantly, the Hadamard transform modifies the distribution of the input to the vector quantizer. A row Y_i of Y is a k-vector, each component of which is the sum of m different samples (or their negation) from $\{X_i\}$; hence, as $m \to \infty$ the probability distribution of each component of Y_i approaches the Gaussian distribution, by the central limit theorem. Thus, the internal k-dimensional quantizer Q may be optimized with respect to the Gaussian distribution, even if k is fixed and small.

5.2 Other generalizations

There are several other improvements for this shape-gain quantizer. For example, instead of using a scalar quantizer for the gain, gains could be blocked together and vector quantized, or, if fixed rate quantization is not required, entropy coded. Or, we may remove the assumption that the gain and shape codebooks operate independently. With a gain-dependent shape codebook, there could be a different shape codebook associated with each quantized gain value. With a shape-dependent gain codebook, we could choose $\hat{g}$ to minimize $\|X - \hat{g}\hat{S}\|$ instead of $\|g - \hat{g}\|$.

6 Conclusions

The wrapped spherical vector quantizer for the memoryless Gaussian source achieves distortions that are in many cases lower than other published results. The operating complexity of the quantizer grows

[2] Paley's Theorem (1933) [47] guarantees that Hadamard matrices exist for orders equal to $n = 2^e(p^m + 1)$, for all positive integers e and m, and every odd prime p (also for $p = 0$ when $e \geq 2$). The orders for which Hadamard matrices exist include every multiple of 4 up to 268, and all powers of 2. It is an open question as to whether they exist for orders equal to all multiples of 4.

linearly with the rate, and for moderate rates is dominated by the complexity of the nearest neighbor algorithm of the underlying lattice. This complexity is comparable or slightly less than other efficient quantization techniques such as pyramid vector quantization of the Laplacian source [14], trellis coded quantization [20], and trellis coded vector quantization [26]. We note that sphere packings other than lattices may be used to create the shape codebook. In this case, more than one type of Voronoi cell results, and an average over all the different Voronoi cells is necessary to compute the MSE of the scaled packing.

Acknowledgement: The authors thank the two reviewers for very thorough readings of this correspondence and their helpful comments.

References

[1] Y. L. Linde, A. Buzo, and R. M. Gray, "An algorithm for vector quantizer design," *IEEE Trans. Commun.*, vol. COM-28, pp. 84–95, Jan. 1980.

[2] S. G. Wilson and D. W. Lytle, "Trellis encoding of continuous-amplitude memoryless sources," *IEEE Trans. Inform. Theory*, vol. IT-23, pp. 404–409, May 1977.

[3] W. A. Pearlman, "Sliding-block and random source coding with constrained size reproduction alphabets," *IEEE Trans. Commun.*, vol. COM-30, pp. 1859–1867, Aug. 1982.

[4] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*. Boston, MA: Kluwer Academic Publishers, 1993.

[5] J. Hamkins and K. Zeger, "Asymptotically efficient spherical codes—Part I: Wrapped spherical codes," *IEEE Trans. Inform. Theory*, vol. 43, pp. 1774–1785, Nov. 1997.

[6] P. Vogel, "Analytical coding of Gaussian sources," *IEEE Trans. Inform. Theory*, vol. 40, pp. 1639–1645, Sept. 1994.

[7] P. F. Swaszek and J. B. Thomas, "Multidimensional spherical coordinates quantization," *IEEE Trans. Inform. Theory*, vol. IT-29, pp. 570–576, July 1983.

[8] H.-C. Tseng and T. R. Fischer, "Transform and hybrid transform/DPCM coding of images using pyramid vector quantization," *IEEE Trans. Commun.*, vol. COM-35, pp. 79–86, 1987.

[9] T. R. Fischer and K. Malone, "Transform coding of speech with pyramid vector quantization," in *Conf. Rec. MILCOM*, pp. 620–623, 1985.

[10] K. Popat and K. Zeger, "Robust quantization of memoryless sources using dispersive FIR filters," *IEEE Trans. Commun.*, vol. 40, pp. 1670–1674, Nov. 1992.

[11] R. Gallager, *Information Theory and Reliable Communication*. New York, NY: Wiley, 1968.

[12] J.-P. Adoul, C. Lamblin, and A. Leguyader, "Base-band speech coding at 2400 bps using spherical vector quantization," in *IEEE Int. Conf. Acoustics, Speech, Signal Processing*, pp. 1.12.1–1.12.4, 1984.

[13] M. V. Eyuboğlu and G. D. Forney, "Lattice and trellis quantization with lattice- and trellis-bounded codebooks—high-rate theory for memoryless sources," *IEEE Trans. Inform. Theory*, vol. 39, pp. 46–59, Jan. 1993.

[14] T. R. Fischer, "A pyramid vector quantizer," *IEEE Trans. Inform. Theory*, vol. IT-32, pp. 568–583, July 1986.

[15] T. R. Fischer, "Geometric source coding and vector quantization," *IEEE Trans. Inform. Theory*, vol. 35, pp. 137–145, Jan. 1989.

[16] D. G. Jeong and J. D. Gibson, "Uniform and piecewise uniform lattice vector quantization for memoryless Gaussian and Laplacian sources," *IEEE Trans. Inform. Theory*, vol. 39, pp. 786–803, May 1993.

[17] N. S. Jayant and P. Noll, *Digital Coding of Waveforms*. Englewood Cliffs, NJ: Prentice-Hall, 1984.

[18] J. Max, "Quantizing for minimum distortion," *IRE Trans. Inform. Theory*, vol. 6, pp. 7–12, March 1960.

[19] M. W. Marcellin, "On entropy-constrained trellis coded quantization," *IEEE Trans. Commun.*, vol. 42, pp. 14–16, Jan. 1994.

[20] M. W. Marcellin and T. Fischer, "Trellis coded quantization of memoryless and Gauss-Markov sources," *IEEE Trans. Commun.*, vol. 38, pp. 82–93, Jan. 1990.

[21] D. Miller, K. Rose, and P. A. Chou, "Deterministic annealing for trellis quantizer and HMM design using Baum-Welch re-estimation," in *IEEE Int. Conf. Acoustics, Speech, Signal Processing*, pp. V–261–V–264, Apr. 1994.

[22] P. Noll and R. Zelinski, "Bounds on quantizer performance in the low bit-rate region," *IEEE Trans. Commun.*, vol. COM-26, pp. 300–304, Feb. 1978.

[23] J. Pan and T. R. Fischer, "Two-stage vector quantization—lattice vector quantization," *IEEE Trans. Inform. Theory*, vol. 41, pp. 155–163, Jan. 1995.

[24] M. C. Rost and K. Sayood, "The root lattices as low bit rate vector quantizers," *IEEE Trans. Inform. Theory*, vol. 34, pp. 1053–1058, Sept. 1988.

[25] D. J. Sakrison, "A geometric treatment of the source encoding of a Gaussian random variable," *IEEE Trans. Inform. Theory*, vol. IT-14, pp. 481–486, May 1968.

[26] H. S. Wang and N. Moayeri, "Trellis coded vector quantization," *IEEE Trans. Commun.*, vol. 40, pp. 1273–1276, Aug. 1992.

[27] S. D. Servetto, "Lattice quantization with side information," in *Proceedings of the Data Compression Conference*, (Snowbird, Utah), pp. 510–519, Mar. 2000.

[28] F. Chen, Z. Gao, and J. Villasenor, "Lattice vector quantization of generalized gaussian sources," *IEEE Trans. Inform. Theory*, vol. 43, pp. 92–103, Jan. 1997.

[29] K. Miller, *Multidimensional Gaussian Distributions*. New York, NY: John Wiley, 1964.

[30] J. Hamkins, *Design and Analysis of Spherical Codes*. PhD thesis, University of Illinois at Urbana-Champaign, Sept. 1996.

[31] J. D. Gibson and K. Sayood, "Lattice quantization," in *Adv. Electronics Electron Phys.* (P. Hawkes, ed.), vol. 72, pp. 259–330, New York: Academic, 1988.

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

[33] S. P. Lloyd, "Least squares quantization in PCM," *IEEE Trans. Inform. Theory*, vol. 28, pp. 129–136, Mar. 1982.

[34] P. E. Fleischer, "Sufficient conditions for achieving minimum distortion in a quantizer," *IEEE Int. Conv. Rec., Part 1*, pp. 104–111, 1964.

[35] A. V. Trushkin, "Sufficient conditions for uniqueness of a locally optimal quantizer for a class of convex error weighting functions," *IEEE Trans. Inform. Theory*, vol. IT-28, pp. 187–198, Mar. 1982.

[36] P. Moo and D. Neuhoff, "An asymptotic analysis of fixed-rate lattice vector quantization," in *Proc. Int. Symp. Inform. Theory and its Applications*, pp. 409–412, Sept. 1996.

[37] J. H. Conway and N. J. A. Sloane, *Sphere Packings, Lattices, and Groups*. Springer-Verlag, 1988.

[38] W. H. Press, W. T. Vetterling, S. A. Teukolsky, and B. P. Flannery, *Numerical Recipes in C*. New York: Cambridge University Press, 1992.

[39] W. H. Beyer, *CRC Standard Mathematical Tables*. Boca Raton, FL: CRC Press, Inc., 26th ed., 1981.

[40] J.-P. Adoul and M. Barth, "Nearest neighbor algorithm for spherical codes from the Leech lattice," *IEEE Trans. Inform. Theory*, vol. 34, pp. 1188–1202, Sept. 1988.

[41] Y. Be'ery, B. Shahar, and J. Snyders, "Fast decoding of the Leech lattice," *IEEE J. Select. Area Commun.*, vol. 7, pp. 959–966, Aug. 1989.

[42] R. Laroia and N. Farvardin, "Trellis-based scalar-vector quantizer for memoryless sources," *IEEE Trans. Inform. Theory*, vol. 40, pp. 860–870, May 1994.

[43] S. G. Wilson and D. W. Lytle, "Trellis encoding of continuous-amplitude memoryless sources," *IEEE Trans. Inform. Theory*, vol. IT-28, pp. 211–226, May 1982.

[44] S. G. Wilson, "Magnitude/phase quantization of independent gaussian variates," *IEEE Trans. Commun.*, vol. COM-28, pp. 1924–1929, Nov. 1980.

[45] A. Vardy and Y. Be'ery, "Maximum likelihood decoding of the Leech lattice," *IEEE Trans. Inform. Theory*, vol. 39, pp. 1435–1444, July 1993.

[46] K. T. Malone and T. R. Fischer, "Contour-gain vector quantization," *IEEE Trans. Acoust. Speech Sig. Process.*, vol. 36, pp. 862–870, June 1988.

[47] R. E. A. C. Paley, "On orthogonal matrices," *J. Math. Phys.*, vol. 12, pp. 311–320, 1933.