Let $a \to 0$ and $T \to 0$ first, and then let $k \to \infty$ to obtain

$$\limsup_{a\to 0}\limsup_{D\to 0}\left[h(X \mid \hat{X}_D, Y, V_a) - \frac{1}{2}\log(2\pi e D)\right] \le -\frac{1}{2}\boldsymbol{E}[\log m(X,Y)]$$

which was to be proved. □

APPENDIX C

Lemma 2: Assume that $I(X;V) < \infty$ and for any $D > 0$, $Y \leftrightarrow V \leftrightarrow X \leftrightarrow Z_D$ forms a Markov chain. Suppose further that there is a measurable function $f(Y, Z_D)$ (which may depend on D) such that $f(Y, Z_D) \to X$ in probability as $D \to 0$. Then

$$\lim_{D\to 0} I(X;V \mid Y, Z_D) = 0.$$

Proof: Use the chain rule twice to obtain

$$\begin{aligned} I(X;V \mid Y, Z_D) &= I(X, Z_D; V \mid Y) - I(Z_D; V \mid Y) \\ &= I(X;V \mid Y) + I(Z_D; V \mid Y, X) - I(Z_D; V \mid Y) \\ &= I(X;V \mid Y) - I(Z_D; V \mid Y) \end{aligned} \tag{C.1}$$

where all quantities are finite since $I(X;V) < \infty$, and the third equality holds because $I(Z_D; V \mid Y, X) = 0$ by the Markov chain condition $Y \leftrightarrow V \leftrightarrow X \leftrightarrow Z_D$. Since

$$I(Z_D; V \mid Y) = I(Y, Z_D; V \mid Y) \ge I(f(Y, Z_D); V \mid Y)$$

we have

$$\liminf_{D\to 0} I(Z_D; V \mid Y) \ge \liminf_{D\to 0} I(f(Y, Z_D); V \mid Y). \tag{C.2}$$

Now the lower semicontinuity of the mutual information [17] and the condition that $f(Y, Z_D) \to X$ in probability imply that

$$\liminf_{D\to 0} I(f(Y, Z_D); V \mid Y = y) \ge I(X; V \mid Y = y) \quad \text{a.e. } [P_Y]$$

and therefore by Fatou's lemma [15] we have

$$\liminf_{D\to 0} I(f(Y, Z_D); V \mid Y) \ge I(X; V \mid Y).$$

The lemma now follows by (C.1) and (C.2). □

ACKNOWLEDGMENT

The authors wish to thank the reviewers and the Associate Editor for their comments.

REFERENCES

[1] T. Berger, *Rate Distortion Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1971.
[2] A. D. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," *IEEE Trans. Inform. Theory*, vol. IT-22, pp. 1–10, Jan. 1976.
[3] R. M. Gray, "A new class of lower bounds to information rates of stationary sources via conditional rate-distortion functions," *IEEE Trans. Inform. Theory*, vol. IT-19, pp. 480–489, July 1973.
[4] A. D. Wyner, "The rate-distortion function for source coding with side information at the decoder—II: General sources," *Inform. Contr.*, vol. 38, pp. 60–80, 1978.
[5] T. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.
[6] S. Shamai (Shitz), S. Verdú, and R. Zamir, "Systematic lossy source/channel coding," *IEEE Trans. Inform. Theory*, vol. 44, pp. 564–579, Mar. 1998.
[7] H. Viswanathan and T. Berger, "Sequential coding of correlated sources," in *Proc. IEEE Int. Symp. Information Theory*, Ulm, Germany, June 1997, p. 272.
[8] T. Linder and R. Zamir, "High-resolution source coding for nondifference distortion measures: The rate distortion function," *IEEE Trans. Inform. Theory*, vol. 45, pp. 533–547, Mar. 1999.
[9] W. R. Gardner and B. D. Rao, "Theoretical analysis of the high-rate vector quantization of LPC parameters," *IEEE Trans. Speech, Audio Processing*, vol. 3, pp. 367–381, Sept. 1995.
[10] J. Li, N. Chaddha, and R. M. Gray, "Asymptotic performance of vector quantizers with the perceptual distortion measure," in *Proc. IEEE Int. Symp. Information Theory*, Ulm, Germany, June 1997, p. 55.
[11] R. Zamir, "The rate loss in the Wyner–Ziv problem," *IEEE Trans. Inform. Theory*, vol. 42, pp. 2073–2084, Nov. 1996.
[12] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*. New York: Academic , 1981.
[13] Y. S. Chow and H. Teicher, *Probability Theory, Independence, Interchangeability, Martingales*. New York: Springer-Verlag, 1988.
[14] T. Linder, R. Zamir, and K. Zeger, "High-resolution source coding for nondifference distortion measures: Multidimensional companding," *IEEE Trans. Inform. Theory*, vol. 45, pp. 548–561, Mar. 1999.
[15] R. B. Ash, *Real Analysis and Probability*. New York: Academic, 1972.
[16] W. Rudin, *Real and Complex Analysis*, 3rd ed. New York: McGraw-Hill, 1987.
[17] I. Csiszár, "On an extremum problem of information theory," *Studia Scient. Math. Hung.*, pp. 57–70, 1974.

Optimal Entropy-Constrained Scalar Quantization of a Uniform Source

András György and Tamás Linder, *Senior Member, IEEE*

***Abstract*—Optimal scalar quantization subject to an entropy constraint is studied for a wide class of difference distortion measures including rth-power distortions with $r > 0$. It is proved that if the source is uniformly distributed over an interval, then for any entropy constraint R (in nats), an optimal quantizer has $N = \lceil e^R \rceil$ interval cells such that $N - 1$ cells have equal length d and one cell has length $c \le d$. The cell lengths are uniquely determined by the requirement that the entropy constraint is satisfied with equality. Based on this result, a parametric representation of the minimum achievable distortion $D_h(R)$ as a function of the entropy constraint R is obtained for a uniform source. The $D_h(R)$ curve turns out to be nonconvex in general. Moreover, for the squared-error distortion it is shown that $D_h(R)$ is a piecewise-concave function, and that a scalar quantizer achieving the lower convex hull of $D_h(R)$ exists only at rates $R = \log N$, where N is a positive integer.**

***Index Terms*—Constrained optimization, difference distortion measures, entropy coding, scalar quantization, uniform source.**

I. INTRODUCTION

Scalar (or zero-memory) quantization is the simplest method for the lossy coding of an information source with real-valued outputs. A scalar quantizer followed by variable-length lossless coding (entropy coding) can perform remarkably well, which makes this method popular in applications where implementation complexity is a decisive factor.

Manuscript received March 26, 2000; revised June 9, 2000. This work was supported in part by the Natural Sciences and Engineering Research Council (NSERC) of Canada. The work of A. György was also supported by the Soros Foundation.

A. György is with the Department of Mathematics and Statistics, Queen's University, Kingston, ON, K7L 3N6 Canada (e-mail: gyorgy@mast.queensu.ca). He is also with the Department of Computer Science and Information Theory, Technical University of Budapest, Budapest, Hungary.

T. Lindar is with the Department of Mathematics and Statistics, Queen's University, Kingston, ON, K7L 3N6 Canada (e-mail:linder@mast.queensu.ca).

Communicated by P. A. Chou, Associate Editor for Source Coding.

Publisher Item Identifier S 0018-9448(00)09674-7.

The two main quantities characterizing a scalar quantizer Q are its distortion and rate. The distortion $D(Q)$ is the average distortion between the source and the quantizer output. If Q is followed by entropy coding, the rate is usually defined as the entropy $H(Q)$ of the output of Q. (For a stationary and memoryless source, $H(Q)$ is indeed the smallest rate asymptotically achievable by variable-length lossless coding of blocks of quantizer outputs.) One would like to make both $H(Q)$ and $D(Q)$ as small as possible, but these quantities are inversely related. A natural design problem is then to minimize $D(Q)$ subject to an entropy constraint $H(Q) \leq R$. Let $D_h(R)$ denote the lowest possible distortion of any scalar quantizer with output entropy not greater than R. A quantizer achieving this minimum is called an ECSQ. It is of interest to determine $D_h(R)$ either analytically or numerically, as well as to find the optimal ECSQ achieving the minimum distortion.

It appears that very few concrete examples for an optimal ECSQ are known in analytic form. In general, efforts have focused on finding necessary conditions for the optimality of an ECSQ with a *fixed* number of output points n [1]–[3]. These conditions give rise to practical algorithms for designing an ECSQ with a fixed number of output points [1], [4], [2], [3], [5]. To determine the overall optimal ECSQ and the corresponding optimal performance curve $D_h(R)$, one must find the optimum performance over all n. Unfortunately, this step is rather hard, even for the most common continuous source distributions. A notable exception is the case of an exponentially distributed source and mean-squared distortion considered by Berger [1]. He derived an analytic expression for $D_h(R)$ based on the observation that for the exponential distribution, the necessary conditions for optimality at any positive rate are satisfied by an infinite-level uniform quantizer. To our knowledge, this is the only case where a correct[1] analytic formula for $D_h(R)$ is known.

In this correspondence we determine analytically the optimal ECSQ for a source which is uniformly distributed over a finite interval. We allow a rather wide class of difference distortion measures including rth-power distortions $d(x,y) = |x-y|^r$ with $r > 0$, and distortion measures of the form $d(x,y) = \rho(|x-y|)$, where ρ is a nonnegative, strictly increasing, and convex function. Our main result proves that an optimal ECSQ for any rate $R \geq 0$ (measured in nats) is an $N = \lceil e^R \rceil$-level quantizer (here $\lceil x \rceil$ denotes the smallest integer not less than x). This quantizer has $N-1$ cells of equal length d and one cell of length $c \leq d$, where d and c are uniquely determined by the requirement that the entropy constraint is satisfied with equality (the optimal quantizer is uniform if $e^R = N$, as expected). Specialized to the squared-error distortion, our result rigorously proves that the ECSQs found for the uniform distribution by Farvardin and Modestino [2] (using a numerical approach) are indeed optimal. In case of the absolute-error distortion our result also agrees with a result independently obtained by Topsoe [14] in a prediction context. Based on the analytic description of an optimal finite-level ECSQ, we then obtain a parametric expression for the $D_h(R)$ curve and investigate its properties. In general, $D_h(R)$ is piecewise-smooth (differentiable everywhere except at the points $R = \log N$). For the squared-error distortion (and more generally for rth-power distortions with $0 < r \leq 3$) we prove that $D_h(R)$ is concave in each interval $[\log(N-1), \log N]$, where $N \geq 2$ is an integer. Thus for such distortion measures, $D_h(R)$ is not convex over any interval. It also follows that in this case an optimal ECSQ achieving the lower convex hull of $D_h(R)$ exists only at rates $R = \log N$, where N is a positive integer.

The question whether $D_h(R)$ for a given source is convex is of particular interest because of the special role of the lower convex hull of $D_h(R)$ in variable-length lossy coding. For example, the lower convex hull of $D_h(R)$ is the minimum achievable distortion in causal lossy coding of a memoryless source [6]. Also, Lloyd–Max type necessary conditions of optimality are known only for an optimal ECSQ which achieves the lower convex hull of $D_h(R)$ [7]. Now for a discrete source, $D_h(R)$ is never convex since it is decreasing and piecewise-constant. On the other hand, it can be shown (using the analytical expression of Berger [1]) that for an exponentially distributed source and the squared-error distortion, $D_h(R)$ is convex. It has also been conjectured [6] that $D_h(R)$ is convex for a wide variety of source distributions and distortion measures. Our results for the uniform source demonstrate that $D_h(R)$ can be nonconvex even for "nice" continuous source distributions.

[1]Although a complete proof that infinite-level uniform quantizers are indeed optimal is missing, the result is widely believed to be correct.

II. Preliminaries

An N-level scalar quantizer Q is a measurable mapping of the real line $\mathbb{R}$ into a finite or countably infinite set of distinct reals $\{y_1, \ldots, y_N\}$ called the *codebook* of Q. (In case the codebook is not finite, we formally define $N = \infty$ and call Q an infinite-level quantizer.) The *codepoints* y_i and the associated *quantization cells* $S_i = \{x : Q(x) = y_i\}, i = 1, \ldots, N$ completely characterize Q since the S_i form a partition of $\mathbb{R}$ and

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

The distortion of Q in quantizing a real random variable X with distribution μ_X is measured by the expectation

$$D(Q) = E[d(X, Q(X))] = \int_{-\infty}^{\infty} d(x, Q(x)) \mu_X\,(dx)$$

where the *distortion measure* $d(\cdot,\cdot)$ is a nonnegative measurable function of two real variables. The entropy-constrained *rate* of Q is the entropy of the discrete random variable $Q(X)$

$$H(Q) = -\sum_{i=1}^{N} P[X \in S_i] \log P[X \in S_i]$$

where $\log$ denotes the natural logarithm ($H(Q)$ is measured in nats). A scalar quantizer whose rate is measured by $H(Q)$ is called an *entropy-constrained scalar quantizer* (ECSQ).

For any $R \geq 0$ let $D_h(R)$ denote the lowest possible distortion of any quantizer with output entropy not greater than R. This function is formally defined by

$$D_h(R) = \inf\{D(Q) : H(Q) \leq R\}$$

where the infimum is taken over all finite- or infinite-level scalar quantizers whose entropy is less than or equal to R. Any Q that achieves $D_h(R)$ in the sense that $H(Q) \leq R$ and $D(Q) = D_h(R)$ is called an *optimal* ECSQ.

III. Optimal ECSQ for a Uniform Source

A scalar quantizer is called *regular* if each cell S_i is an interval and each code point y_i lies inside S_i. Assume that the distortion measure is of the form

$$d(x,y) = \rho(|x-y|) \tag{1}$$

where $\rho : [0,\infty) \to [0,\infty)$ is a strictly increasing function. For such distortion measures, nearest neighbor quantizers (i.e., quantizers which satisfy $d(x, Q(x)) = \min_{1 \leq i \leq N} d(x, y_i)$ for all x) are regular, and thus an optimal fixed-rate N-level quantizer (i.e., a quantizer which has minimum distortion among all N-level quantizers) can be assumed to be regular.

Unfortunately, an optimal ECSQ is not necessarily a nearest neighbor quantizer, and thus in general it is incorrect to restrict attention to regular quantizers (or quantizers with interval cells) when searching for an optimal ECSQ. Indeed, it is not hard to construct a discrete source with three real-valued outputs for which the unique optimal ECSQ is not regular at certain rates. We note here that a nearest neighbor type condition does hold for an optimal ECSQ which achieves the lower convex hull of $D_h(R)$, implying that such a quantizer can be assumed to be regular [7]. However, as Corollary 2 later shows, an ECSQ achieving the lower convex hull of $D_h(R)$ may not exist for most rate constraints. More recently, it has been shown [8] for continuous source distributions and distortion measures in the form $d(x,y) = \rho(|x-y|)$, where ρ is an increasing convex function, that if an optimal finite-level ECSQ exists for a given rate constraint, then there is an optimal ECSQ for the same rate constraint which is regular.

The question of ECSQ regularity is much simpler if the source is uniformly distributed. Let X have a uniform distribution on the unit interval $(0,1)$ and assume that the distortion measure is in the form of (1). Let Q be any finite- or infinite-level quantizer with cells $\{S_1, \ldots, S_N\}$ and code points $\{y_1, \ldots, y_N\}$, and define $p_i = \lambda(S_i \cap (0,1))$ for $i = 1, \ldots, N$, where λ denotes the Lebesgue measure. Then we can define a new quantizer $\hat{Q}$ over $(0,1)$ which has N interval cells of length $p_i, i = 1, \ldots, N$, and N code points which are located at the midpoints of these cells (the definition of $\hat{Q}$ outside $(0,1)$ is immaterial). The distortion of $\hat{Q}$ is

$$D(\hat{Q}) = \sum_{i=1}^{N} \Phi(p_i) \tag{2}$$

where $\Phi(p)$ is defined for all $p \geq 0$ by

$$\Phi(p) = 2\int_0^{p/2} \rho(x)\,dx. \tag{3}$$

Since ρ is increasing, it is easy to see that for all $i = 1, \ldots, N$

$$\int_{S_i \cap (0,1)} \rho(|x-y_i|)\,dx \geq \int_{-p_i/2}^{p_i/2} \rho(|x|)\,dx$$

and so $D(\hat{Q}) \leq D(Q)$. On the other hand,

$$H(\hat{Q}) = -\sum_{i=1}^{N} p_i \log p_i \tag{4}$$

so that $H(\hat{Q}) = H(Q)$. Consequently, when searching for an optimal ECSQ for the uniform distribution over $(0,1)$, it suffices to consider interval partitions of $(0,1)$ and the associated regular quantizers with code points at the midpoints of the intervals. All quantizers in the remainder of this correspondence will be assumed to be of this type. The distortion and rate of any such quantizer are uniquely determined by the cell lengths $\{p_i;\ i = 1, \ldots, N\}$ through (2) and (4). Note that if N is finite and $p_i = 1/N, i = 1, \ldots, N$, the resulting quantizer is the N-level uniform quantizer over $(0,1)$.

In what follows we will consider distortion measures in the form $d(x,y) = \rho(|x-y|)$, where $\rho : [0,\infty) \to [0,\infty)$ is strictly increasing, continuous, and $\rho(e^t), t \in \mathbb{R}$ is strictly convex. Examples for such distortion measures include rth-power distortions $d(x,y) = |x-y|^r$ with $r > 0$ (in this case, $\rho(e^t) = e^{rt}$), and distortion measures $d(x,y) = \rho(|x-y|)$, where ρ is strictly increasing and convex on $[0,\infty)$. Another distortion measure which does not fall into either of these categories but satisfies the requirements is $d(x,y) = \log(1 + |x-y|)$.

If $R = 0$, the optimal ECSQ for any source distribution has only one code point. The next result shows that if the source has a uniform distribution, then for any rate $R > 0$ there exists an optimal finite-level ECSQ with a very simple structure.

Theorem 1: Let the source X have uniform distribution over $(0,1)$ and assume that $d(x,y) = \rho(|x-y|)$, where $\rho : [0,\infty) \to [0,\infty)$ is a strictly increasing continuous function such that $\rho(e^t)$ is strictly convex. Then Q is an optimal ECSQ for a rate constraint $R > 0$ if and only if Q has $N = \lceil e^R \rceil$ cells; one cell of length c and $N-1$ cells of length $(1-c)/(N-1)$, where c is the unique solution of the equation

$$-c\log c - (1-c)\log\left(\frac{1-c}{N-1}\right) = R$$

in the interval $(0, 1/N)$.

Theorem 1 implies the intuitive result that if $R = \log N$, then the unique optimal ECSQ for the uniform source is the N-level uniform quantizer. If $R < \log N$, then $c < (1-c)/(N-1)$, and the optimal ECSQ is no longer unique; there are exactly N such quantizers. Farvardin and Modestino [2] reached the same conclusion for squared-error distortion using numerical methods.

Theorem 1 remains valid (after rescaling) if X is uniformly distributed over an arbitrary interval (a,b). To see this, let

$$\widehat{X} = \frac{X-a}{b-a}$$

and

$$\widehat{\rho}(x) = \rho\left(\frac{x}{b-a}\right).$$

Then $\widehat{X}$ and $\widehat{\rho}$ satisfy the conditions of the theorem, and a quantizer Q is optimal for X and ρ if and only if

$$\hat{Q}(x) = \frac{Q(x(b-a)+a)-a}{b-a}$$

is optimal for $\widehat{X}$ and $\widehat{\rho}$.

The proof of the theorem, given in the next section, has two main parts. First, similarly to [1] and [2], the usual Kuhn–Tucker conditions of constrained optimization are used to identify necessary conditions for the optimality of an n-level ECSQ for a fixed positive integer n. After eliminating all quantizers not satisfying these conditions, we are left with the family of n-level quantizers over $(0,1)$ which satisfy the entropy constraint with equality and whose cell lengths can take only two distinct values (these quantizers were also identified in [2]). The second, harder part of the proof consists of identifying, for a fixed n, the quantizers which have minimal distortion in this family, and then finding the optimal choice of n.

Using (2) and (4), the distortion and entropy of an N-level quantizer Q with cell lengths $p_1 = c$ and $p_i = (1-c)/(N-1), i = 2, \ldots, N$ are given, respectively, by

$$D(Q) = \Phi(c) + (N-1)\Phi\left(\frac{1-c}{N-1}\right).$$

and

$$H(Q) = -c\log c - (1-c)\log\left(\frac{1-c}{N-1}\right).$$

It is easy to see that $H(Q)$ is a strictly increasing function of $c \in [0, 1/N]$. Also, $H(Q) = \log(N-1)$ if $c = 0, H(Q) = \log N$ if $c = 1/N$, and the corresponding quantizers are the $(N-1)$-level and N-level uniform quantizers, respectively. Thus Theorem 1 yields the following parametric description of $D_h(R)$ for the uniform distribution.

Corollary 1: For X and $d(x,y)$ as in Theorem 1, and for any positive integer $N \geq 2$, the parametric representation of $D_h(R)$ in the interval $[\log(N-1), \log N]$ is given by

$$R_c = -c\log c - (1-c)\log\left(\frac{1-c}{N-1}\right)$$

$$D_h(R_c) = \Phi(c) + (N-1)\Phi\left(\frac{1-c}{N-1}\right)$$

where $c \in [0, 1/N]$.

From the parametric representation we can deduce some important properties of $D_h(R)$. For example, it immediately follows that $D_h(R)$

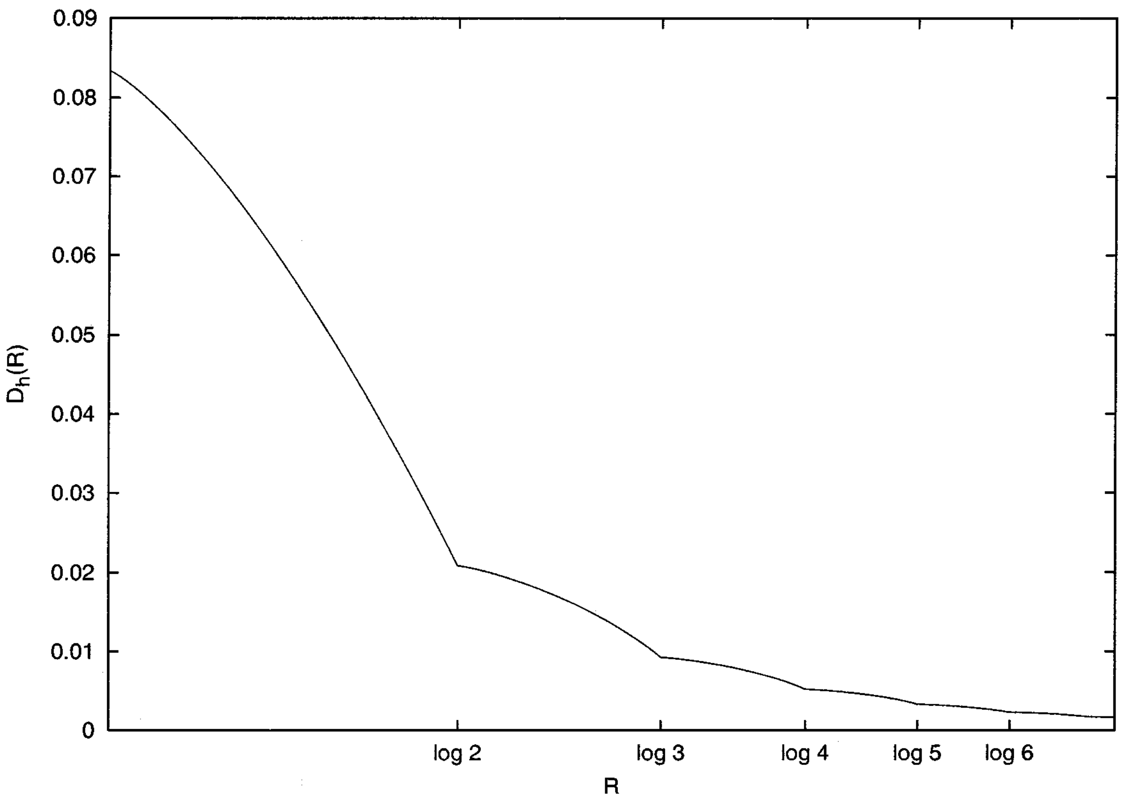

Fig. 1. $D_h(R)$ for the uniform source and squared error distortion.

is everywhere continuous. Moreover, plotting $D_h(R)$ for the squared error distortion $d(x,y)=(x-y)^2$ (see Fig. 1) suggests that the graph of $D_h(R)$ consists of smooth, concave pieces joined in a nonsmooth manner at rates $R=\log N$. The next result proves these properties of $D_h(R)$ under more general conditions.

Corollary 2: With the conditions of Theorem 1, $D_h(R)$ has the following properties.

i) $D_h(R)$ is continuously differentiable on each open interval $(\log(N-1), \log N)$, where $N \geq 2$ is a positive integer. At $R=\log N$, the right derivative of $D_h(R)$ is zero for all $N \geq 1$, and the left derivative of $D_h(R)$ is negative for all $N \geq 2$. Thus $D_h(R)$ is not differentiable at the points $R=\log N$ for $N \geq 2$.

ii) Let $d(x,y)=|x-y|^r$ be the rth-power distortion with $0<r\leq 3$. Then $D_h(R)$ is strictly concave on each interval $[\log(N-1), \log N]$ for $N \geq 2$.

The proof of the corollary is given in the next section. The proof also shows that part ii) cannot be improved in the sense that if $d(x,y)=|x-y|^r$ with $r>3$, then $D_h(R)$ is no longer concave on $[0, \log 2]$.

Part i) of the preceding corollary implies that $D_h(R)$ is not convex. Moreover, part ii) shows that for the squared-error distortion an ECSQ achieving the lower convex hull of $D_h(R)$ exists only at the discrete rate values $R=\log N$. This fact suggests that an ECSQ which achieves the lower convex hull of $D_h(R)$ is the exception rather than the rule.

IV. Proofs

Proof of Theorem 1: Without loss of generality we will assume that $\rho(0)=0$ (otherwise, we can replace $\rho(x)$ by $\widehat{\rho}(x)=\rho(x)-\rho(0)$). Let Ψ be the Gish–Pierce function [9], [10] defined by

$$\Psi(p)=\begin{cases}\dfrac{\Phi(p)}{p}, & \text{if } p>0\\ 0, & \text{if } p=0\end{cases}$$

where $\Phi(p)=2\int_0^{p/2}\rho(x)\,dx$. Notice that $\Psi(p)=E[\rho(pY)]$ for all $p\geq 0$, where Y is a random variable that is uniformly distributed over the interval $(0,1/2)$. Then the strict convexity of $\rho(e^t)$ implies that for all $t_1, t_2 \in \mathbb{R}$ such that $t_1 \neq t_2$, and $0<\alpha<1$

$$\begin{aligned}\Psi\left(e^{\alpha t_1+(1-\alpha)t_2}\right) &= E\left[\rho\left(e^{\alpha t_1+(1-\alpha)t_2}Y\right)\right]\\ &= E\left[\rho\left(e^{\alpha(t_1+\log Y)+(1-\alpha)(t_2+\log Y)}\right)\right]\\ &< E\left[\alpha\rho\left(e^{t_1+\log Y}\right)+(1-\alpha)\rho\left(e^{t_2+\log Y}\right)\right]\\ &= \alpha\Psi(e^{t_1})+(1-\alpha)\Psi(e^{t_2}).\end{aligned}$$

Thus $\Psi(e^t)$ is strictly convex. Since $\Psi(0)=0$ and $\Psi(p)>0$ for $p>0$, the convexity of $\Psi(e^t)$ implies that $\Psi(p)$ is strictly increasing.

By the discussion preceding Theorem 1, we need to find nonnegative cell lengths $\{p_i;\ i=1,2,\ldots\}$ satisfying $\sum_i p_i=1$, which minimize $\sum_i \Phi(p_i)$ subject to $-\sum_i p_i\log p_i \leq R$. For all $\{p_i\}$ satisfying this entropy constraint we have

$$\begin{aligned}\sum_i \Phi(p_i) &= \sum_{i:\,p_i>0} p_i\Psi(e^{\log p_i})\\ &\geq \Psi\left(e^{\sum_i p_i\log p_i}\right) \qquad (5)\\ &\geq \Psi(e^{-R}) \qquad (6)\end{aligned}$$

where (5) follows from Jensen's inequality and the convexity of $\Psi(e^t)$, and (6) follows since Ψ is increasing. Thus $\Psi(e^{-R})$ is a lower bound on the distortion of any quantizer with entropy not greater than R, that is, $D_h(R) \geq \Psi(e^{-R})$. Since $\Psi(e^t)$ is strictly convex and strictly increasing, $\sum_i \Phi(p_i) = \Psi(e^{-R})$ if and only if all positive p_i's are equal and $-\sum_i p_i \log p_i = R$. Equivalently, $R = \log N$ for some positive integer N and $p_i = 1/N$ for (say) $1 \leq i \leq N$, and $p_i = 0$ for $i > N$. The resulting quantizer is the N-level uniform quantizer over $(0,1)$ with entropy $R = \log N$ and distortion $\Psi(1/N) = D_h(\log N)$. This proves the theorem for rates $R = \log N$, where N is a positive integer.

Now consider the case $\log(N-1) < R < \log N$, where $N = \lceil e^R \rceil$. First observe that in the infimum defining $D_h(R)$ it is enough to consider finite-level quantizers; i.e.,

$$D_h(R) = \inf\{D(Q) : Q \text{ is finite-level, } H(Q) \leq R\}. \tag{7}$$

(Let Q be any infinite-level quantizer over $(0,1)$ with cell lengths $\{p_i;\, i = 1, 2, \ldots\}$ and for a positive integer n, let $\hat{Q}$ have cell lengths $\{p_1, \ldots, p_{n-1}, \sum_{i \geq n} p_i\}$. Then $H(\hat{Q}) \leq H(Q)$ and

$$D(\hat{Q}) - D(Q) \leq \Phi\left(\sum_{i \geq n} p_i\right)$$

and now (7) follows since $\sum_{i \geq n} p_i \to 0$ as $n \to \infty$.)

For a positive integer n, let $\mathcal{P}_n$ denote the n-dimensional probability simplex

$$\mathcal{P}_n = \left\{(p_1, \ldots, p_n) : p_i \geq 0, i = 1, \ldots, n; \sum_{i=1}^n p_i = 1\right\}$$

and for all $\boldsymbol{p} = (p_1, \ldots, p_n) \in \mathcal{P}_n$ define

$$h_n(\boldsymbol{p}) = -\sum_{i=1}^n p_i \log p_i,$$
$$\delta_n(\boldsymbol{p}) = \sum_{i=1}^n \Phi(p_i).$$

Let $R > 0$ be fixed. Since h_n and δ_n are continuous and $\mathcal{P}_n$ is compact

$$B_{n,R} = \{\boldsymbol{p} \in \mathcal{P}_n : h_n(\boldsymbol{p}) \leq R\}$$

is compact and δ_n achieves its minimum in $B_{n,R}$. Then (7) implies that $D_h(R)$ is given by

$$D_h(R) = \inf_{n \geq 1} \min\{\delta_n(\boldsymbol{p}) : \boldsymbol{p} \in B_{n,R}\}. \tag{8}$$

Note that $\min\{\delta_n(\boldsymbol{p}) : \boldsymbol{p} \in B_{n,R}\}$ is nonincreasing in n and an optimal entropy-constrained quantizer with a finite codebook exists if and only if the infimum is achieved in (8) for finite n. If $\boldsymbol{p}^* \in \mathcal{P}_n$ minimizes δ_n over $B_{n,R}$ and it has k nonzero components, then by dropping the zero components we obtain

$$\boldsymbol{p}^{**} \in \mathcal{P}_k^+ = \{\boldsymbol{p} \in \mathcal{P}_k : p_i > 0, i = 1, \ldots, k\}$$

which minimizes δ_k over $B_{k,R}$. Since the quantizers associated with $\boldsymbol{p}^*$ and $\boldsymbol{p}^{**}$ are identical, we can conclude that it suffices to find the positive solutions $\boldsymbol{p} \in \mathcal{P}_n^+$ of the constrained minimization problem

$$\begin{aligned} &\text{minimize} && \delta_n(\boldsymbol{p}) \\ &\text{subject to} && h_n(\boldsymbol{p}) \leq R, \quad \sum_{i=1}^n p_i = 1 \end{aligned} \tag{9}$$

for all $n \geq 1$ such that a solution exists. Since h_n and δ_n are continuously differentiable on $B_{n,R} \cap \mathcal{P}_n^+$, we can use the Kuhn–Tucker conditions (see, e.g., [11]) to find all local minimum points in (9) for all n. The collection of these solutions will correspond to a simple family of finite-level quantizers where it will be possible to identify the global optimum.

By the Kuhn–Tucker conditions, if a local minimum $\boldsymbol{p}$ is a regular point of the constraints in (9) in the sense that the gradients of $h_n(\boldsymbol{p})$ and $\sum_{i=1}^n p_i$ are linearly independent, then there exist Lagrange multipliers $\mu \geq 0$ and $\lambda \in \mathbb{R}$ such that

$$J(\boldsymbol{p}, \mu, \lambda) = \delta_n(\boldsymbol{p}) + \mu h_n(\boldsymbol{p}) + \lambda \sum_{i=1}^n p_i$$

satisfies $\partial J / \partial p_i = 0$ for $i = 1, \ldots, n$. The gradients of $h_n(\boldsymbol{p})$ and $\sum_{i=1}^n p_i$ are linearly dependent if and only if $p_i = 1/n$ for $i = 1, \ldots, n$. Otherwise, the Kuhn–Tucker conditions give

$$\frac{\partial J(\boldsymbol{p}, \mu, \lambda)}{\partial p_i} = \rho(p_i/2) - \mu(1 + \log p_i) + \lambda = 0, \qquad i = 1, \ldots, n$$

that is, a minimizing $\boldsymbol{p}$ must solve

$$\lambda = -\rho(p_i/2) + \mu(1 + \log p_i), \qquad i = 1, \ldots, n. \tag{10}$$

Let λ and μ be fixed and for $p > 0$ define $v(p) = -\rho(p/2) + \mu(1 + \log p)$. Then

$$v(2e^t) = -\rho(e^t) + \mu t + \mu(1 + \log 2)$$

is a strictly concave function of t (since $\rho(e^t)$ is strictly convex). Thus the equation $v(2e^t) = \lambda$ has at most two distinct solutions in t, and, consequently, (10) has at most two distinct solutions in p_i.

Hence the components of any optimal $\boldsymbol{p} \in \mathcal{P}_n^+$ take at most two distinct positive values. To uniquely describe such a $\boldsymbol{p}$ (up to permutations of the components), let c denote the smaller of the two values, let l denote the number of components equal to c, and let us specify that $c \in (0, 1/n)$ and $0 \leq l < n$. Then there are $n - l$ components equal to $d = (1 - cl)/(n - l)$ (note that $l = 0$ if and only if all components are equal). An associated quantizer has l cells of length c and $n - l$ cells of length d, and its distortion and entropy are given, respectively, by

$$\delta(c, l, n) = l\Phi(c) + (n - l)\Phi(d)$$

and

$$h(c, l, n) = -lc \log c - (1 - lc) \log d.$$

Therefore, if there exist c, l, and n minimizing $\delta(c, l, n)$ subject to $h(c, l, n) \leq R$, the corresponding quantizer is optimal (i.e, it achieves $D_h(R)$).

Fix $R > 0$ and assume that $\log n < R$. Since the uniform n-level quantizer has minimum distortion (namely, $n\Phi(1/n)$) and maximum entropy (namely, $\log n$) among all n-level quantizers, it is the optimal n-level quantizer with entropy constraint R. Since $\log n < R$, one can easily construct an $(n+1)$-level quantizer Q with n equal-length cells and one sufficiently small cell which has entropy $H(Q) < R$ and distortion $D(Q) < n\Phi(1/n)$. Hence no n-level quantizer can achieve $D_h(R)$ if $\log n < R$. Therefore, we will assume that $\log n > R$ (the case $\log n = R$ was dealt with previously). We will also assume that $l > 0$, since $l = 0$ results in the n-level uniform quantizer with rate $\log n > R$.

Note that

$$\frac{\partial h(c, l, n)}{\partial c} = l(\log d - \log c) > 0 \tag{11}$$

since $c < d = (1 - cl)/(n - l)$, and thus $h(c, l, n)$ is a strictly increasing function of $c \in (0, 1/n)$ for fixed l. Therefore, the constraint $h(c, l, n) \leq R$ can be satisfied with $c > 0$ if and only if $l \geq l_{\min}$, where the integer $l_{\min}$ is defined by

$$l_{\min} = \min\{l \geq 0 : \log(n - l) < R\}.$$

Note that $l_{\min} \geq 1$ since $\log n > R$ by assumption. Next we observe that $\delta(c,l,n)$ is strictly decreasing in c since

$$\frac{\partial \delta(c,l,n)}{\partial c} = l\left(\rho\left(\frac{c}{2}\right) - \rho\left(\frac{d}{2}\right)\right) < 0$$

(recall that ρ is strictly increasing). Thus for fixed n (such that $\log n > R$) and $l \geq l_{\min}$, the unique c minimizing $\delta(c,l,n)$ subject to $h(c,l,n) \leq R$ is the unique solution of the equation

$$-lc\log c - (1-lc)\log d = R$$

in the interval $(0, 1/n)$. Let $c(l,n,R)$ denote this solution, and denote the corresponding distortion by

$$\delta^*(l,n,R) = \delta(c(l,n,R),l,n). \qquad (12)$$

Lemma 1 in the Appendix shows that $\delta^*(l,n,R)$ is strictly increasing in l for fixed n and R, and therefore $\delta^*(l,n,R)$ is uniquely minimized in l by the choice $l = l_{\min}$, the smallest possible value of l.

Now if $\log(N-1) < R < \log N$, then $\log n > R$ if and only if $n \geq N$. For any such n, we have $l_{\min} = n - N + 1$, and now $\delta^*(l,n,R)$ is minimized in l by $l = n - N + 1$. The corresponding distortion is

$$\delta^*(n-N+1,n,R) = (n-N+1)\Phi(c) + (N-1)\Phi(d) \qquad (13)$$

where c is the unique solution of the equation

$$-(n-N+1)c\log c - (1-(n-N+1)c)\log d = R \qquad (14)$$

in $(0,1/n)$ where $d = (1-(n-N+1)c)/(N-1)$. Lemma 2 in the Appendix shows that $\delta^*(n-N+1,n,R)$ is strictly increasing in n for fixed N and R, and thus it is minimized by $n = N$, the smallest possible choice for n. We can conclude that any quantizer over $(0,1)$ with one cell of length c and $N-1$ cells of length $d = (1-c)/(N-1)$, where c satisfies

$$-c\log c - (1-c)\log\left(\frac{1-c}{N-1}\right) = R$$

is optimal; i.e., it achieves $D_h(R)$. It also follows that any other quantizer with a different set of cell lengths is strictly suboptimal. $\square$

Proof of Corollary 2:

i) Recall that by Corollary 1, in the interval $[\log(N-1), \log N]$ the parametric equations of $D_h(R)$ are

$$R_c = -c\log c - (1-c)\log d$$
$$D_h(R_c) = \Phi(c) + (N-1)\Phi(d)$$

where $d = (1-c)/(N-1)$ and $c \in [0, 1/N]$. Note that $c = 0$ corresponds to $R = \log(N-1)$ and $c = 1/N$ corresponds to $R = \log N$. By implicit differentiation, $D_h(R)$ is continuously differentiable in $(\log(N-1), \log N)$ for all $N \geq 2$, and its derivative is

$$D_h'(R_c) = \frac{\partial D_h(R_c)}{\partial c}\left(\frac{\partial R_c}{\partial c}\right)^{-1} = \frac{\rho(c/2) - \rho(d/2)}{\log(d/c)}. \qquad (15)$$

Assume that $N \geq 2$. By L'Hospital's rule, the left derivative at $R = \log N$ is given by the limit

$$\begin{aligned} D_{h^-}'(\log N) &= \lim_{R \to (\log N)^-} D_h'(R) \\ &= \lim_{c \to 1/N} \frac{\rho(c/2) - \rho(d/2)}{\log(d/c)} \\ &= \lim_{c \to 1/N} \frac{\dfrac{\rho(c/2) - \rho(d/2)}{d/2 - c/2}}{\dfrac{\log(d/2) - \log(c/2)}{d/2 - c/2}}. \end{aligned} \qquad (16)$$

The denominator in (16) converges to $2N$. On the other hand, we have

$$\begin{aligned} \frac{\rho(c/2) - \rho(d/2)}{d/2 - c/2} = &\frac{\rho\left(\frac{c}{2}\right) - \rho\left(\frac{1}{2N}\right)}{\frac{c}{2} - \frac{1}{2N}} \frac{\frac{c}{2} - \frac{1}{2N}}{\frac{d}{2} - \frac{c}{2}} \\ &- \frac{\rho\left(\frac{d}{2}\right) - \rho\left(\frac{1}{2N}\right)}{\frac{d}{2} - \frac{1}{2N}} \frac{\frac{d}{2} - \frac{1}{2N}}{\frac{d}{2} - \frac{c}{2}}. \end{aligned} \qquad (17)$$

The convexity of $u(t) = \rho(e^t)$ implies that its left and right derivatives $u_-'(t)$ and $u_+'(t)$ exist for all t, which readily implies the existence of the left and right derivatives $\rho_-'(x)$ and $\rho_+'(x)$ at every $x > 0$. In fact, $u_-'(t) = e^t\rho_-'(e^t)$ and $u_+'(t) = e^t\rho_+'(e^t)$, and therefore $\rho_-'(x)$ and $\rho_+'(x)$ are positive for all $x > 0$. Thus as $c \to \frac{1}{N}$, the first term on the right side of (17) converges to $-\frac{N-1}{N}\rho_-'(\frac{1}{2N})$, and the second term converges to $-\frac{1}{N}\rho_+'(\frac{1}{2N})$. Therefore, by (16), the left derivative of $D_h(R)$ at $R = \log N$ is

$$D_{h^-}'(\log N) = -\frac{1}{2N^2}\rho_+'\left(\frac{1}{2N}\right) - \frac{N-1}{2N^2}\rho_-'\left(\frac{1}{2N}\right) < 0.$$

Now let $N \geq 1$. To determine the right derivative of $D_h(R)$ at $R = \log N$, replace N by $N+1$ in the range of the parameter c (so now we have $c \in (0, 1/(N+1))$ and $d = (1-c)/N$). We obtain

$$D_{h^+}'(\log N) = \lim_{R \to (\log N)^+} D_h'(R) = \lim_{c \to 0} D_h'(R_c) = 0$$

since $d \to 1/N$ as $c \to 0$.

ii) We have $d(x,y) = |x-y|^r$ with $r > 0$, so that Theorem 1 and Corollary 1 apply. Then by (15), the derivative of $D_h(R)$ in $(\log(N-1), \log N)$ is parametrically given by

$$D_h'(R_c) = \frac{(c/2)^r - (d/2)^r}{\log(d/c)}$$

where $c \in (0, 1/N)$. It follows that $D_h(R)$ is twice differentiable in $(\log(N-1), \log N)$ and its second derivative is as shown at the bottom of this page. Next, we show that $D_h''(R) < 0$ in $(\log(N-1), \log N)$

$$\begin{aligned} D_h''(R_c) &= \frac{\partial D_h'(R_c)}{\partial c}\left(\frac{\partial R_c}{\partial c}\right)^{-1} \\ &= \frac{rc[(1-c)(c/2)^{r-1} + d(d/2)^{r-1}]\log(d/c) + 2[(c/2)^r - (d/2)^r]}{2c(1-c)[\log(d/c)]^3}. \end{aligned}$$

for all $N \geq 2$ if $0 < r \leq 3$. By continuity, this implies that $D_h(R)$ is strictly concave on $[\log(N-1), \log N]$ for all $N \geq 2$.

Since $c < d, D_h''(R_c) < 0$ if and only if the numerator of the above quotient is negative. Letting $x = d/c$, after some algebra we obtain that $D_h''(R_c) < 0$ holds for all $c \in (0, 1/N)$ if and only if

$$(x^r - 1)((N-1)x + 1) - r(x^r + (N-1)x)\log x > 0 \tag{18}$$

for all $x > 1$. Now observe that the left side of the preceding inequality is a linear function of $N-1$ such that the coefficient of $N-1$ is $x(x^r - 1 - \log x^r)$. Since this coefficient is positive for all $x > 1$ by the inequality $t - 1 > \log t, (t \neq 1)$, it is enough to prove (18) for $N = 2$. Equivalently, we will show that

$$f_r(x) = \frac{(x^r - 1)(x+1)}{x + x^r} - r \log x > 0$$

for all $x > 1$. Setting $r = 3$ we have

$$f_3'(x) = \frac{(x-1)^4(x^2 + x + 1)}{(x + x^3)^2}$$

and therefore $f_3'(x) > 0$ for all $x > 1$. Since $f_3(1) = 0$, we obtain $f_3(x) > 0$ for all $x > 1$.

On the other hand,

$$\frac{\partial f_r(x)}{\partial r} = \frac{(x^r - 1)(x^2 - x^r)\log x}{(x + x^r)^2}.$$

Hence, for all $x > 1, \partial f_r(x)/\partial r > 0$ if $r < 2$, and $\partial f_r(x)/\partial r < 0$ if $r > 2$. Since $f_0(x) = 0$ and $f_3(x) > 0$ for all $x > 1$, this implies that $f_r(x) > 0$ for all $x > 1$ and $0 < r \leq 3$, proving the claim ii).

Lemma 3 in the Appendix shows that if $r > 3$, then there exists an $x_r > 1$ such that $f_r(x) < 0$ for all $x \in (1, x_r)$. Thus $D_h(R)$ is no longer concave on $[0, \log 2]$ for $r > 3$. □

Appendix

Lemma 1: If $R > 0$ and $\log n > R$, the function $\delta^*(l, n, R)$ defined in (12) is a strictly increasing function of l for $l_{\min} \leq l \leq n-1$.

Proof: Although $\delta^*(l, n, R)$ has been defined for integer l, the defining formulas clearly allow any real $l \in (l_{\min} - \epsilon, n - 1 + \epsilon)$ for $\epsilon > 0$ small enough. We will show that in this interval, $\partial \delta^*/\partial l > 0$.

Since $h(c, l, n)$ has continuous partial derivatives with respect to c and l, and $\partial h/\partial c > 0$ (see (11)), the implicit function theorem implies that the partial derivative of $c(l, n, R)$ with respect to l is

$$\frac{\partial c}{\partial l} = -\frac{\partial h}{\partial l}\left(\frac{\partial h}{\partial c}\right)^{-1}.$$

Now since $\delta^*(l, n, R) = \delta(c(l, n, R), l, n)$, the chain rule gives

$$\frac{\partial \delta^*}{\partial l} = \frac{\partial \delta}{\partial c}\frac{\partial c}{\partial l} + \frac{\partial \delta}{\partial l}.$$

The partial derivatives are

$$\begin{aligned} \frac{\partial h}{\partial l} &= c(\log d - \log c) - d + c \\ \frac{\partial h}{\partial c} &= l(\log d - \log c) \end{aligned}$$

and

$$\begin{aligned} \frac{\partial \delta}{\partial l} &= \Phi(c) - \Phi(d) + (d-c)\rho\left(\frac{d}{2}\right) \\ \frac{\partial \delta}{\partial c} &= l\left(\rho\left(\frac{c}{2}\right) - \rho\left(\frac{d}{2}\right)\right). \end{aligned}$$

Therefore,

$$\frac{\partial \delta^*}{\partial l} = \Phi(c) - \Phi(d) + (d-c)\rho\left(\frac{d}{2}\right) - \frac{(d - c - c\log(d/c))\left(\rho(d/2) - \rho(c/2)\right)}{\log(d/c)}.$$

Since $d = (1 - cl)/(n - l) > c$, we have $\partial \delta^*/\partial l > 0$ if and only if

$$\left[\Phi(c) - \Phi(d) + d\rho\left(\frac{d}{2}\right) - c\rho\left(\frac{c}{2}\right)\right]\log\left(\frac{d}{c}\right) > (d-c)\left[\rho\left(\frac{d}{2}\right) - \rho\left(\frac{c}{2}\right)\right]. \tag{19}$$

In the rest of the proof we will show that (19) holds for all $d > c > 0$ which will imply the claim of the lemma.

By assumption, $u(t) = \rho(e^t)$ is convex, and hence absolutely continuous, i.e., it is the integral of its derivative which exists almost everywhere (see, e.g., [12]). It follows that $\rho(x) = u(\log x)$ is also absolutely continuous, and since its right derivative $\rho_+'(x)$ exists for all $x > 0$ $(\rho_+'(x) = x^{-1}u_+'(\log x))$, we have

$$\rho\left(\frac{d}{2}\right) - \rho\left(\frac{c}{2}\right) = \int_{c/2}^{d/2} \rho_+'(x)\, dx. \tag{20}$$

If $\rho(x)$ is differentiable at some $x > 0$, then $x\rho(x/2) - \Phi(x)$ is also differentiable at this point and

$$\frac{d}{dx}\left[x\rho\left(\frac{x}{2}\right) - \Phi(x)\right] = \frac{x}{2}\rho_+'\left(\frac{x}{2}\right).$$

The absolute continuity of ρ implies that $x\rho(x/2)$ is also absolutely continuous, and since Φ is continuously differentiable, we obtain that

$$\begin{aligned} \Phi(c) - \Phi(d) + d\rho\left(\frac{d}{2}\right) - c\rho\left(\frac{c}{2}\right) &= \int_c^d \frac{x}{2}\rho_+'\left(\frac{x}{2}\right) dx \\ &= 2\int_{c/2}^{d/2} x\rho_+'(x)\, dx. \end{aligned} \tag{21}$$

This and (20) show that (19) can be rewritten as

$$\left(\frac{2}{d-c}\int_{c/2}^{d/2} x\rho_+'(x)\, dx\right)\left(\frac{2}{d-c}\int_{c/2}^{d/2} \frac{1}{x}\, dx\right) > \left(\frac{2}{d-c}\int_{c/2}^{d/2} \rho_+'(x)\, dx\right).$$

Letting $f(x) = x\rho_+'(x)$ and $g(x) = 1/x$, the above is equivalent to

$$E[f(Z)]E[g(Z)] > E[f(Z)g(Z)] \tag{22}$$

where Z is a random variable uniformly distributed over the interval $(c/2, d/2)$. Since $f(x) = u_+'(\log x)$, the strict convexity of u implies that f is strictly increasing. Since g is strictly decreasing, (22) holds by a classical inequality of Chebyshev [13]. (In fact, (22) easily follows by expanding the expectation

$$E[(f(Z) - f(Y))(g(Z) - g(Y))]$$

where Y is independent of Z but has the same distribution, and by noticing that the expectation is negative since $(f(x) - f(y))(g(x) - g(y)) < 0$ for all $x, y > 0, x \neq y$.) □

Lemma 2: Let $R > 0$ and let N be a positive integer such that $R < \log N$. If $n \geq N$, then $\delta^*(n-N+1, n, R)$ defined in (13) is a strictly increasing function of n.

Proof: Just as in Lemma 1, the definition of $\delta^*(n-N+1, n, R)$ can be extended to any real-valued n such that $n \geq N - \epsilon$ (if $\epsilon > 0$ is small enough). We will show that $\partial\delta^*(n-N+1, n, R)/\partial n > 0$ for all $n \geq N$.

To simplify the notation, let

$$\hat{\delta}(c, n) = (n-N+1)\Phi(c) + (N-1)\Phi(d)$$

and

$$\hat{h}(c, n) = -(n-N+1)c\log c - (1-(n-N+1)c)\log d.$$

Then the chain rule implies

$$\frac{\partial\delta^*(n-N+1, n, R)}{\partial n} = \frac{\partial\hat{\delta}}{\partial c}\frac{\partial c}{\partial n} + \frac{\partial\hat{\delta}}{\partial n}. \tag{23}$$

We have

$$\frac{\partial\hat{\delta}}{\partial c} = (n-N+1)\left(\rho\left(\frac{c}{2}\right) - \rho\left(\frac{d}{2}\right)\right)$$
$$\frac{\partial\hat{\delta}}{\partial n} = \Phi(c) - c\rho\left(\frac{d}{2}\right)$$

and by implicit differentiation in (14)

$$\frac{\partial c}{\partial n} = -\frac{\partial\hat{h}}{\partial n}\left(\frac{\partial\hat{h}}{\partial c}\right)^{-1} = -\frac{c(\log d - \log c + 1)}{(n-N+1)(\log d - \log c)}.$$

Substitution into (23) gives

$$\frac{\partial\delta^*(n-N+1, n, R)}{\partial n} = \Phi(c) - c\rho\left(\frac{c}{2}\right) + \frac{c(\rho(d/2) - \rho(c/2))}{\log(d/c)}.$$

Since $d = (1-c(n-N+1))/(N-1) > c$ if $c \in (0, 1/n)$, we have $\partial\delta^*(n-N+1, n, R)/\partial n > 0$ if and only if

$$\frac{\rho(d/2) - \rho(c/2)}{\log(d/2) - \log(c/2)} > \rho\left(\frac{c}{2}\right) - \frac{\Phi(c)}{c}. \tag{24}$$

By (21), the right side of the preceding inequality equals

$$\frac{1}{c/2}\int_0^{c/2} x\rho'_+(x)\,dx = \frac{1}{c/2}\int_0^{c/2} u'_+(\log x)\,dx$$

where $u(t) = \rho(e^t)$. Since u'_+ is strictly increasing, the last expression is less than $u'_+(\log(c/2))$. On the other hand, since u is strictly convex

$$\frac{\rho(d/2) - \rho(c/2)}{\log(d/2) - \log(c/2)} = \frac{u(\log(d/2)) - u(\log(c/2))}{\log(d/2) - \log(c/2)} > u'_+(\log(c/2))$$

which proves (24). □

Lemma 3: Let

$$f_r(x) = \frac{(x^r - 1)(x+1)}{x + x^r} - r\log x.$$

If $r > 3$, then there is an $x_r > 1$ such that $f_r(x) < 0$ for all $x \in (1, x_r)$.

Proof: Let $r > 3$. Since $f_r(1) = 0$, it is enough to prove that $f'_r(x) < 0$ in a nonempty interval $(1, x_r)$. We have

$$f'_r(x) = \frac{x^{2r} - rx^{2r-1} + rx^{r+1} - 2x^r + rx^{r-1} - rx + 1}{(x+x^r)^2} = \frac{g_r(x)}{(x+x^r)^2}.$$

Since $g_r(1) = 0$, it is sufficient to prove that $g'_r(x) < 0$ in some $(1, x_r)$. We have

$$g'_r(x) = 2rx^{2r-1} - r(2r-1)x^{2r-2} + r(r+1)x^r - 2rx^{r-1} + r(r-1)x^{r-2} - r.$$

Since $g'_r(1) = 0$, it is enough to prove that $g''_r(x) < 0$ in some $(1, x_r)$. We have

$$g''_r(x) = rx^{r-3}\big(2(2r-1)x^{r+1} - (2r-1)(2r-2)x^r + r(r+1)x^2 - 2(r-1)x + (r-1)(r-2)\big) = rx^{r-3}h_r(x).$$

Since $h_r(1) = 2r(3-r) < 0$ if $r > 3$, there exists an $x_r > 1$ such that $f_r(x) < 0$ for all $x \in (1, x_r)$, as claimed. □

ACKNOWLEDGMENT

The authors wish to thank an anonymous referee for many constructive comments.

REFERENCES

[1] T. Berger, "Optimum quantizers and permutation codes," *IEEE Trans. Inform. Theory*, vol. IT-18, pp. 759–765, Nov. 1972.
[2] N. Farvardin and J. W. Modestino, "Optimum quantizer performance for a class of non-Gaussian memoryless sources," *IEEE Trans. Inform. Theory*, vol. IT-30, pp. 485–497, May 1984.
[3] J. C. Kieffer, T. M. Jahns, and V. A. Obuljen, "New results on optimal entropy-constrained quantization," *IEEE Trans. Inform. Theory*, vol. 34, pp. 1250–1258, Sept. 1988.
[4] A. N. Netravali and R. Saigal, "Optimum quantizer design using a fixed-point algorithm," *Bell Syst. Tech. J*, vol. 55, pp. 1423–1435, Nov. 1976.
[5] P. A. Chou, T. Lookabaugh, and R. M. Gray, "Entropy-constrained vector quantization," *IEEE Trans. Acoust. Speech, Signal Processing*, vol. 37, pp. 31–42, Jan. 1989.
[6] D. L. Neuhoff and R. K. Gilbert, "Causal source codes," *IEEE Trans. Inform. Theory*, vol. IT-28, pp. 701–713, Sept. 1982.
[7] P. A. Chou and B. J. Betts, "When optimal entropy-constrained quantizers have only a finite number of codewords," in *Proc. IEEE Int. Symp. Information Theory*, Cambridge, MA, USA, Aug. 16–21, 1998, p. 97.
[8] A. György and T. Linder, On the structure of optimal entropy-constrained quantizers, 2000. in preparation.
[9] H. Gish and J. N. Pierce, "Asymptotically efficient quantizing," *IEEE Trans. Inform. Theory*, vol. IT-14, pp. 676–683, Sept. 1968.
[10] R. M. Gray, *Source Coding Theory*. Boston, MA: Kluwer, 1990.
[11] D. G. Luenberger, *Linear and Nonlinear Programming*, 2nd ed. Reading, MA.: Addison-Wesley, 1984.
[12] W. Rudin, *Real and Complex Analysis*, 3rd ed. New York: McGraw-Hill, 1987.
[13] G. H. Hardy, J. E. Littlewood, and G. Polya, *Inequalities*, 2nd ed. Cambridge, U.K.: Cambridge Univ. Press, 1959.
[14] F. Topsoe, "Instances of exact prediction and a new type of inequalities obtained by anchoring," in *Proc. IEEE Information Theory Workshop*, Kruger National Park, South Africa, June 20–25, 1999, p. 99.