ACKNOWLEDGMENT

The author wishes to thank the three anonymous reviewers for their valuable suggestions which greatly improved the quality and readability of this correspondence. In particular, the author would like to thank one of the reviewers for whose detailed advice and the suggestion on Theorem 4.

REFERENCES

[1] H. C. Ferreira and A. J. H. Vinck, "Interference cancellation with permutation trellis codes," in *Proc. IEEE Vehicular Technology Conf.*, Tokyo, Japan, May 2000, pp. 2401–2407.

[2] J.-C. Chang, R.-J. Chen, T. Kløve, and S.-C. Tsai, "Distance-preserving mappings from binary vectors to permutations," *IEEE Trans. Inf. Theory*, vol. 49, no. 4, pp. 1054–1059, Apr. 2003.

[3] K. Lee, "New distance preserving maps of odd length," *IEEE Trans. Inf. Theory*, vol. 50, no. 10, pp. 2539–2543, Oct. 2004.

[4] J.-C. Chang, "Distance-increasing mappings from binary vectors to permutations," *IEEE Trans. Inf. Theory*, vol. 51, no. 1, pp. 359–363, Jan. 2005.

Low-Resolution Scalar Quantization for Gaussian Sources and Squared Error

Daniel Marco, *Member, IEEE,* and David L. Neuhoff, *Fellow, IEEE*

***Abstract*—This correspondence analyzes the low-resolution performance of entropy-constrained scalar quantization. It focuses mostly on Gaussian sources, for which it is shown that for both binary quantizers and infinite-level uniform threshold quantizers, as D approaches the source variance σ^2, the least entropy of such quantizers with mean-squared error D or less approaches zero with slope $-\frac{\log_2 e}{2\sigma^2}$. As the Shannon rate-distortion function approaches zero with the same slope, this shows that in the low-resolution region, scalar quantization with entropy coding is asymptotically as good as any coding technique.**

***Index Terms*—Entropy constrained quantization, Gaussian, low rate, low resolution, scalar quantization, squared error.**

I. INTRODUCTION

This correspondence focuses on the rate-distortion performance of scalar quantization in the low-resolution regime where rate is small. While there are well known, asymptotically accurate, closed-form formulas for the rate-distortion performance of a variety of quantization schemes in the high resolution, i.e., high-encoding rate, regime, there is a shortage of similar formulas for the low resolution, i.e., low rate, regime. As a step in this direction, this correspondence focuses on scalar quantization with entropy coding and squared error distortion measure. For several sources, namely, exponential, Laplacian and uniform, the low-rate performance of such quantizers was found in or derives directly from previous work giving closed-form expressions for the operational rate-distortion function [1], [2]. The main contribution of this correspondence is the derivation of the low-rate performance for a memoryless Gaussian source and squared error distortion measure, for which no closed-form expressions exist or seem feasible.

Manuscript received June 1, 2004; revised December 27, 2005. This work was supported by the National Science Foundation under Grant ANI-0112801. This material in this correspondence was presented at the IEEE International Symposium on Information Theory, Chicago, IL, June/July 2004.

D. Marco was with the Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI 48109 USA. He is now with the Department of Electrical Engineering, California Institute of Technology, Pasadena, CA 91125 USA (e-mail: idaniel@caltech.edu).

D. L. Neuhoff is with the Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI 48109 USA (e-mail: neuhoff@eecs.umich.edu).

Communicated by M. Effros, Associate Editor for Source Coding.

Digital Object Identifier 10.1109/TIT.2006.871610

TABLE I
SLOPE DETERMINING FACTOR s OF THE OPERATIONAL RATE-DISTORTION FUNCTION $R(D)$ AT $D = \sigma^2$

exponential	Laplacian	uniform	Gaussian
0	0	∞	$\frac{\log_2 e}{2}$

To determine the low-resolution performance, we analyze the operational rate-distortion function, $R(D)$, of entropy-constrained scalar quantization in the low-rate region. We focus on squared-error distortion and stationary memoryless sources with absolutely continuous distributions, which are completely characterized by the probability density function (pdf) of an individual random variable. Accordingly, $R(D)$ is defined to be the least output entropy of any scalar quantizer with mean-squared error D or less. As it determines the optimal rate-distortion performance of this kind of quantization, it is important to understand how $R(D)$ depends on the source pdf and how it compares to the Shannon rate-distortion function. For example, the performance of conventional transform coding, which consists of an orthogonal transform followed by a scalar quantizer for each component of the transformed source vector, depends critically on the allocation of rate to component scalar quantizers, and the optimal rate allocation is determined by the operational rate-distortion functions of the components [3, p. 227].

While $R(D)$ can be determined numerically with various quantizer optimization algorithms [1], [4]–[11], closed-form formulas for $R(D)$ are known only for the exponential [1] and uniform [2] pdfs. A general closed form expression is known only for the high resolution, i.e., high rate, region [12], namely

$$R(D) = h - \frac{1}{2}\log(12\, D) + o_{D\to 0} \tag{1}$$

where $h = -\int_{-\infty}^{\infty} f(x)\log_2 f(x)\,dx$ is the differential entropy of the source being quantized, f is its pdf, and $o_{D\to x}$ denotes a quantity that goes to zero as $D \to x$.

In the low-resolution region, it is well understood that $R(D)$ approaches zero as D approaches the variance σ^2. Accordingly, one obtains a first-order approximation to $R(D)$ in this region by finding the slope of $R(D)$ at $D = \sigma^2$, namely

$$R(D) = s\left(1 - \frac{D}{\sigma^2}\right)\left[1 + o_{D\to\sigma^2}\right] \tag{2}$$

where σ^2 is the variance of f, s is a slope determining factor, namely, it is the magnitude of the slope with respect to distortion normalized by variance, and where the assumption throughout this correspondence is that in $o_{D\to\sigma^2}$, D goes to σ^2 from below.

The parametric formula of Sullivan [1] for the exponential pdf and of Gyorgy and Linder [2] for the uniform pdf imply $s = 0$ and $s = \infty$, respectively. Likewise, $s = 0$ for the Laplacian pdf [1]. Whereas these calculations are enabled by the special tractability of exponential and uniform pdfs, the principal result of this correspondence uses, primarily, tail behavior to find the slope determining factor for a Gaussian pdf, which turns out to be $\frac{\log_2 e}{2}$.

As a result, for the aforementioned pdfs, whose low resolution slope determining factors are summarized in Table I, we now have simple,

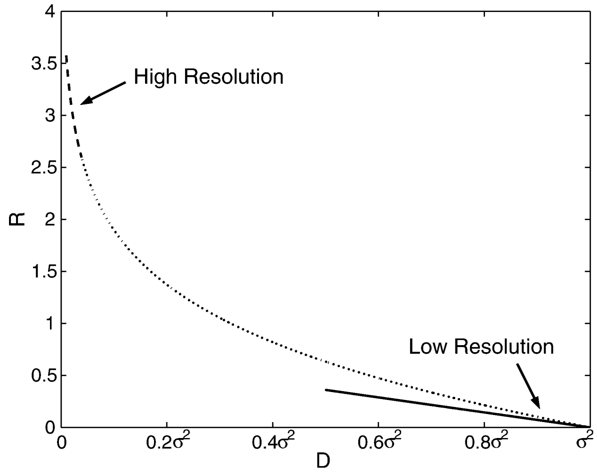

Fig. 1. The dotted line is a qualitative representation of the operational rate-distortion curve of scalar quantization. The dashed line indicates the section of the curve that is well described by (1). The solid line, which shows the tangent of the curve at $D = \sigma^2$, indicates the low resolution performance given by (2).

accurate approximations to the performance of optimal entropy-constrained scalar quantization in both the high- and low-resolution regions, as illustrated in Fig. 1.

It is interesting to compare the low resolution behavior of $R(D)$ for a given pdf to that of the Shannon rate-distortion function, denoted $\mathcal{R}(D)$, of a stationary memoryless source with the same pdf. Since $\mathcal{R}(D)$ represents the best performance attainable by any quantization technique, it must be that $\mathcal{R}(D) \leq R(D)$. It follows from this and the fact that, like $R(D)$, $\mathcal{R}(D) \to 0$ as $D \to \sigma^2$, that the magnitude of the slope of $\mathcal{R}(D)$ at $D = \sigma^2$ is no larger than that of $R(D)$.

We now see from Table I that since the slope of $R(D)$ at $D = \sigma^2$ equals 0 for exponential and Laplacian pdfs, it must equal the slope of the Shannon rate-distortion function (because the latter could be no larger). Therefore, in the low-resolution region, scalar quantization for these pdfs is asymptotically as good as any quantization technique—scalar, vector or otherwise. In contrast, in the high-resolution region, $R(D)$ exceeds $\mathcal{R}(D)$ by $\frac{1}{2}\log_2 \frac{\pi e}{6} = 0.255$ bits/sample, for all pdfs.

For the uniform pdf, the slope of $R(D)$ at $D = \sigma^2$ is infinite, whereas the slope of $\mathcal{R}(D)$ is finite (because it is convex). Therefore, for the uniform pdf at low rate, uniform scalar quantization is nowhere near as good as the best high-dimensional vector quantizers.

For the Gaussian pdf, the rate-distortion function is $\mathcal{R}(D) = \frac{1}{2}\log_2 \frac{\sigma^2}{D}$. Interestingly, as $D \to \sigma^2$, this approaches 0 with precisely the same slope as $R(D)$. Therefore, just as for exponential and Laplacian pdfs, in the low-resolution region scalar quantization for a Gaussian pdf is asymptotically as good as any quantization technique.

We conclude this introduction with a few additional comments. To derive the low resolution slope for a Gaussian pdf, we focus on uniform threshold quantizers with infinitely many cells, optimal reconstruction levels, and increasingly large step sizes Δ. While it is easy to see that, under ordinary conditions, distortion $D \approx \sigma^2$ and quantizer output entropy $H \approx 0$ when Δ is large, the slope at which H approaches 0 as $D \to \sigma^2$ is not obvious. Nevertheless, we find accurate approximations to D and H from which the low resolution slope can be straightforwardly determined.

Whereas the high-resolution formula (1) is based on the fact that the source density can be approximated by a constant on most sufficiently small cells, the low-resolution formula (2) is based on the fact that when the cells are large, the tail of the source probability density decays sufficiently quickly that only a few cells contribute materially to distortion and rate. We will show precisely which cells dominate the distortion and the entropy.

We also analyze binary quantization and show that it has low resolution performance characterized by the same slope. Thus, it, too, is asymptotically as good in the low-resolution region as any quantization technique for the Gaussian memoryless source.

As Laplacian and Gaussian are the two most commonly cited models for transform components (usually called coefficients) [13, pp. 215–218], [14, p. 564], the fact that scalar quantization is asymptotically as good for them as any type of quantization in the low resolution region has interesting ramifications for transform coding. In particular, in situations where transform coding is most effective, a sizable fraction of the coefficients must be coded at low rate. For such coefficients, simple scalar quantization is essentially as effective as any more sophisticated quantization technique. In contrast, to encode the coefficients that must be encoded with high resolution, scalar quantization requires approximately one quarter bit per sample more than optimal vector quantization.

We note that scalar quantization with fixed-rate coding does not attain the rate-distortion performance described in (2). This is because with fixed-rate coding, the smallest nonzero rate is at least 1, which implies that for any $D < \sigma^2$, the least rate of any fixed-rate scalar quantizer with mean-squared error D or less is at least 1. Consequently, the discussion throughout the correspondence is relevant to variable-rate coding, i.e., scalar quantization with entropy coding.

For completeness, we mention other analytical results on low-rate quantization of which we are aware. In [15], [16], the low resolution performance of fixed-rate transform codes is analyzed for Gaussian sources with memory, where low rate is attained with large block lengths. In [17], [18], upper bounds are found to the mean-squared error of dithered scalar quantization. Since these apply to all rates, they also apply to low rates. However, they have little use at asymptotically low rates, and in particular they give no indication of the slope of $R(D)$ at $D = \sigma^2$.

Finally, we note that the method used in this correspondence applies more widely. Specifically, as shown in [19] it can also be used for a Gaussian source with absolute error distortion measure as well as to a Laplacian source with both absolute and squared error distortion measures (the latter has been derived in a different way in [1]).

The remainder of the correspondence focuses on Gaussian sources and squared error distortion measure. Section II introduces and analyzes uniform scalar quantization. Section III does the same for binary quantizers. Section IV offers concluding remarks. Finally, proofs of certain lemmas are relegated to the Appendix , whereas the key results are proved in Sections II and III.

II. Uniform Threshold Quantizers

An infinite-level uniform threshold scalar quantizer (UTQ) with step size Δ and offset $0 \leq \alpha < 1$ is a scalar quantizer with partition having cells $S_k = [(k-\alpha)\Delta, (k+1-\alpha)\Delta), k \in \mathbb{Z}$, along with reconstruction levels $r_k \in S_k$, $k \in \mathbb{Z}$. Its quantization rule is $q(x) = r_k$, when $x \in S_k$. The offset α indicates the fraction of cell S_0 that lies to the left of the origin. For example, when $\alpha = 1/2$, cell S_0 is centered at the origin, whereas when $\alpha = 0$, cell S_0 begins at the origin. Let $\bar{\alpha} \triangleq 1 - \alpha$.

We assume throughout that the source to be quantized is stationary, memoryless and Gaussian with mean zero and variance σ^2, denoted $\mathcal{N}(0, \sigma^2)$ for short, (ordinarily we do not mention stationarity or memorylessness). The (output) entropy of this quantizer on such a source is

$$H(\alpha, \Delta, \sigma^2) = -\sum_{k=-\infty}^{\infty} P_k \log P_k$$

where P_k denotes the probability of the kth cell, and all logarithms in this correspondence have base 2. Since the source is Gaussian

$$P_k = Q\left((k-\alpha)\frac{\Delta}{\sigma}\right) - Q\left((k+1-\alpha)\frac{\Delta}{\sigma}\right)$$

where $Q(x) = \int_x^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}\, dt$. The mean-squared error of this quantizer on such a source is

$$d(\alpha, \Delta, \sigma^2) = \int_{-\infty}^{\infty} (x - q(x))^2 f(x)\, dx$$

where f is the Gaussian density. Except where stated otherwise, we take the reconstruction levels to be the centroids of their respective cells; i.e., $r_k = \int_{S_k} x \frac{f(x)}{P_k}\, dx$. As is well known, this choice minimizes distortion for a given partition. The operational rate-distortion function of infinite-level uniform threshold quantization is the function

$$R_{U,\sigma^2}(D) = \inf_{0\le\alpha<1, \Delta>0: d(\alpha,\Delta,\sigma^2)\le D} H(\alpha, \Delta, \sigma^2) \tag{3}$$

which specifies the least entropy of any such quantizer with mean-squared error D or less. Let $\mathcal{R}_{\sigma^2}(D) = \frac{1}{2}\log\frac{\sigma^2}{D}$ denote the Shannon rate-distortion function of the source.

We set λ to be the ratio Δ/σ and refer to it throughout as the normalized step size. We notice that P_k depends only on α and λ, and for emphasis, we will sometimes denote it $P_k(\alpha, \lambda)$. Consequently, $H(\alpha, \Delta, \sigma^2) = H(\alpha, \Delta/\sigma, 1)$, depends only on α and λ as well. Therefore, we will frequently use the notation $H(\alpha, \lambda)$. Similarly, $d(\alpha, \Delta, \sigma^2) = \sigma^2 d(\alpha, \Delta/\sigma, 1) = \sigma^2 d(\alpha, \lambda, 1)$. It follows from these remarks that $R_{U,\sigma^2}(D) = R_{U,1}\left(\frac{D}{\sigma^2}\right)$.

To find the slope of $R_{U,\sigma^2}(D)$ at $D = \sigma^2$ we need to consider what happens when $\sigma^2 d(\alpha, \lambda, 1) \to \sigma^2$ and $H(\alpha, \lambda) \to 0$. We observe that in order for the latter to happen it is necessary and sufficient that $\alpha\lambda \to \infty$ and $\bar{\alpha}\lambda \to \infty$. Moreover, because of this, for sufficiently large values of D, it suffices to restrict α to be greater than 0 in the definition of $R_{U,\sigma^2}(D)$ in (3).

Sections II-A and II-B find asymptotic low-resolution formulas for $H(\alpha, \lambda)$ and $\sigma^2 - d(\alpha, \Delta, \sigma^2)$, respectively, and Section II-C uses these to find an asymptotic expression for $R_{U,\sigma^2}(D)$ as $D \to \sigma^2$.

Before proceeding, we introduce notation and facts to be used later. Let $G(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ denote the Gaussian density with mean zero and unit variance. Let the entropy function be defined as

$$\mathcal{H}(\ldots, z_{-1}, z_0, z_1, \ldots) = -\sum_{k=-\infty}^{\infty} z_k \log z_k$$

where $0 < z_k \le 1$ for all k are a finite or countably infinite set of numbers that need not sum to one. Let $o_{x\to\infty}$ denote a quantity that converges to zero as $x \to \infty$. More generally, let $o_{x\to x_o, y\to y_o}$ denote a quantity that converges to zero as $x \to x_o$ and $y \to y_o$, where it will be clear from context whether $x \nearrow x_o$, $x \searrow x_o$ or simply $x \to x_o$, and similarly for the variable y. If this quantity depends on parameters other than x and y, its convergence to zero is uniform in such parameters. To keep notation short, we write o_x instead of $o_{x\to\infty}$, when $x_o = \infty$, and we let $o_{x,y}$ denote $o_{x\to\infty, y\to\infty}$.

The following facts provide elementary bounds and approximations to the Q function and closely related functions.

Fact 1: $Q(x) \le \sqrt{\frac{\pi}{2}}\, G(x),\ x \ge 0.$

Fact 2: $Q(x) < \frac{1}{x}\, G(x),\ x > 0.$

Fact 3: $Q(x) > \frac{1}{x}(1 - \frac{1}{x^2})\, G(x),\ x > 0.$

Fact 4: $Q(x) > \begin{cases} \frac{1}{2x}\, G(x), & x \ge \sqrt{2} \\ Q(\sqrt{2}), & x < \sqrt{2}. \end{cases}$

Fact 5: $Q(x) = \frac{1}{x}\, G(x)\left[1 + o_x\right],\ x > 0.$

Fact 6: $Q((x+1)\lambda) = Q(x\lambda)\, o_\lambda,\ x \ge 0$; i.e., $\frac{Q((x+1)\lambda)}{Q(x\lambda)} \to 0$ as $\lambda \to \infty$, uniformly for $x \ge 0$.

Fact 7: For all sufficiently large λ, $Q\Big((x+1)\lambda\Big) < \frac{1}{2}Q(x\lambda)$ for all $x \ge 0$.

Fact 8: For all sufficiently large λ,

$$Q(x\lambda) - Q((x+1)\lambda) > \begin{cases} \frac{1}{4x\lambda}\, G(x\lambda), & x\lambda \ge \sqrt{2} \\ \frac{Q(\sqrt{2})}{2}, & 0 \le x\lambda < \sqrt{2}. \end{cases}$$

Fact 9: $C(x) \triangleq \int_x^\infty t\, G(t)\, dt = G(x).$

Fact 10:

$$\begin{aligned} V(x) &\triangleq \int_x^\infty t^2\, G(t)\, dt \\ &= x\, G(x) + Q(x) = x\, G(x)\,[1 + o_x]. \end{aligned}$$

Fact 11: $C((x+1)\lambda) = C(x\lambda)\, o_\lambda,\ x \ge 0$; i.e., $\frac{C((x+1)\lambda)}{C(x\lambda)} \to 0$ as $\lambda \to \infty$, uniformly for $x \ge 0$.

Fact 12: $V((x+1)\lambda) = V(x\lambda)\, o_\lambda,\ x \ge 0$; i.e., $\frac{V((x+1)\lambda)}{V(x\lambda)} \to 0$ as $\lambda \to \infty$, uniformly for $x \ge 0$.

Facts 1, 2, and 3 are demonstrated in [20, pp. 82–83]. Fact 4 truncates the lower bound of Fact 3. Fact 5 follows from Facts 2 and 3. Fact 6 is derived by upper bounding $\frac{Q((x+1)\lambda)}{Q(x\lambda)}$ using Facts 1 and 4 when $x\lambda < \sqrt{2}$, and using Facts 2 and 4 when $x\lambda \ge \sqrt{2}$. Fact 7 follows from Fact 6, and Fact 8 follows from Facts 4 and 7. Fact 9 and the first equality of Fact 10 derive from elementary integration. The second equality of Fact 10 follows from Fact 5. Fact 11 follows from Fact 9 and simple manipulation of exponentials. Finally, Fact 12 is derived using Facts 10, 4, and 2. Specifically, when $x\lambda < \sqrt{2}$, Fact 4 is used to lower bound $V(x\lambda)$, and the fact that $x\lambda < \sqrt{2}$ is used to upper bound $V((x+1)\lambda)$. When $x\lambda \ge \sqrt{2}$, Fact 2 is used to upper bound $V((x+1)\lambda)$.

A. Asymptotic Entropy

We begin with several lemmas (one is proved here and the rest are proved in the Appendix) that lead to the main result of this section, a low resolution approximation for entropy.

Lemma 1: When a UTQ with offset α, $0 < \alpha < 1$, and sufficiently large normalized step size λ is applied to a $\mathcal{N}(0, \sigma^2)$ source

A. $P_{k+1}(\alpha, \lambda) < P_k(\alpha, \lambda) P_1(\alpha, \lambda)$ for all α and all $k \ge 1$

B. $P_{k-1}(\alpha, \lambda) < P_k(\alpha, \lambda) P_{-1}(\alpha, \lambda)$ for all α and all $k \le -1$.

Lemma 2:

$$\lim_{p\to 0} \frac{\mathcal{H}(1 - p + p\, o_{p\to 0})}{\mathcal{H}(p)} = 0.$$

We comment that this lemma is due to the fact that the entropy function $\mathcal{H}(p) = -p\log p$ has infinite slope at $p = 0$ and finite slope at $p = 1$, as illustrated in Fig. 2. A formal proof is provided in the appendix.

The next lemma shows that in low resolution, quantizer entropy is dominated by the cells adjacent to the center cell.

Lemma 3: For a UTQ with offset α, $0 < \alpha < 1$, and normalized step size λ applied to a $\mathcal{N}(0, \sigma^2)$ source

$$\mathcal{H}(\ldots, P_{-1}(\alpha, \lambda), P_0(\alpha, \lambda), P_1(\alpha, \lambda), \ldots) = \mathcal{H}\Big(P_{-1}(\alpha, \lambda), P_1(\alpha, \lambda)\Big)\left[1 + o_{\alpha\lambda, \bar{\alpha}\lambda}\right].$$

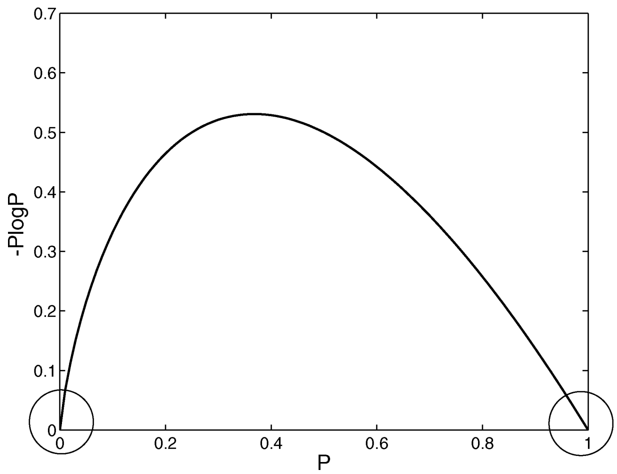

Fig. 2. The entropy function, $-p\log p$.

Proof: For brevity, we omit the parameters α and λ from $P_k(\alpha,\lambda)$. The proof is composed of two main steps. In Step 1, we show that $\mathcal{H}(\ldots,P_{-1},P_0,P_1,\ldots)$ can be asymptotically approximated by the three middle terms; that is,

$$\mathcal{H}(\ldots,P_{-1},P_0,P_1,\ldots)=\mathcal{H}\Big(P_{-1},P_0,P_1\Big)\Big[1+o_{\alpha\lambda,\bar{\alpha}\lambda}\Big].$$

In Step 2, it is shown that these three terms can be asymptotically approximated by only two terms; that is,

$$\mathcal{H}\Big(P_{-1},P_0,P_1\Big)=\mathcal{H}\Big(P_{-1},P_1\Big)\Big[1+o_{\alpha\lambda,\bar{\alpha}\lambda}\Big]$$

where we note that $P_0\to 1$ as $\alpha\lambda\to\infty$ and $\bar{\alpha}\lambda\to\infty$.

Step 1: We first show that for all sufficiently large $\alpha\lambda$ and $\bar{\alpha}\lambda$,

$$1<\frac{\mathcal{H}(\ldots,P_{-1},P_0,P_1,\ldots)}{\mathcal{H}(P_{-1},P_0,P_1)}<1+6P_1+6P_{-1}. \quad (4)$$

The left inequality is trivial. We upper bound the middle term in the following way:

$$\begin{aligned}&\frac{\sum_{k=-\infty}^{\infty}-P_k\log P_k}{\sum_{k=-1}^{1}-P_k\log P_k}\\ &\quad=1+\frac{\sum_{k=-\infty}^{-2}-P_k\log P_k+\sum_{k=2}^{\infty}-P_k\log P_k}{\sum_{k=-1}^{1}-P_k\log P_k(s)}\\ &\quad<1+\frac{\sum_{k=-\infty}^{-2}-P_k\log P_k}{-P_{-1}\log P_{-1}}+\frac{\sum_{k=2}^{\infty}-P_k\log P_k}{-P_1\log P_1}. \quad (5)\end{aligned}$$

Consider the terms in the last summation. We claim that when $\bar{\alpha}\lambda$ is sufficiently large, $-P_k\log P_k<-P_1^k\log P_1^k$ for all $k\geq 2$. To see this, we observe that when $\bar{\alpha}\lambda$ is large, so is λ (since $\bar{\alpha}<1$), and Lemma 1 implies $P_k<P_{k-1}P_1$ for all $k\geq 2$, which in turn implies $P_k<P_1^k<P_1$, for all $k\geq 2$. Next, Fact 1 implies that $P_1<Q(\bar{\alpha}\lambda)<\frac{1}{e}$ when $\bar{\alpha}\lambda$ is sufficiently large. Since $-p\log p$ increases for $p<\frac{1}{e}$, $-P_k\log P_k<-P_1^k\log P_1^k$ for all $k\geq 2$, when $\bar{\alpha}\lambda$ is sufficiently large. Substituting this into the last summation of (5), we have that when $\bar{\alpha}\lambda$ is sufficiently large

$$\begin{aligned}\frac{\sum_{k=2}^{\infty}-P_k\log P_k}{-P_1\log P_1}&<\frac{\sum_{k=2}^{\infty}-P_1^k\log P_1^k}{-P_1\log P_1}\\ &=\sum_{k=2}^{\infty}kP_1^{k-1}\\ &=\frac{2P_1}{(1-P_1)^2}-\frac{P_1^2}{(1-P_1)^2}\\ &<\frac{2P_1}{(1-P_1)^2}<6P_1\end{aligned}$$

where the last inequality derives from the fact that $P_1<\frac{1}{e}$. In much the same way it follows that when $\alpha\lambda$ is sufficiently large

$$\frac{\sum_{k=-\infty}^{-2}-P_k\log P_k}{-P_{-1}\log P_{-1}}<6P_{-1}.$$

This shows (4).Substituting $P_{-1}=o_{\alpha\lambda}$ and $P_1=o_{\bar{\alpha}\lambda}$ into (4), we obtain

$$\mathcal{H}(\ldots,P_{-1},P_0,P_1,\ldots)=\mathcal{H}\Big(P_{-1},P_0,P_1\Big)\Big[1+o_{\alpha\lambda,\bar{\alpha}\lambda}\Big]$$

which completes Step 1.

Step 2: We will show that $\mathcal{H}(P_0)=\mathcal{H}(P_1,P_{-1})\,o_{\alpha\lambda,\bar{\alpha}\lambda}$, from which it will follow that

$$\mathcal{H}\Big(P_{-1},P_0,P_1\Big)=\mathcal{H}\Big(P_{-1},P_1\Big)\Big[1+o_{\alpha\lambda,\bar{\alpha}\lambda}\Big].$$

Define $\widetilde{P}=\sum_{k=-\infty}^{-2}P_k+\sum_{k=2}^{\infty}P_k$. Using the fact that for all sufficiently large λ, $P_k<P_{k-1}P_1$ for all $k\geq 2$, we upper bound the second sum as

$$\sum_{k=2}^{\infty}P_k<\sum_{k=2}^{\infty}P_1^k=\frac{P_1^2}{1-P_1}<2P_1^2,$$

where the last inequality is due to $P_1<\frac{1}{2}$ for all sufficiently large $\bar{\alpha}\lambda$. The first sum in the definition of $\widetilde{P}$ can be upper bounded in much the same way, except that it holds for all sufficiently large $\alpha\lambda$. Thus when $\alpha\lambda$ and $\bar{\alpha}\lambda$ are both sufficiently large, $\widetilde{P}<2(P_{-1}^2+P_1^2)$. Therefore,

$$P_0>1-P_{-1}-P_1-2(P_{-1}^2+P_1^2)>1-P_{-1}-P_1-2(P_{-1}+P_1)^2.$$

Since $P_{-1}+P_1=o_{\alpha\lambda}+o_{\bar{\alpha}\lambda}$, it follows that when $\alpha\lambda$ and $\bar{\alpha}\lambda$ are sufficiently large, $P_0>1-P_{-1}-P_1-2(P_{-1}+P_1)^2>\frac{1}{e}$, which since $\mathcal{H}(p)$ decreases monotonically for $p>\frac{1}{e}$, implies that

$$\mathcal{H}(P_0)<\mathcal{H}\Big(1-P_{-1}-P_1-2(P_{-1}+P_1)^2\Big).$$

Consequently

$$\begin{aligned}\frac{\mathcal{H}(P_0)}{\mathcal{H}\Big(P_{-1},P_1\Big)}&<\frac{\mathcal{H}(1-P_{-1}-P_1-2(P_{-1}+P_1)^2\Big)}{\mathcal{H}(P_{-1},P_1)}\\ &<\frac{\mathcal{H}\Big(1-P_{-1}-P_1-2(P_{-1}+P_1)^2\Big)}{\mathcal{H}(P_{-1}+P_1)}\\ &=\frac{\mathcal{H}(1-p-2p^2)}{\mathcal{H}(p)}\end{aligned}$$

where $p\triangleq P_{-1}+P_1$, and where the second inequality follows from the easy to prove fact that for any $a,b\in\mathbb{R}^+$, $\mathcal{H}(a+b)<\mathcal{H}(a,b)$. We observe that as $\alpha\lambda$ and $\bar{\alpha}\lambda$ tend to infinity, p goes to zero. Therefore, by Lemma 2 it follows that $\frac{\mathcal{H}(1-p-2p^2)}{\mathcal{H}(p)}\to 0$ as $p\to 0$. This shows that $\mathcal{H}(P_0)=\mathcal{H}(P_1,P_{-1})\,o_{\alpha\lambda,\bar{\alpha}\lambda}$, which completes Step 2 and the proof of the lemma. $\square$

Lemma 4: Let $a(s)$ and $b(s)$ be positive functions on $\mathbb{R}$ such that $a(s)=b(s)[1+o_s]$, and for some $\varepsilon>0$, $|b(s)-1|>\varepsilon$ for all s. Then

$$\mathcal{H}(a(s))=\mathcal{H}(b(s))\Big[1+o_s\Big].$$

Lemma 5:

$$\mathcal{H}\Big(Q(x)\Big)=\frac{\log e}{2}x\,G(x)\Big[1+o_x\Big].$$

The following theorem gives the low resolution approximation to the entropy of uniform quantization.

Theorem 6: For a UTQ with offset α, $0<\alpha<1$, and normalized step size λ applied to a $\mathcal{N}(0,\sigma^2)$ source

$$H(\alpha,\lambda)=\frac{\log e}{2}\Big(\alpha\lambda\,G(\alpha\lambda)+\bar{\alpha}\lambda\,G(\bar{\alpha}\lambda)\Big)\Big[1+o_{\alpha\lambda,\bar{\alpha}\lambda}\Big] \quad (6)$$

where $H(\alpha,\lambda) = \mathcal{H}(\ldots, P_{-1}(\alpha,\lambda), P_0(\alpha,\lambda), P_1(\alpha,\lambda), \ldots)$ is the quantizer entropy.

If one fixes α, this theorem shows the rate at which entropy converges to 0 as $\lambda \to \infty$. However, the convergence is not uniform in α, and this theorem shows how entropy depends on α as well as λ. In particular, it gives an accurate approximation to quantizer entropy when both $\alpha\lambda$ and $\bar{\alpha}\lambda$ are large. Notice that $\alpha = 0$ is not allowed since $H(0,\lambda) = 1 + o_\lambda$, namely, the output entropy does not go to zero as $\lambda \to \infty$.

Proof: For brevity, we omit the parameters α and λ from $P_k(\alpha,\lambda)$. Lemma 3 shows that

$$H(\alpha,\lambda) = \mathcal{H}(\ldots, P_{-1}, P_0, P_1, \ldots) = \mathcal{H}\Big(P_{-1}, P_1\Big)\Big[1 + o_{\alpha\lambda,\bar{\alpha}\lambda}\Big]. \tag{7}$$

Since $P_{-1} = Q(\alpha\lambda) - Q\Big((1+\alpha)\lambda\Big)$, Fact 6 implies that $P_{-1} = Q(\alpha\lambda)[1 + o_\lambda]$, and thus in particular $P_{-1} = Q(\alpha\lambda)[1 + o_{\alpha\lambda}]$, since $0 < \alpha < 1$. Since $|Q(\alpha\lambda) - 1| > \frac{1}{2}$ for all $\alpha\lambda$, it follows from Lemma 4 that $\mathcal{H}(P_{-1}) = \mathcal{H}(Q(\alpha\lambda))\,[1 + o_{\alpha\lambda}]$. Next, applying Lemma 5, we obtain

$$\mathcal{H}(Q(\alpha\lambda)) = \left(\frac{1}{2}\log e\right)\alpha\lambda\, G(\alpha\lambda)\,\Big[1 + o_{\alpha\lambda}\Big].$$

Combining these yields

$$\mathcal{H}(P_{-1}) = \left(\frac{1}{2}\log e\right)\alpha\lambda\, G(\alpha\lambda)\,\Big[1 + o_{\alpha\lambda}\Big].$$

In a similar way

$$\mathcal{H}(P_1) = \left(\frac{1}{2}\log e\right)\bar{\alpha}\lambda\, G(\bar{\alpha}\lambda)\,\Big[1 + o_{\bar{\alpha}\lambda}\Big].$$

Combining the expressions for $\mathcal{H}(P_{-1})$ and $\mathcal{H}(P_1)$ together with (7) completes the proof of the theorem. □

We now comment on the cell or cells that dominate entropy when it is small. The entropy $H(\alpha,\lambda)$ will be small if and only if $P_0 \approx 1$ and $P_k \approx 0$, $k \neq 0$, which makes $-P_k \log P_k \approx 0$ for all k, and which happens if and only if $\alpha\lambda$ and $\bar{\alpha}\lambda$ are both large. Lemma 3 shows that $H(\alpha,\lambda)$ is dominated by the cells, S_{-1} and S_1, immediately adjacent to the center cell. This is not coincidental; rather, as mentioned earlier, it follows from the fact, illustrated in Fig. 2, that the entropy function, $\mathcal{H}(p) = -p\log p$, has infinite slope at $p = 0$ and finite slope at $p = 1$. Thus, when entropy is nearly zero, it is dominated by the largest of the nearly zero probabilities, which are P_{-1} and/or P_1. Indeed the two terms within the large parentheses in (6) correspond to $\mathcal{H}(P_{-1})$ and $\mathcal{H}(P_1)$, respectively. If $\alpha\lambda \ll \bar{\alpha}\lambda$, e.g., if $\alpha < \frac{1}{2}$ and λ is very large, then $P_{-1} \gg P_1$, and it is cell S_{-1} and the first term within the parentheses that dominate the entropy. Conversely, if $\bar{\alpha}\lambda \ll \alpha\lambda$, then $P_1 \gg P_{-1}$, and it is cell S_1 and the second term within the parentheses that dominate. Finally, if $\alpha\lambda \approx \bar{\alpha}\lambda$, then the two dominating cells contribute roughly the same to the entropy.

B. Asymptotic Distortion

The following theorem provides the low-resolution approximation to distortion.

Theorem 7: For a UTQ with offset α, $0 < \alpha < 1$, normalized step size λ, centroid reconstruction levels, and a $\mathcal{N}(0,\sigma^2)$ source

$$\frac{\sigma^2 - d(\alpha,\Delta,\sigma^2)}{\sigma^2} = \Big(\alpha\lambda\, G(\alpha\lambda) + \bar{\alpha}\lambda\, G(\bar{\alpha}\lambda)\Big)\Big[1 + o_{\alpha\lambda,\bar{\alpha}\lambda}\Big].$$

Like Theorem 6, this theorem gives an accurate approximation when both $\alpha\lambda$ and $\bar{\alpha}\lambda$ are large. The proof will be structured in a way that makes evident which cell or cells dominate $\sigma^2 - d(\alpha,\Delta,\sigma^2)$.

Proof: For brevity we omit the arguments of $d(\alpha,\Delta,\sigma^2)$. Let

$$\sigma_k^2 \triangleq \int_{S_k} x^2 f(x)\,dx = \sigma^2\Big(V((k-\alpha)\lambda) - V((k+1-\alpha)\lambda)\Big)$$

where f is the pdf of the Gaussian source and $V(x)$ is defined in Fact 10. Let

$$d_k \triangleq \int_{S_k} (x - r_k)^2 f(x)\,dx = \sigma_k^2 - r_k^2 P_k$$

be the contribution of the kth cell to the distortion (recalling that $r_k = \int_{S_k} x\frac{f(x)}{P_k}\,dx$). We observe that $\sigma^2 = \sum_k \sigma_k^2$ and $d = \sum_k d_k$. We now write

$$\sigma^2 - d = \big(\sigma^2 - d_0\big) - d_{-1} - d_1 - \sum_{|k|\geq 2} d_k. \tag{8}$$

We deal with these terms in reverse order. First, since $d_k \leq \sigma_k^2$

$$\begin{aligned}\sum_{|k|\geq 2} d_k \leq \sum_{|k|\geq 2}\sigma_k^2 &= \int_{-\infty}^{-(\alpha+1)\Delta} x^2 f(x)\,dx + \int_{(2-\alpha)\Delta}^{\infty} x^2 f(x)\,dx\\ &= \sigma^2 V((\alpha+1)\lambda) + \sigma^2 V((2-\alpha)\lambda)\\ &\overset{(a)}{=} \sigma^2 V(\alpha\lambda)\,o_\lambda + \sigma^2 V((1-\alpha)\lambda)\,o_\lambda\\ &\overset{(b)}{=} \sigma^2\alpha\lambda\, G(\alpha\lambda)\,o_{\alpha\lambda} + \sigma^2\bar{\alpha}\lambda\, G(\bar{\alpha}\lambda)\,o_{\bar{\alpha}\lambda}\end{aligned} \tag{9}$$

where (a) follows from Fact 12, and (b) is obtained using Fact 10. Next, we have (10) shown at the bottom of the page, where (a) is due to the definition of $C(x)$ given in Fact 10, (b) follows from Facts 6, 11, and 12, and (c) follows from Facts 5, 9, and 10. By symmetry it follows that

$$d_{-1} = \sigma^2\alpha\lambda\, G(\alpha\lambda)\,o_{\alpha\lambda}. \tag{11}$$

Finally

$$\sigma^2 - d_0 = (\sigma^2 - \sigma_0^2) + (\sigma_0^2 - d_0) \tag{12}$$

where as in (9) above

$$\begin{aligned}\sigma^2 - \sigma_0^2 = \sum_{k\neq 0}\sigma_k^2 &= \sigma^2 V(\alpha\lambda) + \sigma^2 V((1-\alpha)\lambda\Big)\\ &= \sigma^2\alpha\lambda G(\alpha\lambda)\Big[1 + o_{\alpha\lambda}\Big] + \sigma^2\bar{\alpha}\lambda G(\bar{\alpha}\lambda)\Big[1 + o_{\bar{\alpha}\lambda}\Big]\end{aligned} \tag{13}$$

$$\begin{aligned} d_1 &= \sigma_1^2 - r_1^2 P_1\\ &\overset{(a)}{=} \sigma^2(V((1-\alpha)\lambda) - V((2-\alpha)\lambda)) - \left(\frac{\sigma C((1-\alpha)\lambda) - \sigma C((2-\alpha)\lambda)}{Q((1-\alpha)\lambda) - Q((2-\alpha)\lambda)}\right)^2\Big(Q((1-\alpha)\lambda) - Q((2-\alpha)\lambda)\Big)\\ &\overset{(b)}{=} \sigma^2 V((1-\alpha)\lambda)\Big[1 + o_\lambda\Big] - \frac{\sigma^2\Big(C((1-\alpha)\lambda)\Big[1 + o_\lambda\Big]\Big)^2}{Q\Big((1-\alpha)\lambda\Big)\Big[1 + o_\lambda\Big]}\\ &\overset{(c)}{=} \sigma^2\bar{\alpha}\lambda\, G(\bar{\alpha}\lambda)\,[1 + o_{\bar{\alpha}\lambda}] - \frac{\sigma^2 G^2(\bar{\alpha}\lambda)\Big[1 + o_\lambda\Big]}{\frac{1}{\bar{\alpha}\lambda}G(\bar{\alpha}\lambda)\Big[1 + o_{\bar{\alpha}\lambda}\Big]} = \sigma^2\bar{\alpha}\lambda\, G(\bar{\alpha}\lambda)\,o_{\bar{\alpha}\lambda}\end{aligned} \tag{10}$$

where the third equality uses Fact 10, and where as in (10)

$$\begin{aligned}
&\sigma_0^2 - d_0 \\
&= r_0^2 P_0 \\
&= \left(\frac{\sigma C(\alpha\lambda) - \sigma C((1-\alpha)\lambda)}{1 - Q(\alpha\lambda) - Q((1-\alpha)\lambda)}\right)^2 \left(1 - Q(\alpha\lambda) - Q(\bar{\alpha}\lambda)\right) \\
&\overset{(a)}{=} \frac{\sigma^2\Big(G(\alpha\lambda) - G(\bar{\alpha}\lambda)\Big)^2}{1 + o_{\alpha\lambda} + o_{\bar{\alpha}\lambda}} \\
&= \frac{\sigma^2\Big(G(\alpha\lambda)\Big(G(\alpha\lambda) - G(\bar{\alpha}\lambda)\Big) + G(\bar{\alpha}\lambda)\Big(G(\bar{\alpha}\lambda) - G(\alpha\lambda)\Big)\Big)}{1 + o_{\alpha\lambda} + o_{\bar{\alpha}\lambda}} \\
&\overset{(b)}{=} \frac{\sigma^2\Big(\alpha\lambda G(\alpha\lambda) o_{\alpha\lambda} + \bar{\alpha}\lambda G(\bar{\alpha}\lambda) o_{\bar{\alpha}\lambda}\Big)}{1 + o_{\alpha\lambda} + o_{\bar{\alpha}\lambda}} \\
&= \Big(\sigma^2 \alpha\lambda\, G(\alpha\lambda) + \sigma^2 \bar{\alpha}\lambda\, G(\bar{\alpha}\lambda)\Big) o_{\alpha\lambda, \bar{\alpha}\lambda} \qquad (14)
\end{aligned}$$

where (a) is due to Fact 9, and (b) follows from having

$$G(\alpha\lambda) - G(\bar{\alpha}\lambda) = \alpha\lambda \frac{G(\alpha\lambda) - G(\bar{\alpha}\lambda)}{\alpha\lambda} = \alpha\lambda o_{\alpha\lambda}$$

and similarly

$$G(\bar{\alpha}\lambda) - G(\alpha\lambda) = \bar{\alpha}\lambda o_{\bar{\alpha}\lambda}.$$

Substituting (13) and (14) into (12) yields

$$\sigma^2 - d_0 = \Big(\sigma^2 \alpha\lambda\, G(\alpha\lambda) + \sigma^2 \bar{\alpha}\lambda\, G(\bar{\alpha}\lambda)\Big) \Big[1 + o_{\alpha\lambda, \bar{\alpha}\lambda}\Big]. \qquad (15)$$

Substituting (9), (10), (11), and (15) into (8) yields

$$\sigma^2 - d = \sigma^2 \Big(\alpha\lambda\, G(\alpha\lambda) + \bar{\alpha}\lambda\, G(\bar{\alpha}\lambda)\Big) \Big[1 + o_{\alpha\lambda, \bar{\alpha}\lambda}\Big].$$

Dividing the above by σ^2 gives the desired result. □

We now consider which cell or cells make the dominating contribution to $\sigma^2 - d$, when the latter is very small. When $d \approx \sigma^2$, both $\alpha\lambda$ and $\bar{\alpha}\lambda$ are large. From (15), we see that $d_0 \approx \sigma^2$, and from (9)–(11), we see that $d_k \approx 0$ for $k \neq 0$. We are interested, however, in finding the cells that dominate the rate at which distortion converges to variance. Since $d_0 \to \sigma^2$ and $d_k \to 0$, $k \neq 0$, it makes most sense to compare $\sigma^2 - d_0$ and the sum of the d_k's, $k \neq 0$. Comparing (15) to (9)–(11), reveals that $\sum_{k \neq 0} d_k$ is asymptotically negligible relative to $\sigma^2 - d_0$. We conclude that when $d \approx \sigma^2$, $\sigma^2 - d_0$ is the dominant component of $\sigma^2 - d$.

C. Asymptotic Rate-Distortion

The following lemma, which is used in the proof of Theorem 9 below, derives directly from Theorems 6 and 7.

Lemma 8: For a UTQ with offset α, $0 < \alpha < 1$, normalized step size λ, centroid reconstruction levels, and a $\mathcal{N}(0, \sigma^2)$ source

$$\frac{H(\alpha, \lambda)}{\sigma^2 - d(\alpha, \Delta, \sigma^2)} = \frac{\log e}{2\sigma^2} \Big[1 + o_{\alpha\lambda, \bar{\alpha}\lambda}\Big]$$

The following is the principal result of this correspondence.

Theorem 9: In the low resolution region, the operational rate-distortion function of infinite-level uniform threshold scalar quantization for a Gaussian source with variance σ^2 is

$$R_{U,\sigma^2}(D) = \frac{\log e}{2}\left(1 - \frac{D}{\sigma^2}\right)\Big[1 + o_{D \to \sigma^2}\Big]. \qquad (16)$$

Proof: Since $R_{U,\sigma^2}(D) = R_{U,1}\left(\frac{D}{\sigma^2}\right)$, it suffices to show

$$\lim_{D \to 1} \frac{R_{U,1}(D)}{1 - D} = \frac{\log e}{2}. \qquad (17)$$

Next, we rewrite the operational rate-distortion function for a unit variance source for all sufficiently large D as

$$R_{U,1}(D) = \inf_{0 < \alpha < 1} R_{U,1,\alpha}(D)$$

where

$$R_{U,1,\alpha}(D) \triangleq \inf_{\Delta > 0 : d(\alpha, \Delta, 1) \leq D} H(\alpha, \Delta)$$

is the operational rate-distortion function of UTQ with fixed offset α for a source with unit variance, and as mentioned before, $\alpha = 0$ can be omitted from the constraint since D is sufficiently large. As a preliminary to showing (17), we will show $R_{U,1,\alpha}(D)$ satisfies (17) for any fixed $\alpha \in (0, 1)$. First

$$\begin{aligned}
\limsup_{D \to 1} \frac{R_{U,1,\alpha}(D)}{1 - D} &\overset{(a)}{=} \limsup_{\lambda \to \infty} \frac{R_{U,1,\alpha}\Big(d(\alpha, \lambda, 1)\Big)}{1 - d(\alpha, \lambda, 1)} \\
&\overset{(b)}{\leq} \limsup_{\lambda \to \infty} \frac{H(\alpha, \lambda)}{1 - d(\alpha, \lambda, 1)} \\
&\overset{(c)}{=} \frac{\log e}{2} \qquad (18)
\end{aligned}$$

where (a) derives from the fact that $d(\alpha, \lambda, 1)$ goes continuously to 1 as $\lambda \to \infty$, (b) follows from the definition of $R_{U,1,\alpha}\Big(d(\alpha, \lambda, 1)\Big)$, and (c) is obtained from Lemma 8. Next

$$\liminf_{D \to 1} \frac{R_{U,1,\alpha}(D)}{1 - D} \overset{(a)}{\geq} \liminf_{\lambda \to \infty} \frac{H(\alpha, \lambda)}{1 - d(\alpha, \lambda, 1)} \overset{(b)}{=} \frac{\log e}{2} \qquad (19)$$

where (b) is obtained from Lemma 8 and (a) is shown as follows. By the definition of $R_{U,1,\alpha}(D)$, for any $D \in (0, 1)$, there exists $\lambda(D)$ such that

$$H(\alpha, \lambda(D)) \leq R_{U,1,\alpha}(D) + \varepsilon(D) \text{ and } d(\alpha, \lambda(D), 1) \leq D \qquad (20)$$

where $\varepsilon(D) > 0$ and $\lim_{D \to 1} \frac{\varepsilon(D)}{1 - D} = 0$. (The choices of $\varepsilon(D)$ and $\lambda(D)$ are not unique, but any fixed choices will do.) From (18) it follows that $R_{U,1,\alpha}(D) \to 0$ as $D \to 1$. Thus, by (20)$H(\alpha, \lambda(D)) \to 0$ as $D \to 1$, which implies that $\lambda(D) \to \infty$ as $D \to 1$, since $H(\alpha, \lambda) \to 0$ if and only if $\lambda \to \infty$. This and (20) yield

$$\begin{aligned}
\liminf_{D \to 1} \frac{R_{U,1,\alpha}(D)}{1 - D} &\geq \liminf_{D \to 1} \frac{H(\alpha, \lambda(D)) - \varepsilon(D)}{1 - d(\alpha, \lambda(D), 1)} \\
&\geq \liminf_{\lambda \to \infty} \frac{H(\alpha, \lambda)}{1 - d(\alpha, \lambda, 1)}.
\end{aligned}$$

It now follows from (18) and (19) that

$$\lim_{D \to 1} \frac{R_{U,1,\alpha}(D)}{1 - D} = \frac{\log e}{2}. \qquad (21)$$

Finally, to obtain the result of the theorem we proceed as follows:

$$\limsup_{D\to 1} \frac{R_{U,1}(D)}{1-D} \overset{(a)}{=} \limsup_{D\to 1} \frac{\inf_\alpha R_{U,1,\alpha}(D)}{1-D} \overset{(b)}{\le} \inf_\alpha \limsup_{D\to 1} \frac{R_{U,1,\alpha}(D)}{1-D} \overset{(c)}{=} \frac{\log e}{2} \tag{22}$$

where (a) follows from the definition of $R_{U,1}(D)$, (b) is elementary, and (c) follows from (21). Similarly

$$\liminf_{D\to 1} \frac{R_{U,1}(D)}{1-D} \overset{(a)}{\ge} \liminf_{D\to 1} \frac{\mathcal{R}_1(D)}{1-D} \overset{(b)}{=} \liminf_{D\to 1} \frac{\frac{1}{2}\log\frac{1}{D}}{1-D} \overset{(c)}{=} \frac{\log e}{2} \tag{23}$$

where $\mathcal{R}_1(D)$ is the Shannon rate-distortion function of a unit variance Gaussian source, and where (a) follows from the converse rate-distortion theorem, (b) uses the well-known formula for $\mathcal{R}_1(D)$ [21, p. 477], and (c) is obtained by taking the limit, for example, by using L'Hospital's rule. Equation (17) and the theorem now follow from (22) and (23). We note that (19) could have been shown using Shannon's rate-distortion function as a lower bound, as was done in (23). However, the approach taken above, demonstrates that

$$\lim_{D\to 1} \frac{R_{U,1,\alpha}(D)}{1-D} = \lim_{\lambda\to\infty} \frac{H(\alpha,\lambda)}{1-d(\alpha,\lambda,1)}$$

without using either Gaussianity or Shannon's rate-distortion function. It requires only that the latter limit exist. □

As it is easy to see,

$$\mathcal{R}_{\sigma^2}(D) = \frac{1}{2}\log\frac{\sigma^2}{D} = \frac{\log e}{2}\left[1-\frac{D}{\sigma^2}\right]\left[1+o_{D\to\sigma^2}\right].$$

Comparing this to the theorem statement reveals that for a Gaussian source and the low resolution region, the operational rate-distortion function of infinite-level uniform threshold scalar quantization and the Shannon rate-distortion function approach 0 with the same slope as $D \to \sigma^2$. Therefore, in the low resolution region, such quantizers are asymptotically as good as any quantization technique — scalar, block, or otherwise. Additionally, from (21) and the relation between $R_{U,\sigma^2,\alpha}(D)$ and $R_{U,1,\alpha}(D)$, one concludes that for any $\alpha \ne 0$, the operational rate-distortion function $R_{U,\sigma^2,\alpha}(D)$ of uniform threshold quantization with offset α also approaches zero with the same slope as the Shannon rate-distortion function, as does the operational rate-distortion function of scalar quantization in general. Finally, we note that from the dominance results presented previously, the slope is approximately equal to $\frac{H(P_{-1})+H(P_1)}{\sigma^2-D_0}$, i.e., the distortion term is dominated by the center cell and the entropy is dominated by the two adjacent cells.

Remark: We assumed throughout that the quantizers' reconstruction levels were centroids, which is necessary for optimal performance. It turns out, however, that this assumption can be relaxed somewhat, asymptotically in the low-resolution region. That is, for Theorem 9 to hold it is only necessary that the reconstruction levels be sufficiently close to the centroids. More specifically, there is very little sensitivity to the reconstruction levels for $k \ne -1, 0, 1$, in the sense that the requirement that $r_k \in S_k$ is sufficient. For $k = -1, 0, 1$, the requirement depends on the behavior of $\alpha\lambda$ relative to $\bar{\alpha}\lambda$ as both quantities tend to infinity. If $\alpha\lambda \ll \bar{\alpha}\lambda$, then r_{-1} needs to be close to the centroid of S_{-1} and there is no restriction on r_1 (except for lying in S_1). Similarly, if $\bar{\alpha}\lambda \ll \alpha\lambda$, then r_1 needs to be close to the centroid of S_1 and there is no restriction on r_{-1} (except for lying in S_{-1}). Lastly, if $\alpha\lambda \approx \bar{\alpha}\lambda$, then both r_{-1} and r_1 need to be close to the centroids of S_{-1} and S_1, respectively. In all cases, r_0 needs to be sufficiently close to zero. A formal derivation of these statements can be found in [22, pp. 98–104].

III. Binary Quantizers

A binary (two-level) scalar quantizer is characterized by three numbers: a threshold t and two reconstruction levels $r_0 < t$ and $r_1 \ge t$. Let $S_0(t) = (-\infty, t)$ and $S_1(t) = [t, \infty)$ be the two quantization cells, and let the quantization rule be $q(x) = r_k$ when $x \in S_k$, $k = 0, 1$.

As in the previous section, the source considered is stationary, memoryless Gaussian with mean zero and variance σ^2, and the reconstruction levels r_0 and r_1 are taken to be the cell centroids, unless otherwise specified. We let P_k or $P_k(t,\sigma^2)$ denote the probability of the source value lying in S_k, $k = 0, 1$.

Let $H(t,\sigma^2) = \mathcal{H}(P_0(t,\sigma^2), P_1(t,\sigma^2))$ denote the entropy of the quantizer output with threshold t for the Gaussian source. Let

$$d(t,\sigma^2) \triangleq \int_{-\infty}^{\infty} (x-q(x))^2 f(x)\,dx$$

denote the mean-squared error distortion of this quantizer on this source. The operational rate-distortion function of binary quantization for this source is $R_{B,\sigma^2}(D) = \inf_{t:d(t,\sigma^2)\le D} H(t,\sigma^2)$, which specifies the least entropy of any such quantizer with mean-squared error D or less.

It is easy to see that $P_k(t,\sigma^2) = P_k(\frac{t}{\sigma},1)$, $H(t,\sigma^2) = H(\frac{t}{\sigma},1)$, $d(t,\sigma^2) = \sigma^2 d(\frac{t}{\sigma},1)$, and $R_{B,\sigma^2}(D) = R_{B,1}(\frac{D}{\sigma^2})$. Hence it is convenient to parameterize P_k and H by $\lambda = \frac{t}{\sigma}$, i.e., $P_k(\lambda)$ and $H(\lambda)$. Due to the symmetry of the Gaussian density, it suffices to restrict attention to $t > 0$.

As before, we will find asymptotic low-resolution approximations to entropy and distortion, and then combine these to determine the asymptotic low resolution expression for $R_{B,\sigma^2}(\mathrm{D})$. We also determine which cells dominate entropy and distortion, and we relax the requirement that the levels be centroids. Since the derivations in the binary case are similar in spirit to those in the uniform case, we will only state the results and provide no proofs, so as to spare the reader repetitive details.

Theorem 10: For a binary scalar quantizer with threshold t applied to a $\mathcal{N}(0,\sigma^2)$ source

$$H(\lambda) = \frac{\log e}{2}\lambda\, G(\lambda)\left[1+o_\lambda\right],$$

where as indicated earlier $\lambda = \frac{t}{\sigma}$, $H(\lambda) = \mathcal{H}(P_1(\lambda), P_2(\lambda))$ and $G(x)$ denotes a zero-mean, unit-variance Gaussian density.

Theorem 11: For a binary scalar quantizer with threshold t and reconstruction levels at cell centroids applied to a $\mathcal{N}(0,\sigma^2)$ source

$$\frac{\sigma^2 - d(t,\sigma^2)}{\sigma^2} = \lambda\, G(\lambda)\left[1+o_\lambda\right]$$

where $\lambda = \frac{t}{\sigma}$.

Theorem 12: In the low resolution region, the operational rate-distortion function of binary scalar quantization for a Gaussian source with variance σ^2 is

$$R_{B,\sigma^2}(D) = \frac{\log e}{2}\left(1 - \frac{D}{\sigma^2}\right)\left[1 + o_{D\to\sigma^2}\right].$$

Notice that the expression given in this theorem for binary quantization is precisely the same as that given in Theorem 9 for infinite-level uniform threshold quantization, which in turn matches the Shannon rate-distortion function in the low resolution region. We conclude that binary quantization is another type of quantization that is asymptotically optimal in the low resolution region.

We now comment on the cells that dominate the entropy and distortion. As before, when entropy is small, it is dominated by the cell that has largest probability not close to one, which is S_1. And just as with uniform quantizers, when distortion is close to σ^2, $\sigma^2 - D$ is dominated by the cell whose probability is nearly one, namely, S_0. That is, $\frac{\sigma^2 - D_0}{\sigma^2 - D} \approx 1$.

As with uniform quantizers, the requirement for cell centroids can be relaxed somewhat. Specifically, the reconstruction levels need not lie exactly at cell centroids, but they do need to be sufficiently close to them. This is formally stated in [22, pp. 106–107].

IV. Conclusion

This correspondence considered the asymptotic performance of scalar quantizers in the low-resolution domain, which is determined by the slope of the operational rate-distortion function of such quantizers at distortion equal σ^2. For the cases of exponential, Laplacian and uniform sources and difference distortion measures, this slope has been provided in or can be determined from [1], [2]. The focus of this correspondence has been on the Gaussian source and squared error distortion measure. We considered infinite-level uniform threshold and binary scalar quantizers in the asymptotic case that the cell sizes tend to infinity (for the uniform case) and that the quantizer threshold tends to infinity (for the binary case). We derived simple formulas for the rate of convergence of entropy to zero and of mean-squared error distortion to the source variance.

The convergence of entropy and distortion as $\lambda \to \infty$ for uniform quantization is not uniform in the offset α. The derived formulas show how entropy and distortion depend on α as well as λ. Specifically, they provide accurate approximations when both $\alpha\lambda$ and $\bar{\alpha}\lambda$ are large.

Using these convergence formulas, the operational rate-distortion function of infinite-level uniform threshold and binary scalar quantization has been shown to approach zero as $D \to \sigma^2$ with the same slope as that of the Shannon rate-distortion function. This shows that in the low resolution region scalar quantization (in particular uniform and binary) is asymptotically optimal for Gaussian sources and squared error distortion measure.

Last, the method used in this correspondence can also be applied to a Gaussian source with absolute error distortion measure and to a Laplacian source with both absolute and squared error distortion measures [19].

Appendix

A. Proof of Lemma 1

We will show Part A; Part B follows by symmetry. To simplify notation, we omit the parameters α and λ from $P_k(\alpha, \lambda)$. Consider $k \geq 1$. First, Fact 8, with $(k - \alpha)$ playing the role of x, shows that for all sufficiently large λ, the following lower bound to P_k holds for all $k \geq 1$:

$$\begin{aligned} P_k &= Q\Big((k-\alpha)\lambda\Big) - Q\Big((k+1-\alpha)\lambda\Big) \\ &> \begin{cases} \frac{1}{4}\frac{1}{(k-\alpha)\lambda}G((k-\alpha)\lambda), & (k-\alpha)\lambda \geq \sqrt{2} \quad (a) \\ \frac{Q(\sqrt{2})}{2}, & (k-\alpha)\lambda < \sqrt{2} \quad (b). \end{cases} \end{aligned} \tag{A1}$$

Next, we upper bound P_{k+1} using Fact 2

$$\begin{aligned} P_{k+1} &= Q\Big((k+1-\alpha)\lambda\Big) - Q\Big((k+2-\alpha)\lambda\Big) \\ &< \frac{1}{(k+1-\alpha)\lambda} G((k+1-\alpha)\lambda). \end{aligned}$$

Combining the lower bound to P_k with the upper bound to P_{k+1}, we obtain

$$\frac{P_{k+1}}{P_k} < \begin{cases} 4e^{-\frac{(2(k-\alpha)+1)\lambda^2}{2}}, & (k-\alpha)\lambda \geq \sqrt{2} \quad (a) \\ \frac{2\,G((k+1-\alpha)\lambda)}{Q(\sqrt{2})(k+1-\alpha)\lambda}, & (k-\alpha)\lambda < \sqrt{2} \quad (b). \end{cases} \tag{A2}$$

It now suffices to show that for all sufficiently large λ, the above upper bound to $\frac{P_{k+1}}{P_k}$ is smaller than the lower bound to P_1 obtained from (A1). We do so by considering two cases.

Case 1: $(k-\alpha)\lambda < \sqrt{2}$—In this case, $(1-\alpha)\lambda < \sqrt{2}$. Thus, by (A1b), $P_1 > \frac{Q(\sqrt{2})}{2}$. Next, by (A2b),

$$\frac{P_{k+1}}{P_k} < \frac{2\,G((k+1-\alpha)\lambda)}{Q(\sqrt{2})(k+1-\alpha)\lambda} < \frac{2\,G(\lambda)}{Q(\sqrt{2})\lambda}$$

where the last inequality uses $k + 1 - \alpha > 1$. Since $P_1 > \frac{Q(\sqrt{2})}{2}$ and $\frac{P_{k+1}}{P_k} < \frac{2\,G(\lambda)}{Q(\sqrt{2})\lambda} \to 0$ as $\lambda \to \infty$, we conclude that for all sufficiently large λ, $\frac{P_{k+1}}{P_k} < P_1$, for all k, α such that $(k-\alpha)\lambda < \sqrt{2}$.

Case 2: $(k-\alpha)\lambda \geq \sqrt{2}$ — We consider two subcases. First, suppose $(1-\alpha)\lambda < \sqrt{2}$. Then by (A1b), $P_1 > \frac{Q(\sqrt{2})}{2}$. Next, by (A2a)

$$\frac{P_{k+1}}{P_k} < 4e^{-\frac{(2(k-\alpha)+1)\lambda^2}{2}} < 4e^{-\frac{\lambda^2}{2}}.$$

We conclude that for all sufficiently large λ, $\frac{P_{k+1}}{P_k} < P_1$, for all of the k, α such that $(k-\alpha)\lambda \geq \sqrt{2}$ and $(1-\alpha)\lambda < \sqrt{2}$. Next, suppose $(1-\alpha)\lambda \geq \sqrt{2}$. Then by (A1a)

$$P_1 > \frac{1}{4}\frac{1}{(1-\alpha)\lambda} G((1-\alpha)\lambda) > \frac{1}{4}\frac{1}{\lambda} G(\lambda)$$

using $1 - \alpha < 1$. By (A2a)

$$\frac{P_{k+1}}{P_k} < 4e^{-\frac{(2(k-\alpha)+1)\lambda^2}{2}} < 4e^{-\frac{\lambda^2}{2}} e^{-\sqrt{2}\lambda}$$

using $(k-\alpha)\lambda \geq \sqrt{2}$. Since $e^{-\sqrt{2}\lambda} \to 0$ faster than $\frac{1}{\lambda} \to 0$, we conclude that for all sufficiently large λ, $\frac{P_{k+1}}{P_k} < P_1$, for all k, α such that $(k-\alpha)\lambda \geq \sqrt{2}$ and $(1-\alpha)\lambda \geq \sqrt{2}$. This completes the proof of Part A and the lemma. □

B. Proof of Lemma 2

We need to show that

$$\lim_{p\to 0} \frac{-(1-p+p\,o_{p\to 0})\ln(1-p+p\,o_{p\to 0})}{-p\ln p} = 0.$$

The fact that $\lim_{x\to 0} \frac{\ln(1-x)}{-x} = 1$, or equivalently, that $\frac{\ln(1-x)}{-x} = 1 + o_{x\to 0}$, implies

$$\begin{aligned}
&\frac{-(1-p+p\,o_{p\to 0})\ln(1-p+p\,o_{p\to 0})}{-p\ln p}\\
&= \left[\frac{-(1-p+p\,o_{p\to 0})\ln(1-p+p\,o_{p\to 0})}{(1-p+p\,o_{p\to 0})(p+p\,o_{p\to 0})}\right]\\
&\cdot \left[\frac{(1-p+p\,o_{p\to 0})(p+p\,o_{p\to 0})}{-p\ln p}\right]\\
&= \left[1+o_{p\to 0}\right]\cdot\left[\left(1-p+p\,o_{p\to 0}\right)\frac{1+o_{p\to 0}]}{-\ln p}\right]\longrightarrow 0 \text{ as } p\to 0
\end{aligned}$$

which proves the lemma. □

C. Proof of Lemma 4

It is sufficient to show

$$\lim_{s\to s_o} \frac{\mathcal{H}(a(s))}{\mathcal{H}(b(s))} = 1.$$

We have the following string of equalities:

$$\begin{aligned}
\frac{\mathcal{H}(a(s))}{\mathcal{H}(b(s))} &= \frac{-a(s)\log a(s)}{-b(s)\log b(s)} = \frac{a(s)}{b(s)}\frac{\log\left[\frac{a(s)}{b(s)}b(s)\right]}{\log b(s)}\\
&= \frac{a(s)}{b(s)}\left[1+\frac{\log\frac{a(s)}{b(s)}}{\log b(s)}\right] = \frac{a(s)}{b(s)} + \frac{\frac{a(s)}{b(s)}\log\frac{a(s)}{b(s)}}{\log b(s)}.
\end{aligned}$$

Since $|b(s)-1| > \varepsilon$ for all s, it follows that either $\log b(s) > \log(1+\varepsilon)$ or $\log b(s) < \log(1-\varepsilon)$ for all s. Therefore, $\log b(s)$ is bounded away from zero. Combining this with the fact that $\frac{a(s)}{b(s)} \to 1$ as $s\to\infty$, and that $\frac{a(s)}{b(s)}\log\frac{a(s)}{b(s)} \to 0$ as $s\to\infty$, the result follows. □

D. Proof of Lemma 5

The lemma is obtained by using Fact 5 in the first equality as follows:

$$\begin{aligned}
\mathcal{H}\Big(Q(x)\Big) &= \mathcal{H}\left(\frac{1}{x}\,G(x)\Big[1+o_x\Big]\right)\\
&= -\frac{1}{x}\,G(x)\Big[1+o_x\Big]\log\left(\frac{1}{x}\,G(x)\Big[1+o_x\Big]\right)\\
&= \frac{1}{x}G(x)\Big[1+o_x\Big]\Big[\log\sqrt{2\pi}x+\frac{x^2}{2}\log e-\log\Big[1+o_x\Big]\Big]\\
&= \frac{\log e}{2}x\,G(x)\Big[1+o_x\Big].
\end{aligned}$$

□

References

[1] G. J. Sullivan, "Efficient scalar quantization of exponential and Laplacian random variables," *IEEE Trans. Inf. Theory*, vol. 42, no. 5, pp. 1365–1374, Sep. 1996.
[2] A. Gyorgy and T. Linder, "Optimal entropy-constrained scalar quantization of a uniform source," *IEEE Trans. Inf. Theory*, vol. 46, no. 7, pp. 2704–2711, Nov. 2000.
[3] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*. Boston, MA: Kluwer, 1992.
[4] T. Berger, "Optimum quantizers and permutation codes," *IEEE Trans. Inf. Theory*, vol. IT-18, no. 6, pp. 759–765, Nov. 1972.
[5] A. N. Netravali and R. Saigal, "Optimal quantizer design using a fixed-point algorithm," *Bell Syst. Tech. J.*, vol. 55, pp. 1423–1435, Nov. 1976.
[6] P. Noll and R. Zelinski, "Bounds on quantizer performance in the low bit-rate region," *IEEE Trans. Commun.*, vol. COM-26, no. 2, pp. 300–305, Feb. 1978.
[7] N. Farvardin and J. W. Modestino, "Optimum quantizer performance for a class of non-Gaussian memoryless sources," *IEEE Trans. Inf. Theory*, vol. IT-30, no. 3, pp. 485–497, May 1984.
[8] J. C. Kieffer, T. M. Jahns, and V. A. Obuljen, "New results on optimal entropy-constrained quantization," *IEEE Trans. Inf. Theory*, vol. 34, no. 5, pp. 1250–1258, Sep. 1988.
[9] P. A. Chou, T. Lookabaugh, and R. M. Gray, "Entropy-constrained vector quantization," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 1, pp. 31–42, Jan. 1989.
[10] J. Buhmann and H. Kuhnel, "Vector quantization with complexity costs," *IEEE Trans. Inf. Theory*, vol. 39, no. 4, pp. 1133–1145, Jul. 1993.
[11] K. Rose and D. Miller, "A deterministic annealing algorithm for entropy-constrained vector quantizer design," in *Proc. Asilomar Conf. Signals, Systems, and Computing*, vol. 2, Pacific Grove, CA, Nov. 1993, pp. 1651–1655.
[12] H. Gish and J. N. Pierce, "Asymptotically efficient quantization," *IEEE Trans. Inf. Theory*, vol. IT-14, no. 5, pp. 676–683, Sep. 1968.
[13] A. N. Netravali and B. G. Haskell, *Digital Pictures: Representation, Compression and Standards*, 2nd ed. New York: Plenum, 1988.
[14] N. S. Jayant and P. Noll, *Digital Coding of Waveforms: Principles and Applications to Speech and Video*. Englewood Cliffs, NJ: Prentice-Hall, 1988.
[15] D. F. Lyons and D. L. Neuhoff, "A coding theorem for low-rate transform codes," in *Proc. IEEE Int. Symp. Information Theory*, San Antonio, TX, Jan. 1993, p. 333.
[16] D. F. Lyons, "Fundamental Limits of Low-Rate Transform Codes," Ph.D. dissertation, EECS Dept., Univ. Michigan, Ann Arbor, 1992.
[17] J. Ziv, "On universal quantization," *IEEE Trans. Inf. Theory*, vol. IT-31, no. 3, pp. 344–347, May 1985.
[18] R. Zamir and M. Feder, "On universal quantization by randomized uniform/lattice quantizers," *IEEE Trans. Inf. Theory*, vol. 38, no. 2, pp. 428–436, Mar. 1992.
[19] D. Marco and D. L. Neuhoff, "Low Resolution Scalar Quantization for Gaussian and Laplacian Sources with Absolute and Squared Error Distortion Measures," Univ. Michigan, Ann Arbor, CSPL Tech. Rep., TR 372, Jan. 2006.
[20] J. M. Wozencraft and I. M. Jacobs, *Principles of Communication Engineering*. New York: Wiley, 1967.
[21] R. G. Gallager, *Information Theory and Reliable Communication*. New York: Wiley, 1968.
[22] D. Marco, "Asymptotic Quantization and Applications to Sensor Networks," Ph.D. dissertation, EECS Dept., Univ. Michigan, Ann Arbor, 2004.