

## MIT Open Access Articles

*Distributed scalar quantization for computing:  
High-resolution analysis and extensions*

The MIT Faculty has made this article openly available. **Please share** how this access benefits you. Your story matters.

**Citation:** Misra, Vinith, Vivek K. Goyal, and Lav R. Varshney. "Distributed Scalar Quantization for Computing: High-Resolution Analysis and Extensions." IEEE Transactions on Information Theory 57.8 (2011): 5298–5325.

**As Published:** <http://dx.doi.org/10.1109/TIT.2011.2158882>

**Publisher:** Institute of Electrical and Electronics Engineers (IEEE)

**Persistent URL:** <http://hdl.handle.net/1721.1/73159>

**Version:** Author's final manuscript: final author's manuscript post peer review, without publisher's formatting or copy editing

**Terms of use:** Creative Commons Attribution-Noncommercial-Share Alike 3.0



# Distributed Scalar Quantization for Computing: High-Resolution Analysis and Extensions

Vinith Misra, Vivek K Goyal, *Senior Member, IEEE*, and Lav R. Varshney, *Member, IEEE*

## Abstract

Communication of quantized information is frequently followed by a computation. We consider situations of *distributed functional scalar quantization*: distributed scalar quantization of (possibly correlated) sources followed by centralized computation of a function. Under smoothness conditions on the sources and function, companding scalar quantizer designs are developed to minimize mean-squared error (MSE) of the computed function as the quantizer resolution is allowed to grow. Striking improvements over quantizers designed without consideration of the function are possible and are larger in the entropy-constrained setting than in the fixed-rate setting. As extensions to the basic analysis, we characterize a large class of functions for which regular quantization suffices, consider certain functions for which asymptotic optimality is achieved without arbitrarily fine quantization, and allow limited collaboration between source encoders. In the entropy-constrained setting, a single bit per sample communicated between encoders can have an arbitrarily-large effect on functional distortion. In contrast, such communication has very little effect in the fixed-rate setting.

## Index Terms

Asymptotic quantization theory, distributed source coding, optimal point density function, rate-distortion theory

## I. INTRODUCTION

CONSIDER a collection of  $n$  spatially-separated sensors, each measuring a scalar  $X_j$ ,  $j = 1, 2, \dots, n$ . As shown in Fig. 1, the measurements are encoded and communicated over rate-limited links to a sink node without any interaction between the sensors. The sink node computes an estimate of the function  $g(X_1^n) = g(X_1, X_2, \dots, X_n)$  from the received data. This may be interpreted as a special case of the *distributed source coding* problem in which distortion is measured as the mean-squared error of the function estimate. We refer to this special case as *distributed functional source coding* to emphasize that it is the function  $g(X_1^n)$  and not the source vector  $X_1^n$  that is being reconstructed. Similarly, we will refer to approximate representation of  $X_1^n$  under mean-squared error distortion as *ordinary source coding*. Restricting to scalar quantization, this *distributed functional scalar quantization* (DFSQ) problem is the central subject of this paper. Compared to ordinary source coding, DFSQ can provide performance improvements in addition to any that are rooted in statistical dependence of the  $X_j$ s; for clarity, most examples presented here are for cases with independent  $X_j$ s.

### A. Summary of Main Contributions

The primary aim of this paper is to develop a high resolution approach to the analysis of DFSQ. To this end, we consider for each source variable  $X_j$  a sequence of companding quantizers  $\{Q_K^j\}$  of increasing resolution  $K$ . Under fairly loose smoothness requirements on the function  $g(x_1^n)$  and the source probability density function (pdf)  $f(x_1^n)$ , high-resolution analysis yields a choice for  $\{(Q_K^1, \dots, Q_K^n)\}_{K=1}^\infty$  that outperforms any other choice of companding quantizer sequences at sufficiently high resolution. This analysis also gives an approximation for the resulting distortion-rate function that has relative error which vanishes as  $K \rightarrow \infty$ .

There are situations in which designing quantizers to minimize the MSE of the function estimate is no different than designing them for low MSEs  $\mathbf{E}[(X_j - \hat{X}_j)^2]$ ,  $j = 1, 2, \dots, n$ . Our analysis will show, for example, that there is little advantage from accounting for  $g$  when  $g$  is linear. However, there are also cases in which the improvement is very large for large values of  $n$ ; examples in Section V feature distortion improvement over ordinary source coding by a factor that is polynomial in  $n$  in the fixed-rate case and exponential in  $n$  in the variable-rate case.

In addition to developing a basic theory in which there are no interactions between quantizers and certain limitations on  $g$  simplify our analysis, we consider several extensions. First, we permit nonregular quantizers and demonstrate that if the function  $g(x_1^n)$  satisfies a loose *equivalence-free* condition then optimal quantizers are regular at sufficiently high rate. Next, we explore a situation in which the high-resolution analysis breaks down because there is an interval where the marginal

This material is based upon work supported by the National Science Foundation under Grant No. 0729069.

The material in this paper was presented in part at the Information Theory and its Applications Workshop, La Jolla, California, January/February 2008; and the IEEE Data Compression Conference, Snowbird, Utah, March 2008.

V. Misra (email: vinith@stanford.edu) was with the Massachusetts Institute of Technology when this work was completed and is now with the Department of Electrical Engineering, Stanford University, Stanford, CA 94305 USA. V. K. Goyal (email: vgoyal@mit.edu), and L. R. Varshney (email: lrv@mit.edu) are with the Department of Electrical Engineering and Computer Science and the Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, MA 02139 USA. L. R. Varshney is also with the Laboratory for Information and Decision Systems.

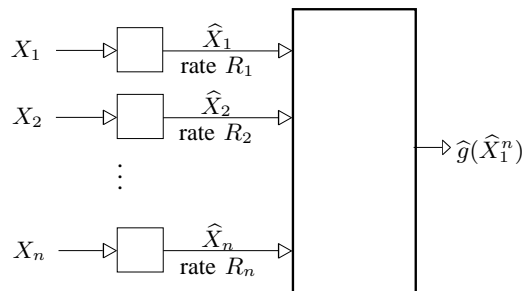


Fig. 1. Distributed functional source coding.

density  $f_{X_j}$  is positive but the optimal companding quantizer sequence for  $X_j$  is not arbitrarily fine. This prompts the concept of a *don't care interval*, a mixture of low- and high-resolution, and connections with [1]. Finally, we allow rate-constrained information communicated from encoder 2 to encoder 1 to affect the encoding of  $X_1$ . We call this *chatting* and bound its effect on the distortion  $D$ . In the fixed-rate setting, the reduction in distortion can be no more than if  $R_1$  were increased by the same rate; in the variable-rate setting, the reduction in distortion can be arbitrarily large.

For ordinary quantization problems, high-resolution analysis is not interesting for a discrete source because the distortion reaches zero at some finite resolution. Indeed, as in most works using high-resolution analysis, we assume that the source random variables are jointly continuous, i.e., that a joint probability density function for  $X_1^n$  exists. Similarly, high-resolution analysis of DFSQ may be uninteresting when  $g(X_1^n)$  is discrete because zero functional distortion may be achieved at some finite resolution. We do not explicitly require  $g(X_1^n)$  to be a continuous random variable, but the continuity of  $g$  that we do require eliminates many situations in which zero functional distortion may be achieved at some finite resolution.

## B. Related Work

DFSQ has strong connections to several problems that have been studied in prior work on quantization and distributed source coding. We provide a brief summary of some of these connections here. This paper is restricted to high-resolution analysis of companding scalar quantizers for real-valued sources. Contrarily, some related works deal with lossless source coding or lossy vector quantization, often in the (Shannon-theoretic) limit of large block length, at any rate.

Consider the situation depicted in Fig. 1 with  $n = 2$ . In general,  $X_1$  and  $X_2$  are memoryless, stationary random processes and  $g$  is a function of the two. Several topics arise by considering special cases of this formulation.

When  $g$  is the identity function, the goal is to reconstruct the source variables themselves; often the correlation between  $X_1$  and  $X_2$  is of primary interest. Slepian and Wolf solve this problem in the infinite blocklength regime for lossless representation of sources drawn from a discrete alphabet [2]. The lossy problem for sources from a discrete alphabet, restricted to scalar quantization followed by block entropy coding, is considered in [3].

In the setting with lossy representation of continuous sources, one might consider applying Slepian–Wolf coding to the output of local quantizers for each of the sources. This approach, with vector quantization performed on blocks of each of the sources, is optimal at all rates for jointly Gaussian sources and MSE distortion [4]. This approach is also optimal in the asymptotic regime of both large block length and high resolution [5]. The general lossy multiterminal source coding problem for large block length but finite rates, whether for discrete or continuous alphabet sources, is open.

While this paper restricts to scalar quantization of the sources, the use of Slepian–Wolf coding on the output of these quantizers is considered (Sec. IV-D). Note that since the identity function has a vector output, our DFSQ formulation technically does not permit this choice of  $g$ , but that only minor modification of the proofs are required to permit vector-valued functions.

If  $g(X_1, X_2) = X_1$  and  $R_2$  is unconstrained, then  $X_2$  can be viewed as receiver side information available at the decoder. The trade-off between  $R_1$  and distortion (of  $X_1$  alone) in the large block length regime is given by the Wyner–Ziv rate-distortion function [6], [7]. Rebollo-Monedero *et al.* examined this scenario at high resolution but any block length, and showed that providing receiver side information to the encoder yields no improvement in performance [8], cf. [9]. Under suitable constraints on the distortion metric, one may also view  $X_2$  as receiver side information that determines the distortion measure on  $X_1$ , drawing a connection to [10] and to work on non-MSE distortion functions [11].

For general  $g$  and unconstrained  $R_2$ , the lossy problem has been studied by Yamamoto [12] and later by Feng *et al.* [13], who provide an assortment of rate-loss bounds on performance in the large block length setting. The lossless setting has been explored by Orlitsky and Roche [14].

In the large block length regime for lossless coding, Han and Kobayashi [15] studied the classification of functions according to whether the rate region is the same as that for the identity function (i.e., the same as the Slepian–Wolf rate region). Their results are conclusive when  $n = 2$  and the source alphabets are finite. This distributed version of the problem for general  $g$ , minimizing the sum-rate  $R_1 + R_2$ , was later investigated by Doshi *et al.* [16].

Let  $Y = g(X_1, X_2)$ . Then  $Y$  may be interpreted as a *remote source* that is observed only through  $X_1$  and  $X_2$ , leading to a remote source multiterminal source coding problem [17]. Alternatively,  $\{Y = X_0, X_1, X_2\}$ , can be thought of as a source triple and the problem in Fig. 1 as a two-help-one problem with  $R_0 = 0$  [18].

Most of the above examples involve block coding of  $X_1$  and  $X_2$ , and results are obtained by allowing the block length to grow arbitrarily large. While the variable-length DFSQ analysis does utilize block entropy coding and Slepian–Wolf coding,  $X_1$  and  $X_2$  must first pass through scalar quantizers. Even though the samples of  $X_1$  and  $X_2$  are i.i.d., there would still be geometric benefits to using vector quantization over blocks of samples; this is left to future work.

Quantization with a functional motive bears strong resemblance to the idea of “task-oriented quantization.” There has been considerable work in this direction for detection, classification, and estimation, including high-rate analysis [19]–[21]. The use of a function at the decoder can be seen as inducing a non-MSE distortion measure on the source data. In this sense, a thread may be drawn to perceptual source coding [22], where a non-MSE distortion reflects human sensitivity to audio or video.

Under appropriate constraints on the function  $g$ , one may consider it as having introduced a *locally quadratic* distortion measure on the source  $X_1^n$ . In [23], Linder *et al.* consider quantization via companding functions for locally quadratic distortion measures. We say more about connections to this work in Section IV-E.

Interesting related problems have also arisen without a requirement of distributed coding. Rather than having a single function  $g$ , one may consider a set of functions  $\{g_\alpha\}_{\alpha \in \mathcal{A}}$  and define

$$D_g = \mathbf{E} \left[ d(g_\alpha(X_1^n), g_\alpha(\hat{X}_1^n)) \right],$$

where  $\alpha$  is a random variable taking values in index set  $\mathcal{A}$ . One may consider this a special case of the Wyner-Ziv problem with  $\alpha$  as decoder side information and a functional distortion measure. In such a setting, fixed- and variable-rate quantization to minimize MSE was studied by Bucklew in the high-rate regime [24]. Note that if the function were known deterministically to the encoder, one could do no better than to simply compute the function and encode the result.

### C. Structure of Paper

We start in Section II by reviewing the high-resolution approximation techniques used in our analysis. In Section III we obtain optimal fixed- and variable-rate functional quantizers for the  $n = 1$  case; while not important in practice, this case illustrates the role of monotonicity and smoothness of  $g$ . Generalizations to arbitrary  $n$ , under similar restrictions on  $g(\cdot)$ , are given in Section IV. Some notable examples in Section V are those that show dramatic scaling of distortion with respect to  $n$ . Some arguments in Sections II and III are meant only to build intuition; the technical results of those sections are rigorously justified as special cases of statements in Section IV.

The second half of the paper extends the basic theory of Section IV. Section VI addresses the use of non-regular companding quantizers and shows that a weak *equivalence-free* condition guarantees regularity of the optimal companding quantizer sequence. In the process we develop the notion of high-resolution non-regular quantization. In Section VII, we consider certain conditions that cause the high-resolution approach to lead to an optimal quantizer for  $X_j$  that does not have high resolution over the entire support of  $f_{X_j}$ . A modified analysis and design procedure yields a “rate amplification” in the variable-rate case. Limited communication between encoders, or chatting, is studied in Section VIII, and concluding comments appear in Section IX.

## II. UNIVARIATE ORDINARY QUANTIZATION

To introduce both notation and techniques, the high-resolution analysis of scalar quantizers under MSE distortion is reviewed in this section.

### A. Definitions

A  $K$ -level quantizer on  $[0, 1]$  is a function  $Q_K : [0, 1] \rightarrow [0, 1]$  with a range consisting of  $K$  points. The expected distortion of  $Q_K$  applied to random variable  $X$  taking values in  $[0, 1]$  is given by  $D(Q_K) = \mathbf{E} [d(X, Q(X))]$ , where  $d : [0, 1] \times [0, 1] \rightarrow [0, \infty)$  is an appropriately chosen distortion function. Squared-error distortion  $d(x, y) = (x - y)^2$  is both a frequent and analytically-tractable choice. In *fixed-rate* (or codebook-constrained) quantization, the rate is defined as the logarithm of the number of levels,  $R = \log K$ , where all logarithms have base 2. In *variable-rate* (or entropy-constrained) quantization, the rate is defined as the entropy of the quantizer output,  $R = H(Q_K(X))$ . An optimal fixed-rate or variable-rate quantizer minimizes distortion subject to a constraint on the applicable rate.

A value in the range of  $Q_K$  is called a *quantizer point* or *reconstruction point*, and the inverse image under  $Q_K$  of a quantizer point is called a *cell* or *partition region*. If each cell is an interval and the associated reconstruction point lies within the interval, the quantizer is called *regular*. For a distortion function that increases with the difference of its arguments (e.g. squared-error distortion), the optimal fixed-rate quantizer is regular. If the distortion function is also convex in the difference of its arguments and the source distribution is non-atomic, the optimal variable-rate quantizer is regular as well [25, Sect. 6.2] [26].

A *componder* function  $w : [0, 1] \rightarrow [0, 1]$  is continuous, increasing, differentiable almost everywhere, and invertible on  $[0, 1]$ . Furthermore,  $w(0) = 0$  and  $w(1) = 1$ . The  $K$ -level uniform quantizer on  $[0, 1]$  is defined as

$$Q_K^U(x) = \begin{cases} \frac{2i-1}{2K}, & \text{for } x \in \left(\frac{i-1}{K}, \frac{i}{K}\right], i = 1, 2, \dots, K; \\ \frac{1}{2K}, & \text{for } x = 0. \end{cases}$$

For squared-error distortion and more generally, optimal quantizers satisfy a stronger condition than regularity:

$$x < y \quad \text{implies} \quad Q_K(x) \leq Q_K(y).$$

They can thus be realized in companding form:

$$Q_K(x) = w^{-1}(Q_K^U(w(x))).$$

A quantizer that has a companding form may equivalently be defined by its *point density function*  $\lambda(x)$ :

$$\lambda(x) = w'(x),$$

which always satisfies  $\int_0^1 \lambda(x) dx = w(1) - w(0) = 1 - 0 = 1$  by the fundamental theorem of calculus. For small  $\delta$  and large resolution  $K$ , one may observe that  $\delta \lambda(x)$  approximates the fraction of quantizer points in an interval of length  $\delta$  around  $x$ . Because of this intuitive relationship to quantizer structure, we will use the point density description instead of the compander description whenever possible, with  $Q_K^\lambda(x)$  denoting a quantizer of resolution  $K$  and point density function  $\lambda$ . A *companding quantizer sequence*  $\{Q_K^\lambda\}_{K=1}^\infty$  refers to a sequence of quantizers generated with the same point density  $\lambda$  and indexed by resolution  $K$ . Our interest will be in optimizing these quantizer sequences.

The *distortion-resolution function*  $d(K; \lambda)$  for a companding quantizer sequence  $\{Q_K^\lambda\}$  indexes the distortion of the sequence by the resolution  $K$ :

$$d(K; \lambda) = \mathbf{E} \left[ |X - Q_K^\lambda(X)|^2 \right].$$

The fixed-rate *resolution-rate* function  $K_{\text{fr}}(R; \lambda) = \lfloor 2^R \rfloor$  is the largest resolution that satisfies a fixed-rate constraint. Similarly, the variable-rate resolution-rate function  $K_{\text{vr}}(R; \lambda)$  is the largest resolution that satisfies a variable-rate constraint. Specifically,  $K_{\text{vr}}(R; \lambda)$  is the largest resolution such that the entropy of the quantized output  $H(Q_{K_{\text{vr}}}^\lambda(X))$  is less than the rate constraint  $R$ :

$$K(\lambda; R) = \max_{H(Q_K^\lambda(X)) \leq R} K.$$

The quality of a quantizer sequence  $\{Q_K^\lambda\}$  is measured by its distortion-rate function. The fixed-rate distortion-rate function measures the distortion of the highest-resolution element of the sequence that satisfies the fixed-rate constraint:  $D_{\text{fr}}(R; \lambda) = d(K_{\text{fr}}(R; \lambda), \lambda)$ . Similarly, the variable-rate distortion-rate function measures the distortion of the highest-resolution element of the sequence that satisfies the variable-rate constraint:  $D_{\text{vr}}(R; \lambda) = d(K_{\text{vr}}(R; \lambda), \lambda)$ .

Under a fixed-rate constraint, we say that a companding quantizer sequence  $\{Q_K^{\lambda^*}\}$  is *asymptotically better* than another  $\{Q_K^\lambda\}$  if

$$\limsup_{R \rightarrow \infty} \frac{D_{\text{fr}}(R; \lambda^*)}{D_{\text{fr}}(R; \lambda)} \leq 1.$$

Essentially, we compare the best rate- $R$  quantizers from each sequence. If  $\{Q_K^{\lambda^*}\}$  is asymptotically better than all other quantizer sequences, we say  $\{Q_K^{\lambda^*}\}$  and  $\lambda^*$  are asymptotically *fixed-rate optimal*.

Analogously, an asymptotically *variable-rate optimal* quantizer sequence  $\{Q_K^{\lambda^*}\}$  is asymptotically better than any other  $\{Q_K^\lambda\}$ :

$$\limsup_{R \rightarrow \infty} \frac{D_{\text{vr}}(R; \lambda^*)}{D_{\text{vr}}(R; \lambda)} \leq 1.$$

Note that while we only consider optimality among the set of regular companding quantizer sequences, Linder [27] provided conditions for a source probability distribution function under which a companding quantizer sequence can be optimal in a more general sense.

## B. Problem Statement

A sequence of quantizers is to be applied to a source  $X$  with pdf  $f_X$  supported on the interval  $[0, 1]$ . The distortion of the quantizers is measured by squared error. For any fixed- or variable-rate constraint, the optimal quantizer can be realized in companding form, so we seek an asymptotically optimal companding function.

For high-resolution techniques to be valid, both the companding function and the source pdf must satisfy certain smoothness requirements. We assume the source satisfies conditions UO1 and UO2, and we optimize only among companding functions that satisfy UO3 and UO4:

UO1. The source pdf  $f$  is bounded and supported on the interval  $[0, 1]$ .

- UO2. The first derivative of the source pdf  $f'$  is defined and bounded on all but a finite number of points in  $[0, 1]$ .  
 UO3. We optimize among companding functions that are differentiable.  
 UO4. The integral  $\int_0^1 f(x)w'(x)^{-2} dx$  is finite.

### C. Solution via High-Resolution Analysis

The quantities of fundamental interest in the analysis of companding quantizer sequences are the fixed- and variable-rate distortion-rate functions  $D_{\text{fr}}(R; \lambda)$  and  $D_{\text{vr}}(R; \lambda)$ , which describe the distortion of fixed- and variable-rate companding quantizers with rate  $R$  and point density  $\lambda$ . High resolution analysis consists of several approximations that allow one to derive asymptotically accurate versions of both  $D_{\text{fr}}^{\text{HR}}(R; \lambda)$  and  $D_{\text{vr}}^{\text{HR}}(R; \lambda)$ . Specifically, under appropriate restrictions on the source pdf we will show that

$$\lim_{R \rightarrow \infty} \frac{D_{\text{fr}}^{\text{HR}}(R; \lambda)}{D_{\text{fr}}(R; \lambda)} = \lim_{R \rightarrow \infty} \frac{D_{\text{vr}}^{\text{HR}}(R; \lambda)}{D_{\text{vr}}(R; \lambda)} = 1. \quad (1)$$

In Sec. II-C1, the approximate distortion-resolution function  $d^{\text{HR}}(K; \lambda)$  is derived. Then, in Sec. II-C2, the approximate resolution-rate function  $K^{\text{HR}}(R; \lambda)$  is obtained for both fixed- and variable-rate constraints. Finally, in Sec. II-C3 these two quantities yield the approximate distortion-rate functions  $D_{\text{fr}}^{\text{HR}}(R; \lambda)$  and  $D_{\text{vr}}^{\text{HR}}(R; \lambda)$ . The derivation we provide is left informal and is not intended to prove that assumptions UO1–UO4 yield (1); this follows either from Linder [27] or as a special case of Theorem 6 in Sec. III. For further technical details and references to original sources, see [28]. Finally, in Sec. II-C4, the approximate distortion-rate functions are optimized through choice of point density (companding function). The sequences of companding quantizers yielded by this optimization are shown to be asymptotically fixed- or variable-rate optimal.

1) *The Distortion-Resolution Function:* As previously defined,  $d(K; \lambda)$  is the distortion of the companding quantizer with resolution  $K$ . We now define an approximation  $d^{\text{HR}}(K; \lambda)$ , known as the approximate distortion-resolution function. For rigorous proof that

$$\lim_{K \rightarrow \infty} d^{\text{HR}}(K; \lambda)/d(K; \lambda) = 1, \quad (2)$$

we refer to the main result of Linder [27], or to Theorem 9 with  $g(x) = x$ .

Let  $X$  be a random variable with pdf  $f_X(x)$ , let  $Q_K^\lambda$  be a  $K$ -point companding quantizer, and suppose  $\lambda$  and  $f$  satisfy assumptions UO1–UO4. Let  $\{\beta_i\}_{i \in \mathcal{I}} = Q_K^\lambda([0, 1])$  be the reconstruction points, and let  $S_i = (Q_K^\lambda)^{-1}(\beta_i)$ ,  $i \in \mathcal{I}$ , be the corresponding partition regions.

The distortion of the quantizer is

$$\begin{aligned} d(K; \lambda) &= \mathbf{E} \left[ (X - \hat{X})^2 \right] \\ &= \sum_{i \in \mathcal{I}} \mathbf{E} \left[ (X - \beta_i)^2 \mid X \in S_i \right] \mathbf{P}(X \in S_i) \end{aligned} \quad (3)$$

by the law of total expectation. The initial aim of high-resolution theory is to express this distortion as an integral involving  $f_X$ . To that end, we make the following approximations about the source and quantizer:

HR1.  $f_X$  may be approximated as constant on each  $S_i$ .

HR2. The size of the cell containing  $x$  is approximated with the help of the point density function:

$$x \in S_i \quad \Rightarrow \quad \text{length}(S_i) \sim (K\lambda(x))^{-1}, \quad (4)$$

where  $\sim$  means that the ratio of the two quantities goes to 1 with increasing resolution  $K$ . This is the meaning of “ $\sim$ ” for the remainder of the paper.

The first approximation follows from the smoothness of  $f_X$  (assumptions UO1 and UO2), while the second follows from the smoothness of  $w(x)$  (assumption UO3).

Now we can approximate each non-boundary term in (3). By HR1,  $\beta_i$  should be approximately at the center of  $S_i$ , and the length of  $S_i$  then makes the conditional expectation approximately  $\frac{1}{12}(K\lambda(\beta_i))^{-2}$ . Invoking Assumption HR1 again, the  $i$ th term in the sum is  $\int_{x \in S_i} \frac{1}{12}(K\lambda(\beta_i))^{-2} f_X(x) dx$ . Finally,

$$\begin{aligned} d(K; \lambda) &\sim \int_0^1 \frac{(K\lambda(x))^{-2}}{12} f_X(x) dx = \frac{1}{12K^2} \mathbf{E} [\lambda^{-2}(X)] \\ &= d^{\text{HR}}(K; \lambda). \end{aligned} \quad (5)$$

2) *The Resolution-Rate Function:* For a fixed-rate quantizer, the resolution-rate relationship is given simply by  $K_{\text{fr}}(R; \lambda) = \lfloor 2^R \rfloor$ , and it is approximated with vanishing relative error by  $K_{\text{fr}}^{\text{HR}}(R; \lambda) = 2^R$ . The variable-rate resolution-rate function is more difficult to approximate.

As long as the quantization is fine ( $\lambda(x) > 0$ ) wherever the density is positive, we can approximate the output entropy of a quantizer using the point density. Defining  $p(x)$  as  $\mathbf{P}(X \in S_i)$  for  $x \in S_i$ , and letting  $h(X)$  denote the differential entropy of  $X$ ,

$$\begin{aligned}
H(Q_K^\lambda(X)) &= - \sum_{i \in \mathcal{I}} \mathbf{P}(X \in S_i) \log \mathbf{P}(X \in S_i) \\
&\stackrel{(a)}{=} - \int_0^1 f_X(x) \log p(x) dx \\
&\stackrel{(b)}{\approx} - \int_0^1 f_X(x) \log(f_X(x)/(K\lambda(x))) dx \\
&= - \int_0^1 f_X(x) \log f_X(x) dx \\
&\quad + \int_0^1 f_X(x) \log(K\lambda(x)) dx \\
&= h(X) + \log K + \mathbf{E}[\log \lambda(X)], \tag{6}
\end{aligned}$$

where (a) follows from the definition of  $p(x)$ ; and (b) involves approximating the source pdf as constant in each cell and (4).

A generalized version of this approximation is proven rigorously in [23]. We state it here as a lemma.

*Lemma 1:* Suppose the source  $X$  has a density over  $[0, 1]$  and a finite differential entropy  $h(X)$ . Then if  $\mathbf{E}[\log \lambda(X)]$  is finite,

$$\lim_{R \rightarrow \infty} \left[ H(Q_{K(R; \lambda)}^\lambda(X)) - \log K(R; \lambda) \right] = h(X) + \mathbf{E}[\log \lambda(X)].$$

*Proof:* Follows as a special case of Proposition 2 in [23]. ■

With the insight of this approximation, we define:

*Definition 1:* The variable-rate approximate resolution-rate function  $K_{\text{vr}}^{\text{HR}}(R; \lambda)$  is given by

$$\log K_{\text{vr}}^{\text{HR}}(R; \lambda) = R - h(X) - \mathbf{E}[\log \lambda(X)].$$

*Lemma 2:* The error between the log of the variable-rate approximate resolution-rate function  $\log K_{\text{vr}}^{\text{HR}}(R; \lambda)$  and the log of the actual resolution-rate function  $K_{\text{vr}}(R; \lambda)$  goes to zero, i.e.

$$\lim_{R \rightarrow \infty} \log K_{\text{vr}}^{\text{HR}}(R; \lambda) - \log K_{\text{vr}}(R; \lambda) = 0.$$

*Proof:* The error of the approximation  $K_{\text{vr}}^{\text{HR}}$  may be written as

$$\log K_{\text{vr}}(R; \lambda) - \log K_{\text{vr}}^{\text{HR}}(R; \lambda) = \epsilon_R + H(Q_{K_{\text{vr}}(R; \lambda)}^\lambda(X)) - R,$$

where  $\epsilon_R$  goes to zero by Lemma 1. Furthermore, by definition  $K_{\text{vr}}(R; \lambda)$  has been chosen to be the largest resolution such that  $H(Q_{K_{\text{vr}}(R; \lambda)}^\lambda(X)) \leq R$ . We then have that

$$R - H(Q_{K_{\text{vr}}(R; \lambda)}^\lambda(X)) < H(Q_{K_{\text{vr}}(R; \lambda)+1}^\lambda(X)) - H(Q_{K_{\text{vr}}(R; \lambda)}^\lambda(X)),$$

i.e. the second term in the rate approximation error is bounded by the increment in entropy from an increment in resolution. By Lemma 1 once again, the increment in entropy may be bounded as

$$\begin{aligned}
&H(Q_{K_{\text{vr}}(R; \lambda)+1}^\lambda(X)) - H(Q_{K_{\text{vr}}(R; \lambda)}^\lambda(X)) \\
&= h(X) + \log(K_{\text{vr}}(R; \lambda) + 1) + \mathbf{E}[\log \lambda(X)] - h(X) - \log K_{\text{vr}}(R; \lambda) - \mathbf{E}[\log \lambda(X)] + \delta(R) \\
&= \log(K_{\text{vr}}(R; \lambda) + 1) - \log K_{\text{vr}}(R; \lambda) + \delta(R) \\
&= \log \frac{K_{\text{vr}}(R; \lambda) + 1}{K_{\text{vr}}(R; \lambda)} + \delta(R),
\end{aligned}$$

where  $\delta(R)$  goes to zero. Since  $K_{\text{vr}}(R; \lambda)$  diverges to infinity with  $R$ , this error goes to zero. ■

3) *The Distortion-Rate Functions*: The high-resolution distortion-rate function can be obtained by combining the distortion-resolution and resolution-rate functions. For fixed-rate,

$$D_{\text{fr}}^{\text{HR}}(R) = \frac{1}{12} \mathbf{E} [\lambda^{-2}(X)] 2^{-2R}, \quad (7a)$$

whereas for variable-rate

$$D_{\text{vr}}^{\text{HR}}(R) = \frac{1}{12} \mathbf{E} [\lambda^{-2}(X)] 2^{-2(R-h(X)-\mathbf{E}[\log \lambda(X)])}. \quad (7b)$$

Asymptotic validity in the sense of (1) follows in the fixed-rate case from (2) and from the fact that  $(K_{\text{fr}}(R; \lambda)/K_{\text{fr}}^{\text{HR}}(R; \lambda))^2$  goes to 1. In the variable-rate case, we may bound the error from use of  $K^{\text{HR}}(R; \lambda)$  in place of  $K(R; \lambda)$  as a multiplying factor of  $2^{2|K^{\text{HR}}(R; \lambda) - K(R; \lambda)|}$ , which by Lemma 2 goes to 1.

4) *Asymptotically-Optimal Companding Quantizer Sequences*: We seek asymptotically-optimal companding quantizer sequences for both fixed-rate and variable-rate constraints. By the following lemma, this reduces to minimizing the high-resolution distortion-rate functions of (7a) and (7b).

*Lemma 3*: Suppose  $\lambda_{\text{fr}}^*$  and  $\lambda_{\text{vr}}^*$  minimize  $D_{\text{fr}}^{\text{HR}}(R; \lambda)$  and  $D_{\text{vr}}^{\text{HR}}(R; \lambda)$  respectively. Then the quantizer sequences  $\{Q_{K_{\text{fr}}}^{\lambda_{\text{fr}}^*}\}$  and  $\{Q_{K_{\text{vr}}}^{\lambda_{\text{vr}}^*}\}$  are asymptotically fixed- and variable-rate optimal.

*Proof*: As the proof is virtually identical for fixed- and variable-rate cases, we only provide it for the variable-rate case.

Let  $\{Q_K^\lambda\}$  be any companding quantizer sequence. We are interested in proving that

$$\limsup_{R \rightarrow \infty} \frac{D_{\text{vr}}(R; \lambda_{\text{vr}}^*)}{D_{\text{vr}}(R; \lambda)} \leq 1.$$

The supremum limit on the left may be factored:

$$\begin{aligned} \limsup_{R \rightarrow \infty} \frac{D_{\text{vr}}(R; \lambda_{\text{vr}}^*)}{D_{\text{vr}}(R; \lambda)} &= \limsup_{R \rightarrow \infty} \frac{D_{\text{vr}}(R; \lambda_{\text{vr}}^*)}{D_{\text{vr}}^{\text{HR}}(R; \lambda_{\text{vr}}^*)} \frac{D_{\text{vr}}^{\text{HR}}(R; \lambda_{\text{vr}}^*)}{D_{\text{vr}}^{\text{HR}}(R; \lambda)} \frac{D_{\text{vr}}^{\text{HR}}(R; \lambda)}{D_{\text{vr}}(R; \lambda)} \\ &\stackrel{(a)}{\leq} \limsup_{R \rightarrow \infty} \frac{D_{\text{vr}}(R; \lambda_{\text{vr}}^*)}{D_{\text{vr}}^{\text{HR}}(R; \lambda_{\text{vr}}^*)} \limsup_{R \rightarrow \infty} \frac{D_{\text{vr}}^{\text{HR}}(R; \lambda_{\text{vr}}^*)}{D_{\text{vr}}^{\text{HR}}(R; \lambda)} \limsup_{R \rightarrow \infty} \frac{D_{\text{vr}}^{\text{HR}}(R; \lambda)}{D_{\text{vr}}(R; \lambda)} \end{aligned}$$

because the supremum limit of a product of positive sequences is upper-bounded by the product of their individual supremum limits. We can now bound each of these factors.

We have, by optimality of  $\lambda_{\text{vr}}^*$ , that  $D_{\text{vr}}^{\text{HR}}(R; \lambda) \geq D_{\text{vr}}^{\text{HR}}(R; \lambda_{\text{vr}}^*)$  for any  $R$  and therefore that

$$\limsup_{R \rightarrow \infty} \frac{D_{\text{vr}}^{\text{HR}}(R; \lambda_{\text{vr}}^*)}{D_{\text{vr}}^{\text{HR}}(R; \lambda)} \leq 1.$$

Furthermore, by (1), we have that

$$\lim_{R \rightarrow \infty} \frac{D_{\text{vr}}(R; \lambda_{\text{vr}}^*)}{D_{\text{vr}}^{\text{HR}}(R; \lambda_{\text{vr}}^*)} = \lim_{R \rightarrow \infty} \frac{D_{\text{vr}}^{\text{HR}}(R; \lambda)}{D_{\text{vr}}(R; \lambda)} = 1.$$

This proves the lemma. ■

Now we optimize the distortion-rate expressions. Because analogous optimizations appear in Sections III and IV, we explicitly derive both the optimizing point densities and the resulting distortion-rate functions. Our approach follows [29].

In the fixed-rate case, the problem is to minimize (7a) for a given value of  $R$ . This minimization may be performed with the help of Hölder's inequality:

$$\begin{aligned} D_{\text{fr}}^{\text{HR}}(R; \lambda) &= \frac{1}{12} 2^{-2R} \int_0^1 f_X(x) \lambda^{-2}(x) dx \\ &= \frac{1}{12} 2^{-2R} \int_0^1 f_X(x) \lambda^{-2}(x) dx \left( \int_0^1 \lambda(x) dx \right)^2 \\ &\geq \frac{1}{12} 2^{-2R} \int_0^1 (f_X(x) \lambda^{-2}(x))^{1/3} (\lambda(x))^{2/3} dx \\ &= \frac{1}{12} 2^{-2R} \left( \int_0^1 f_X^{1/3}(x) dx \right)^3, \end{aligned}$$

with equality only if  $\lambda(x) \propto f_X^{1/3}(x)$ . Thus,  $D_{\text{fr}}^{\text{HR}}$  is minimized by

$$\lambda(x) = f_X^{1/3}(x) / \left( \int_0^1 f_X^{1/3}(t) dt \right). \quad (8)$$

The resulting minimal distortion is

$$D_{\text{fr}}^{\text{HR}}(R) = \frac{1}{12} 2^{-2R} \left( \int_0^1 f_X^{1/3}(x) dx \right)^3 = \frac{1}{12} \|f_X\|_{1/3}^3 2^{-2R}, \quad (9)$$



where we have introduced a notation for the  $\mathcal{L}^{1/3}$  quasinorm.

For the variable-rate optimization, we use Jensen's inequality rather than Hölder's inequality:

$$\begin{aligned} D_{\text{vr}}^{\text{HR}}(R; \lambda) &= \frac{1}{12} 2^{-2(R-h(X))} \mathbf{E} [\lambda^{-2}(X)] 2^{2\mathbf{E}[\log \lambda(X)]} \\ &\stackrel{(a)}{\geq} \frac{1}{12} 2^{-2(R-h(X))} \mathbf{E} [\lambda^{-2}(X)] 2^{2\log \mathbf{E}[\lambda(X)]} \\ &= D_{\text{vr}}^{\text{HR}}(R), \end{aligned}$$

where (a) follows from the convexity of  $-\log(\cdot)$ . This lower bound is achieved when  $\lambda(X)$  is a constant. Thus  $\lambda(x) = 1$  is asymptotically optimal, i.e., the quantizer should be uniform.

Note that both variable- and fixed-rate quantization have  $\Theta(2^{-2R})$ , or  $-6$  dB/bit, dependence of distortion on rate. This is a common feature of ordinary quantizers with MSE distortion, but we demonstrate in Section VII that certain functional scenarios can cause distortion to fall even faster with the rate.

5) *Optimal Bit Allocation*: As a final preparatory digression, we state the solution to a typical resource allocation problem that arises several times in Section IV.

*Lemma 4*: Suppose  $D = \sum_{j=1}^n c_j 2^{-2R_j}$  for some positive constants  $\{c_j\}_{j=1}^n$ . Then the minimum of  $D$  over the choice of  $\{R_j\}_{j=1}^n$  subject to the constraint  $\sum_{j=1}^n R_j \leq nR$  is attained with

$$R_j = R + \frac{1}{2} \log \frac{c_j}{\left(\prod_{j=1}^n c_j\right)^{1/n}}, \quad j = 1, 2, \dots, n,$$

resulting in

$$D = n \left(\prod_{j=1}^n c_j\right)^{1/n} 2^{-2R}.$$

*Proof*: The result can be shown using the inequality for arithmetic and geometric means. It appeared first in the context of bit allocation in [30]; a full proof appears in [25, Sect. 8.3]. ■

The lemma does not restrict the  $R_j$ s to be nonnegative or to be integers. Such restrictions are discussed in [31].

### III. UNIVARIATE FUNCTIONAL QUANTIZATION

Let  $X$  be a random variable with pdf  $f_X(x)$  defined over  $[0, 1]$ , and let  $g : [0, 1] \rightarrow \mathbb{R}$  be the function of interest. A sequence of companding quantizers  $\{Q_K^\lambda\}$  is applied to the source  $X$ , and an estimate  $\hat{g}(Q_K^\lambda(X))$  is formed at the decoder, where  $\hat{g}$  is the estimator function. Functional distortion is measured by squared error  $D = \mathbf{E}[(g(X) - \hat{g}(Q_K^\lambda(X)))^2]$ . We seek an asymptotically-optimal estimator  $\hat{g}$  and companding function  $w$  that satisfy certain constraints.

Since we seek to answer this design question with high-resolution techniques, the function  $g$  and the source  $X$  must be restricted in a manner similar to conditions UO1–4 in Section II-B. For the moment we err on the side of being too strict. Sections VI and VII will significantly loosen these requirements.

UF1.  $g$  is monotonic.

UF2.  $g$  is Lipschitz continuous on  $[0, 1]$ , and the first- and second- derivatives of  $g$  are defined except possibly on a set of zero Jordan measure.

UF3. The source pdf  $f$  is continuous, bounded, and supported on the interval  $[0, 1]$ .

UF4. We optimize among companding functions  $w$  that are piecewise differentiable (and therefore a point density description  $\lambda$  is appropriate).

UF5. The integral  $\int_0^1 f(x)g'(x)^2 \lambda(x)^{-2} dx$  is defined and finite.

Throughout this paper, we assume that  $\hat{g}(t) = \mathbf{E}[g(X) | X \in S_i]$  for all  $t \in S_i$ . This achieves the minimum possible functional distortion  $\mathbf{E}[\text{var}(g(X) | Q_K^\lambda(X))]$ .

#### A. Sufficiency of Regular Quantizers

The following lemma relates monotonicity to regularity of optimal quantizers, thus justifying the optimization among companding quantizers:

*Lemma 5*: If  $g$  is monotonic, there exists an optimal functional quantizer of  $X$  that is regular.

*Proof*: The optimal functional quantizer in one dimension is induced by the optimal ordinary quantizer for the variable  $Y = g(X)$ . That is, one may compute the function  $g(X)$  and quantize it directly. Since the optimal ordinary quantizer for a real-valued source is regular, the optimal quantizer for  $Y$ , denoted by  $Q_Y(y)$  and having points  $\{\hat{y}_i\}_{i \in \mathcal{I}}$ , is regular.

$Q_Y(y)$  may be implemented by a quantizer for  $X$  with cells given by  $g^{-1}(Q_Y^{-1}(\hat{y}_i))$ . We know that  $Q_Y^{-1}(\hat{y}_i)$  is an interval since  $Q_Y$  is regular. Also, since  $g$  is monotonic, the inverse map  $g^{-1}$  applied to any interval in the range of  $g$  gives an interval. Thus  $g^{-1}(Q_Y^{-1}(\hat{y}_i))$  is an interval, which demonstrates that there exists a regular quantizer in  $X$  that is optimal. ■

### B. The Distortion-Resolution Function

Assumption UF2 is introduced so that a piecewise linear approximation of  $g$  suffices in estimating the functional distortion of the quantizer. More specifically, recalling the notation  $\{\beta_i\}_{i \in \mathcal{I}}$  for the quantizer points and  $\{S_i\}_{i \in \mathcal{I}}$  for the partition,

$$g_{\text{PL}}(x) = g(\beta_i) + g'(\beta_i)(x - \beta_i), \quad \text{for } x \in S_i, \quad i \in \mathcal{I}$$

may be interpreted as an approximation of  $g$  that leads to the high-resolution approximate distortion-resolution function.

The use of  $g_{\text{PL}}$  prompts us to give a name to the magnitude of the derivative of  $g$ . The distortion is then expressed using this function.

*Definition 2:* The *univariate functional sensitivity profile* of  $g$  is defined as  $\gamma(x) = |g'(x)|$ .

*Theorem 6:* Suppose a source  $X \in [0, 1]$  is quantized by a sequence of companding quantizers  $\{Q_K^\lambda\}$  with point density  $\lambda(x)$  and increasing resolution  $K$ . Further suppose that the source, quantizer, and function  $g : [0, 1] \rightarrow \mathbb{R}$  satisfy Assumptions UF1–5. Then the high-resolution distortion-resolution function is an *asymptotically accurate* approximation of the true distortion-resolution function:

$$d(K; \lambda) = \mathbf{E} [\text{var}(g(X) | Q_K^\lambda(X))] \sim \frac{1}{12K^2} \mathbf{E} [(\gamma(X)/\lambda(X))^2] = d^{\text{HR}}(K; \lambda). \quad (10)$$

*Proof:* Follows as a special case of Theorem 9. ■

### C. The Resolution-Rate Functions

The relationship between resolution and rate in the functional context is unchanged from the ordinary context. For a fixed-rate constraint, the resolution-rate function is given by  $K_{\text{fr}}(R; \lambda) = \lfloor 2^R \rfloor$  and is approximated at high-resolution by  $K_{\text{fr}}^{\text{HR}}(R; \lambda) = 2^R$ . For a variable-rate constraint, the resolution-rate function is given by the highest resolution such that the entropy of the quantized output is less than the rate constraint. This is approximated as before by  $\log K_{\text{vr}}^{\text{HR}}(R; \lambda) = R - h(X) - \mathbf{E}[\log \lambda(X)]$ . Both of these approximations continue to be asymptotically accurate, regardless of the distortion measure in use.

### D. The Distortion-Rate Functions

By combining the distortion-rate function with the resolution-rate function, the high-resolution distortion-rate function can be obtained. For fixed-rate,

$$D_{\text{fr}}^{\text{HR}}(R; \lambda) = \frac{1}{12} \mathbf{E} [(\gamma(X)/\lambda(X))^2] 2^{-2R}, \quad (11a)$$

whereas for variable-rate,

$$D_{\text{vr}}^{\text{HR}}(R; \lambda) = \frac{1}{12} \mathbf{E} [(\gamma(X)/\lambda(X))^2] 2^{-2(R - h(X) - \mathbf{E}[\log \lambda(X)])}. \quad (11b)$$

The asymptotic validity of these two expressions, as in (1), holds as it did in the ordinary case. For the fixed-rate expression, this follows from Theorem 6 and the fact that  $\lfloor 2^R \rfloor 2^{-R}$  approaches 1. For the variable-rate expression, the error from use of  $K^{\text{HR}}(R; \lambda)$  in the distortion-rate expression instead of  $K(R; \lambda)$  can be bounded as a multiplying factor of  $2^{2|K^{\text{HR}}(R; \lambda) - K(R; \lambda)|}$ , which by Lemma 2 goes to 1.

### E. Asymptotically-Optimal Companding Quantizer Sequences

We seek asymptotically-optimal companding quantizer sequences for fixed- and variable-rate constraints under a functional distortion measure. The lemma below demonstrates that it suffices to optimize the high-rate distortion-rate functions  $D_{\text{fr}}^{\text{HR}}$  and  $D_{\text{vr}}^{\text{HR}}$ .

*Lemma 7:* Suppose  $\lambda_{\text{fr}}^*$  and  $\lambda_{\text{vr}}^*$  minimize  $D_{\text{fr}}^{\text{HR}}(R; \lambda)$  and  $D_{\text{vr}}^{\text{HR}}(R; \lambda)$  respectively. Then the quantizer sequences  $\{Q_K^{\lambda_{\text{fr}}^*}\}$  and  $\{Q_K^{\lambda_{\text{vr}}^*}\}$  are asymptotically fixed- and variable-rate optimal.

*Proof:* The proof is virtually identical to that of Lemma 3. ■

The distortion expression (10) bears strong resemblance to (5), but with the probability density  $f_X(x)$  replaced with a *weighted density*  $\gamma^2(x)f_X(x)$ . Unlike the density  $f_X(x)$ , the weighted density  $\gamma^2(x)f_X(x)$  need not integrate to one. Optimal point densities and the resulting distortions now follow easily.

For fixed-rate coding, we are attempting to minimize the distortion (10) for a given value of  $K$ . Following the arguments in Section II-C4, the optimal point density is proportional to the cube root of the weighted density:

$$\lambda_{\text{fr}}^*(x) = \frac{(\gamma^2(x)f_X(x))^{1/3}}{\int_0^1 (\gamma^2(t)f_X(t))^{1/3} dt}. \quad (12)$$

The admissibility of this point density (assumption UF5) requires positivity of  $\lambda(x)$  everywhere  $f_X$  is positive. This excludes the possibility that  $\gamma(x) = 0$  for an interval  $x \in (a, b)$  such that  $\mathbf{P}(X \in (a, b)) > 0$  because in this case the quantization is

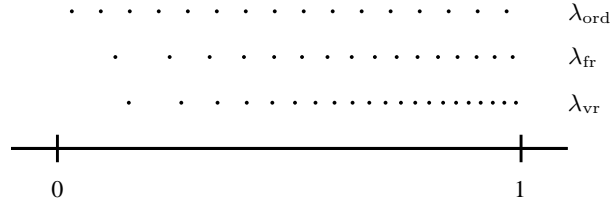


Fig. 2. Quantizer points illustrating the point densities derived in Example 1 at rate  $R = 4$ .

not fine for  $X \in (a, b)$ . We revisit this restriction in Section VII. By evaluating (11a) with point density (12), the resulting distortion is

$$D_{\text{fr}}^{\text{HR}}(R) = D_{\text{fr}}^{\text{HR}}(R; \lambda_{\text{fr}}^*) = \frac{1}{12} \|\gamma^2 f_X\|_{1/3} 2^{-2R}. \quad (13)$$

For variable-rate coding, a derivation very similar to that of ordinary variable-rate quantization may be performed. This yields an optimal point density that is proportional to the functional sensitivity profile:

$$\lambda_{\text{vr}}^*(x) = \frac{\gamma(x)}{\int_0^1 \gamma(t) dt}. \quad (14)$$

The restriction for  $\lambda$  to be positive wherever  $f_X$  is positive takes the same form as above (assumption UF5). The resulting distortion is

$$D_{\text{vr}}^{\text{HR}}(R) = D_{\text{vr}}^{\text{HR}}(R; \lambda_{\text{vr}}^*) = \frac{1}{12} 2^{2h(X) + 2\mathbf{E}[\log \gamma(X)]} 2^{-2R}. \quad (15)$$

The example below shows that even for univariate functions, there are benefits from functional quantization. It also illustrates the difference between the fixed- and variable-rate cases. While quantizing  $X$  instead of  $g(X)$  seems naïve, as we move to the distributed multivariate case it will not be possible to compute the function before quantization.

*Example 1:* Suppose  $X$  is uniformly distributed over  $[0, 1]$  and  $g(x) = x^2$ . For both fixed- and variable-rate, the optimal ordinary quantizer is uniform, i.e.,  $\lambda_{\text{ord}} = 1$ . With  $\gamma(x) = 2x$ , evaluating (11a) gives  $D_{\text{fr}}^{\text{HR}}(R; \lambda_{\text{ord}}) = D_{\text{vr}}^{\text{HR}}(R; \lambda_{\text{ord}}) = \frac{1}{9} 2^{-2R} \approx 0.111 \cdot 2^{-2R}$ .

The optimal point density for fixed-rate functional quantization is  $\lambda_{\text{fr}}^*(x) = \frac{5}{3} x^{2/3}$  and yields distortion

$$D_{\text{fr}}^{\text{HR}}(R) = \frac{1}{12} \|(2x)^2\|_{1/3} \cdot 2^{-2R} = \frac{9}{125} 2^{-2R} \approx 0.072 \cdot 2^{-2R}.$$

The optimal point density for variable-rate functional quantization is  $\lambda_{\text{vr}}^*(x) = 2x$ . With  $h(X) = 0$  and  $\mathbf{E}[\log \gamma(X)] = 1 - 1/(\ln 2)$ , the resulting distortion is

$$D_{\text{vr}}^{\text{HR}}(R) = \frac{1}{12} \cdot 4e^{-2} \cdot 2^{-2R} \approx 0.045 \cdot 2^{-2R}.$$

Quantizers designed with the three derived optimal point densities are illustrated in Fig. 2 for rate  $R = 4$ . The functionally-optimized quantizers put more points at higher values of  $x$ , where the function varies more quickly. In addition, the variable-rate quantizer is allowed more points ( $K = 21$ ) while meeting the rate constraint.

The interested reader can verify that  $D_{\text{fr}}^{\text{HR}}(R)$  and  $D_{\text{vr}}^{\text{HR}}(R)$  exactly match the performance obtained by designing optimal quantizers for  $Y = X^2$ .  $\square$

In the second example, we use a nonuniform source pdf with the same nonlinear function  $g$  to illustrate various quantities.

*Example 2:* Suppose  $X$  has the pdf  $f_X(x) = 3x^2$  over  $[0, 1]$  and  $g(x) = x^2$ . We illustrate a codebook-constrained quantizer with rate  $R = 2$  designed with the high-resolution analysis.

By evaluating (12), the asymptotically-optimal point density for fixed-rate functional quantization is  $\lambda_{\text{fr}}^*(x) = \frac{7}{3} x^{4/3}$ . Integrating the point density gives the corresponding compander function  $w_{\text{fr}}^*(x) = x^{7/3}$ . As shown in the top panel of Fig. 3, the points are given by

$$\beta_i = w_{\text{fr}}^*{}^{-1}((2i - 1)/8), \quad i = 1, 2, 3, 4,$$

and the cell boundaries are given by  $w_{\text{fr}}^*{}^{-1}(\{0, 1/4, 1/2, 3/4, 1\})$ . The middle panel shows  $f_X$  and an approximation  $\hat{f}_X$  that is constant on each cell of the quantizer. The bottom panel shows  $g$  and the approximation  $g_{\text{PL}}$ , which is linear on each cell of the quantizer and tangent to  $g$  at each point.  $\square$

Referring to Fig. 3 for examples, the high-resolution distortion-resolution function  $d^{\text{HR}}(K; \lambda)$  can be thought of as a computation of the MSE of  $g_{\text{PL}}$  when the source with piecewise constant pdf  $\hat{f}_X$  is quantized with companding quantizer employing compander  $w_{\text{fr}}^*$ . In this case  $\hat{g}$ , the optimal function estimate, is given by evaluating  $g_{\text{PL}}$  at the center of the cell containing the source variable. Informally, as resolution  $K$  increases,  $\hat{f}_X \rightarrow f$ ,  $g_{\text{PL}} \rightarrow g$ , and the centers of the cells approach the corresponding quantizer points. These intuitions extend to multivariate functions as well, but our formal justifications in Section IV use techniques that do not explicitly form approximations  $\hat{f}_X$  or  $g_{\text{PL}}$ .

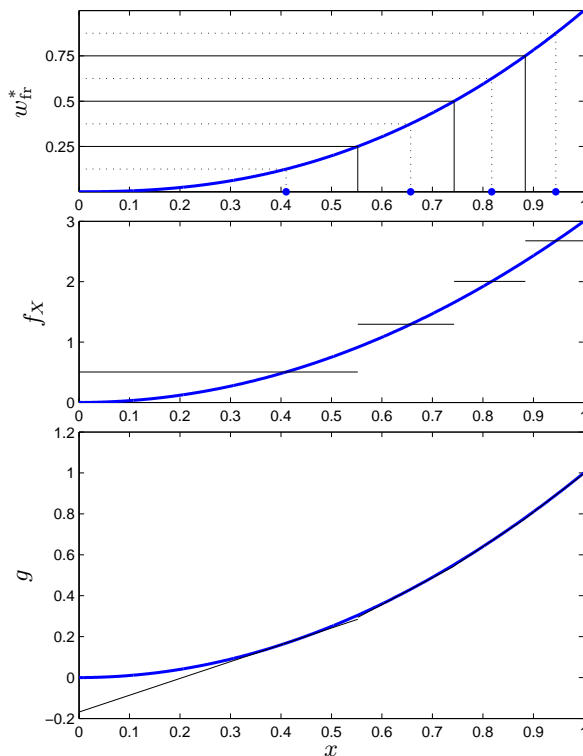


Fig. 3. Illustrations for Example 2. Top panel: points and cell boundaries of the quantizer are determined by the companding function  $w_{\text{fr}}^*$ . Middle panel: source pdf  $f_X$  and its piecewise constant approximation. Bottom panel: function  $g$  and its piecewise linear approximation.

#### F. Discontinuous Functions

Our main result on univariate functional quantization, Theorem 6, assumes the continuity of  $g$ . One can effectively sidestep this assumption, but doing so requires the quantizer to be described more precisely than by a point density function alone.

For simplicity, assume  $f_X$  is strictly positive on  $[0, 1]$ . Suppose we were to allow  $g$  to have a point of discontinuity  $x_0 \in (0, 1)$  with

$$c_0 = \lim_{\delta \rightarrow 0} |g(x_0 + \delta) - g(x_0 - \delta)| > 0.$$

The difficulty that arises is that if  $x_0$  is an interior point of a partition cell  $S_i$ , this cell produces a component of the functional distortion proportional to  $c_0^2 \mathbf{P}(X \in S_i)$ . Since  $c_0^2 \mathbf{P}(X \in S_i) = \Theta(K^{-1})$ , it is not negligible in comparison to the (best case)  $\Theta(K^{-2})$  functional distortion. Thus having a point of discontinuity of  $g$  in the interior of a partition cell disrupts the asymptotic distortion calculation (10).

The representation of quantizers by number of levels  $K$  and point density function  $\lambda$  cannot prevent a point of discontinuity from falling in the interior of a partition cell. However, if we augment the description of the quantizer with specified partition boundaries, we can still obtain the distortion estimate (10).

*Corollary 8:* Suppose a companding quantizer sequence for a source  $X \in [0, 1]$  is described by point density function  $\lambda(x)$ . Further suppose that the source, quantizer, and function  $g : [0, 1] \rightarrow \mathbb{R}$  satisfy Assumptions UF1–5 with the exception of discontinuities at  $M$  points  $\{x_m\}_{m=1}^M$ . Then a quantizer sequence obtained by adding partition cell boundaries at  $\{x_m\}_{m=1}^M$  will have distortion

$$d_g = \mathbf{E} \left[ (g(X) - g(\hat{X}))^2 \right] \sim \frac{1}{12K^2} \mathbf{E} \left[ (\gamma(X)/\lambda(X))^2 \right].$$

*Proof:* This follows from Theorem 6 applied separately to each of the subintervals where  $g$  is continuous. ■

In the sequel, we will not consider discontinuous functions. The multivariate extension of Corollary 8 requires points of discontinuity to be in the Cartesian product of finite sets of discontinuity for each variable. Such separable sets of points of discontinuity are not general and can be handled rather intuitively.

## IV. MULTIVARIATE FUNCTIONAL QUANTIZATION

With Section III as a warm-up, we may now establish the central results of distributed functional quantization.

### A. Definitions

An  $n$ -dimensional *distributed companding quantizer*  $Q_{\mathbf{K}}^{\mathbf{w}}$  is specified by  $n$  companding functions  $\mathbf{w} = (w_1, w_2, \dots, w_n)$  and an  $n$ -vector of resolutions  $\mathbf{K} = (K_1, K_2, \dots, K_n)$ . When applied to an  $n$ -tuple  $x_1^n \in [0, 1]^n$ ,  $Q_{\mathbf{K}}^{\mathbf{w}}$  quantizes each component  $x_j$  of  $x_1^n$  separately with compander  $w_j$  and resolution  $K_j$ :

$$Q_{\mathbf{K}}^{\mathbf{w}}(x_1^n) = (Q_{K_1}^{w_1}(x_1), Q_{K_2}^{w_2}(x_2), \dots, Q_{K_n}^{w_n}(x_n)).$$

A distributed companding quantizer may equivalently be specified by  $n$  point density functions  $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \dots, \lambda_n)$ , in which case it is denoted by  $Q_{\mathbf{K}}^{\boldsymbol{\lambda}}$ .

An estimation function  $\hat{g} : [0, 1]^n \rightarrow \mathbb{R}$  estimates the value of  $g(X_1^n)$  from the quantized representation  $Q_{\mathbf{K}}^{\boldsymbol{\lambda}}(X_1^n)$ . The distortion of a distributed quantizer paired with an estimator  $\hat{g}$  is given by the distortion-resolution function

$$d_{\hat{g}}(\mathbf{K}; \boldsymbol{\lambda}) = \mathbf{E}[|g(X_1^n) - \hat{g}(Q_{\mathbf{K}}^{\boldsymbol{\lambda}}(X_1^n))|^2].$$

In this paper, use of the optimal estimator

$$\hat{g}(x_1^n) = \mathbf{E}[g(X_1^n) \mid Q_{\mathbf{K}}^{\boldsymbol{\lambda}}(X_1^n) = Q_{\mathbf{K}}^{\boldsymbol{\lambda}}(x_1^n)]$$

will be indicated by omitting the subscript:  $d(\mathbf{K}; \boldsymbol{\lambda})$ .

The *rate*  $R$  of a distributed quantizer takes on three different meanings. A *fixed-rate* constraint limits the total resolution  $K = \prod_{j=1}^n K_j \leq 2^R$ , and we assume that the  $j$ th quantizer communicates to the decoder with rate  $R_j = \log K_j$ . A *variable-rate* (marginal entropy) constraint limits the sum of the marginal entropies  $\sum_{j=1}^n H(Q_{K_j}^{\lambda_j}(X_j)) \leq R$ , and we assume that the  $j$ th quantizer utilizes entropy-coding to the decoder to attain rate  $R_j = H(Q_{K_j}^{\lambda_j}(X_j))$ . A *Slepian–Wolf* (joint entropy) constraint limits the joint entropy  $H(Q_{\mathbf{K}}^{\boldsymbol{\lambda}}(X_1^n)) \leq R$ , and we assume that the  $j$ th quantizer utilizes Slepian–Wolf coding to the decoder to attain rate  $R_j = H(Q_{K_j}^{\lambda_j}(X_j) \mid Q_{K^{j-1}}^{\lambda^{j-1}}(X^{j-1}))$ , where  $Q_{K^{j-1}}^{\lambda^{j-1}}(X^{j-1})$  is used to represent  $(Q_{K_1}^{\lambda_1}(X_1), \dots, Q_{K_{j-1}}^{\lambda_{j-1}}(X_{j-1}))$ . Note that the choice of this particular point on the Slepian–Wolf rate boundary is arbitrary. The resulting performance is measured by the distortion-rate functions

$$D_{\text{fr}}(R; \boldsymbol{\lambda}) = \min_{\mathbf{K}: \prod_{j=1}^n K_j \leq 2^R} d(\mathbf{K}; \boldsymbol{\lambda}),$$

$$D_{\text{vr}}(R; \boldsymbol{\lambda}) = \min_{\mathbf{K}: \sum_{j=1}^n H(Q_{K_j}^{\lambda_j}(X_j)) \leq R} d(\mathbf{K}; \boldsymbol{\lambda}),$$

and

$$D_{\text{sw}}(R; \boldsymbol{\lambda}) = \min_{\mathbf{K}: \sum_{j=1}^n H(Q_{K_j}^{\lambda_j}(X_j) \mid Q_{K^{j-1}}^{\lambda^{j-1}}(X^{j-1})) \leq R} d(\mathbf{K}; \boldsymbol{\lambda}).$$

A quantizer point density  $\boldsymbol{\lambda}^*$  is *asymptotically better* than another  $\boldsymbol{\lambda}$  under a fixed-rate, variable-rate, or Slepian–Wolf constraint if the ratio of the distortion-rate functions is at most one:

$$\lim_{R \rightarrow \infty} \frac{D_{\text{fr}}(R; \boldsymbol{\lambda}^*)}{D_{\text{fr}}(R; \boldsymbol{\lambda})} \leq 1, \quad (16a)$$

$$\lim_{R \rightarrow \infty} \frac{D_{\text{vr}}(R; \boldsymbol{\lambda}^*)}{D_{\text{vr}}(R; \boldsymbol{\lambda})} \leq 1, \text{ or} \quad (16b)$$

$$\lim_{R \rightarrow \infty} \frac{D_{\text{sw}}(R; \boldsymbol{\lambda}^*)}{D_{\text{sw}}(R; \boldsymbol{\lambda})} \leq 1. \quad (16c)$$

If  $\boldsymbol{\lambda}$  is asymptotically better than any other distributed quantizer sequence, it is *asymptotically optimal*.

### B. Problem Statement

Let  $X_1^n$  be a random vector with joint pdf  $f_{X_1^n}(x_1^n)$  defined over  $[0, 1]^n$ , and let  $g : [0, 1]^n \rightarrow \mathbb{R}$  be the function of interest. A distributed companding quantizer  $\{Q_{\mathbf{K}}^{\boldsymbol{\lambda}}\}$  is applied to  $X_1^n$ . Equivalently, a companding quantizer  $Q_{K_j}^{\lambda_j}$  is applied to each component of the source  $X_j$ . The decoder then forms an estimate  $\hat{g}(Q_{\mathbf{K}}^{\boldsymbol{\lambda}}(X_1^n))$ , where  $\hat{g}(Q_{\mathbf{K}}^{\boldsymbol{\lambda}}(X_1^n)) = \mathbf{E}[g(X_1^n) \mid Q_{\mathbf{K}}^{\boldsymbol{\lambda}}(X_1^n)]$  is the optimal estimation function. Distortion is measured by squared error in the function  $D = \mathbf{E}[(g(X_1^n) - \hat{g}(Q_{\mathbf{K}}^{\boldsymbol{\lambda}}(X_1^n)))^2]$ , which for the optimal estimator reduces to  $D = \mathbf{E}[\text{var}(g(X_1^n) \mid Q_{\mathbf{K}}^{\boldsymbol{\lambda}}(X_1^n))]$ . Fig. 1 depicts this scenario, with  $R_j = \log K_j$  in the fixed-rate case,  $R_j = H(Q_{K_j}^{\lambda_j}(X_j))$  in the variable-rate case, and  $R_j = H(Q_{K_j}^{\lambda_j}(X_j) \mid Q_{K^{j-1}}^{\lambda^{j-1}}(X^{j-1}))$  in the Slepian–Wolf case. We wish to choose  $\boldsymbol{\lambda}$  to be asymptotically optimal.

As in Section III, we will impose restrictions on the function  $g$  and the joint probability distribution function of  $X_1^n$  so that a local affine approximation is effective.

MF1.  $g$  is Lipschitz continuous, and the first and second derivatives of  $g$  are defined except possibly on a set of zero Jordan measure.

MF2. The source pdf  $f$  is continuous and supported on  $[0, 1]^n$ , and is therefore bounded.

MF3. We optimize among companding functions  $w_j$  that are piecewise differentiable (and therefore a point density description  $\lambda_j$  is appropriate).

MF4. Letting  $g_j(x_1^n)$  denote  $\partial g(x_1^n)/\partial x_j$ , the integrals

$$\int_0^1 f(x_j) \mathbf{E} \left[ |g_j(X^n)|^2 \mid X_j = x_j \right] \lambda_j(x_j)^{-2} dx_j$$

are defined and positive for all  $j \in \{1, 2, \dots, n\}$ .

Constraints MF1–MF4 are more restrictive than they need to be, but this helps in simplifying proofs. For instance, condition MF4 guarantees that every source variable must be finely quantized for distortion to approach zero. If this is violated for the  $j$ th source variable, it merely implies that a finite-resolution quantization of  $X_j$  suffices.

Note that there is no analogue to the monotonicity assumption UF1 in the multivariate case. It can be shown that if  $g$  is monotonic in each of its variables the optimal fixed-rate distributed quantizer is regular. With the added restriction that the source variables be independent, it can be shown that the optimal variable-rate distributed quantizer is also regular, via techniques similar to those of [26]. Rather than constraining the function  $g$  and the source pdf  $f$  in this manner, however, assumption MF3 explicitly restricts optimization to the space of regular companding quantizer sequences, regardless of whether regularity is optimal. In Sec. VI it is shown that nonregular companding quantizer sequences are asymptotically suboptimal for a wide variety of functions  $g$ , giving this constraint some validity.

### C. High-Resolution Analysis

1) *The Distortion-Resolution Function*: Our main technical task in finding the optimal quantizers is to justify an approximation of the distortion in terms of point density functions. Since the quantization is distributed, our concept of functional sensitivity is now extended to each variable separately, with averaging performed over the remaining variables.

*Definition 3*: The  $j$ th functional sensitivity profile of  $g$  is defined as

$$\gamma_j(x) = \left( \mathbf{E} \left[ |g_j(X_1^n)|^2 \mid X_j = x \right] \right)^{1/2}. \quad (17)$$

*Theorem 9*: Suppose  $n$  sources  $X_1^n \in [0, 1]^n$  are quantized by a distributed companding quantizer  $Q_{\mathbf{K}}^\lambda$ , and suppose that the source, quantizers, and function  $g : [0, 1]^n \rightarrow \mathbb{R}$  satisfy assumptions MF1–4. Let  $d(\mathbf{K}; \lambda) = \mathbf{E} [\text{var}(g(X_1^n) \mid Q_{\mathbf{K}}^\lambda(X_1^n))]$  denote the true distortion-resolution function, and let  $d^{\text{HR}}$  denote the high-resolution approximate distortion-resolution function:

$$d^{\text{HR}}(\mathbf{K}; \lambda) = \sum_{j=1}^n \frac{1}{12K_j^2} \mathbf{E} \left[ \left( \frac{\gamma_j(X_j)}{\lambda_j(X_j)} \right)^2 \right]. \quad (18)$$

Then  $d(\mathbf{K}; \lambda) \sim d^{\text{HR}}(\mathbf{K}; \lambda)$ , where  $\sim$  indicates that the ratio of the two quantities approaches one as the smallest element of the vector  $\mathbf{K}$  grows without bound.

*Proof*: See Appendix A. ■

2) *Connecting Resolution to Rate*: To convert the distortion-resolution function to a distortion-rate function, we first introduce a slight generalization of the high-resolution resolution-rate relationship.

*Lemma 10*: If the source  $X_1^n$  has a density over  $[0, 1]^n$  with finite differential entropy  $h(X_1^n)$  and if  $\mathbf{E} [\log \lambda_j(X_j)]$  is finite for all  $j \in \{1, \dots, n\}$ , then as each component of the resolution vector  $\mathbf{K}$  diverges,

$$H(Q_{\mathbf{K}}^\lambda(X_1^n)) - \sum_{j=1}^n \log K_j \rightarrow h(X_1^n) + \sum_{j=1}^n \mathbf{E} [\log \lambda_j(X_j)].$$

*Proof*: Suppose  $W = (w_1(x_1), w_2(x_2), \dots, w_n(x_n))$  is an  $n$ -dimensional companding function that is applied to the source  $X_1^n$  prior to quantization by a rectangular lattice quantizer  $Q^U$  with side length  $K_j^{-1}$  on the  $j$ th side, and furthermore suppose  $W^{-1}$  is then applied to estimate the source. The output of this quantization process  $W^{-1}(Q^U(W(X_1^n)))$  is identical to the scenario we consider, and since  $W^{-1}$  is one-to-one, the joint discrete entropy of the outputs are identical as well:  $H(Q^U(W(X_1^n))) = H(Q_{\mathbf{K}}^\lambda(X_1^n))$ .

Since the volume of each cell of the rectangular lattice  $Q^U$  is equal to  $K^{-1}$ , and since the diameter of each cell falls to zero, a special case of a result by Csiszár [32], [33] tells us that

$$\lim_{K \rightarrow \infty} H(Q^U(W(X_1^n))) - \log K = h(W(X_1^n)).$$

Since the differential entropy of a continuously differentiable function of  $X$  is given by  $h(f(X)) = h(X) + \mathbf{E} [\log \det J_f(X)]$ , where  $J_f(X)$  denotes the Jacobian matrix for the function  $f$ , we may reduce the expression to

$$\lim_{K \rightarrow \infty} H(Q^U(W(X_1^n))) - \log K = h(X_1^n) + \sum_{j=1}^n \mathbf{E} [\log \lambda_j(X_j)].$$

Recalling that  $H(Q^U(W(X_1^n))) = H(Q_K^\lambda(X_1^n))$ , the proof is complete.  $\blacksquare$

Armed with this, the distortion-resolution function may be modified to include considerations of rate.

*Lemma 11:* Define the fixed-rate, variable-rate, and Slepian-Wolf distortion-resolution functions as

$$\begin{aligned} d_{\text{fr}}^{\text{HR}}(\mathbf{K}; \boldsymbol{\lambda}) &= \sum_{j=1}^n \frac{1}{12K_j^2} \mathbf{E} \left[ \left( \frac{\gamma_j(X_j)}{\lambda_j(X_j)} \right)^2 \right], \\ d_{\text{vr}}^{\text{HR}}(\mathbf{K}; \boldsymbol{\lambda}) &= \sum_{j=1}^n \frac{1}{12} 2^{-2H(Q_{K_j}^{\lambda_j}(X_j)) + 2h(X_j) + 2\mathbf{E}[\log \lambda_j(X_j)]} \mathbf{E} \left[ \left( \frac{\gamma_j(X_j)}{\lambda_j(X_j)} \right)^2 \right], \\ d_{\text{sw}}^{\text{HR}}(\mathbf{K}; \boldsymbol{\lambda}) &= \sum_{j=1}^n \frac{1}{12} 2^{-2H(Q_{K_j}^{\lambda_j}(X_j) | Q_{K^{j-1}}^{\lambda^{j-1}}(X^{j-1})) + 2h(X_j | X^{j-1}) + 2\mathbf{E}[\log \lambda_j(X_j)]} \mathbf{E} \left[ \left( \frac{\gamma_j(X_j)}{\lambda_j(X_j)} \right)^2 \right]. \end{aligned}$$

Then  $d(\mathbf{K}; \boldsymbol{\lambda}) \sim d_{\text{fr, vr, sw}}^{\text{HR}}(\mathbf{K}; \boldsymbol{\lambda})$ .

*Proof:* By Theorem 9,  $d(\mathbf{K}; \boldsymbol{\lambda}) \sim d_{\text{fr}}^{\text{HR}}(\mathbf{K}; \boldsymbol{\lambda})$ . This establishes the first of the asymptotic equalities.

For the second (variable-rate) asymptotic equality, we observe that by Lemma 1,

$$K_j \sim 2^{-2H(Q_{K_j}^{\lambda_j}(X_j)) + 2h(X_j) + 2\mathbf{E}[\log \lambda_j(X_j)]}$$

and therefore that  $d_{\text{vr}}^{\text{HR}}(\mathbf{K}; \boldsymbol{\lambda}) \sim d^{\text{HR}}(\mathbf{K}; \boldsymbol{\lambda})$ . Again, by Theorem 9,  $d(\mathbf{K}; \boldsymbol{\lambda}) \sim d^{\text{HR}}(\mathbf{K}; \boldsymbol{\lambda})$ .

For the third (Slepian-Wolf) asymptotic equality, we start by noting that by Lemma 10,

$$\prod_{i=1}^j K_i \sim 2^{-2H(Q_{K^j}^{\lambda^j}(X^j)) + 2h(X^j) + 2\sum_{i=1}^j \mathbf{E}[\log \lambda^i(X^i)]},$$

and similarly

$$\prod_{i=1}^{j-1} K_i \sim 2^{-2H(Q_{K^{j-1}}^{\lambda^{j-1}}(X^{j-1})) + 2h(X^{j-1}) + 2\sum_{i=1}^{j-1} \mathbf{E}[\log \lambda^{i-1}(X^{i-1})]}.$$

Dividing the first by the second yields that

$$K_j \sim 2^{-2H(Q_{K_j}^{\lambda_j}(X_j) | Q_{K^{j-1}}^{\lambda^{j-1}}(X^{j-1})) + 2h(X_j | X^{j-1}) + 2\mathbf{E}[\log \lambda_j(X_j)]}$$

and therefore that  $d_{\text{sw}}^{\text{HR}}(\mathbf{K}; \boldsymbol{\lambda}) \sim d^{\text{HR}}(\mathbf{K}; \boldsymbol{\lambda}) \sim d(\mathbf{K}; \boldsymbol{\lambda})$ .  $\blacksquare$

3) *The Distortion-Rate Functions:* We may now establish high-resolution approximations to the distortion-rate function under each of the three rate constraints.

*Lemma 12:* Define the fixed-rate, variable-rate, and Slepian-Wolf high-resolution distortion-rate functions as

$$D_{\text{fr}}^{\text{HR}}(R; \boldsymbol{\lambda}) = \frac{n}{12} 2^{-2R/n} \left( \prod_{j=1}^n \mathbf{E} \left[ \left( \frac{\gamma_j(X_j)}{\lambda_j(X_j)} \right)^2 \right] \right)^{1/n}, \quad (19a)$$

$$D_{\text{vr}}^{\text{HR}}(R; \boldsymbol{\lambda}) = \frac{n}{12} 2^{-2R/n} \left( \prod_{j=1}^n 2^{2h(X_j) + 2\mathbf{E}[\log \lambda_j(X_j)]} \mathbf{E} \left[ \left( \frac{\gamma_j(X_j)}{\lambda_j(X_j)} \right)^2 \right] \right)^{1/n}, \quad (19b)$$

$$D_{\text{sw}}^{\text{HR}}(R; \boldsymbol{\lambda}) = \frac{n}{12} 2^{-2R/n} \left( 2^{2h(X_1^n)} \prod_{j=1}^n 2^{2\mathbf{E}[\log \lambda_j(X_j)]} \mathbf{E} \left[ \left( \frac{\gamma_j(X_j)}{\lambda_j(X_j)} \right)^2 \right] \right)^{1/n}. \quad (19c)$$

Then  $D_{\text{fr, vr, sw}}^{\text{HR}}(R; \boldsymbol{\lambda}) \sim D_{\text{fr, vr, sw}}^{\text{HR}}(R; \boldsymbol{\lambda})$ .

*Proof:* See Appendix B.  $\blacksquare$

4) *Asymptotically Optimal Distributed Quantizers:* The expressions (19) decouple the problem of designing  $n$  point densities  $\boldsymbol{\lambda}$  into  $n$  separate problems of designing a single point density  $\lambda_j$ . Furthermore, each design problem (the minimization of an expression in (19)) is of a familiar form. Thus we obtain the following theorem.

*Theorem 13:* The asymptotic fixed-rate (codebook-constrained) distortion-rate expression (19a) is minimized by the choice

$$\lambda_j^*(x) = \frac{(\gamma_j^2(x) f_{X_j}(x))^{1/3}}{\int_0^1 (\gamma_j^2(t) f_{X_j}(t))^{1/3} dt}, \quad j = 1, 2, \dots, n, \quad (20)$$

yielding distortion

$$D_{\text{fr}}^{\text{HR}}(R) = \frac{n}{12} \left( \prod_{j=1}^n \|\gamma_j^2 f_{X_j}\|_{1/3} \right)^{1/n} 2^{-2R/n}. \quad (21)$$

The asymptotic variable-rate (marginal entropy-constrained) distortion-rate expression (19b) is minimized by the choice

$$\lambda_j^*(x) = \frac{\gamma_j(x)}{\int_0^1 \gamma_j(t) dt}, \quad j = 1, 2, \dots, n, \quad (22)$$

yielding distortion

$$D_{\text{vr}}^{\text{HR}}(R) = \frac{n}{12} \left( \prod_{j=1}^n 2^{2h(X_j) + 2\mathbf{E}[\log \gamma_j(X_j)]} \right)^{1/n} 2^{-2R/n}. \quad (23)$$

The asymptotic Slepian–Wolf (joint entropy-constrained) distortion-rate expression (19c) is optimized by a choice of point densities identical to the variable-rate case (22). The resulting distortion is

$$D_{\text{sw}}^{\text{HR}}(R) = \frac{n}{12} \left( 2^{2h(X_1^n)} \prod_{j=1}^n 2^{2\mathbf{E}[\log \gamma_j(X_j)]} \right)^{1/n} 2^{-2R/n}. \quad (24)$$

The distributed quantizer point densities yielded by the above optimizations are asymptotically optimal.

*Proof:* To prove (20) gives the optimal point density for fixed-rate coding and (22) gives the optimal point density for both variable-rate and Slepian–Wolf coding, it suffices to note that minimizing the  $n$  terms in (19a), (19b), and (19c) separately gives problems identical to those in Section III.

The proof that the choice of  $\lambda$  that minimizes the high-resolution expression is asymptotically optimal is virtually identical to that of Lemma 3, so it is omitted. ■

#### D. Variation: Joint Entropy Constraint

Distortion expressions (21) and (23) are minimum distortions subject to a sum-rate constraint. The individual rates given by  $R_j = \log K_j$  (fixed-rate) or by (6) (variable-rate) implicitly specify no entropy coding or separate entropy coding of the  $\hat{X}_j$ s, respectively.

If the  $\hat{X}_j$ s are not independent—which is anticipated whenever the  $X_j$ s are not independent—one may employ Slepian–Wolf coding of the  $\hat{X}_j$ s without violating the distributed coding requirement implicit in Fig. 1. This lowers the total rate from  $\sum_{j=1}^n H(\hat{X}_j)$  to  $H(\hat{X}_1, \hat{X}_2, \dots, \hat{X}_n)$  and changes the marginal entropy constraint into a joint entropy constraint. While the optimal compander choice (22) is unchanged by this modification, the resulting distortion-rate function reduces from (23) to (24).

Some remarks:

- 1) By comparing (24) to (23), we see that the inclusion of Slepian–Wolf coding has reduced the sum rate to achieve any given distortion by

$$\left( \sum_{j=1}^n h(X_j) \right) - h(X_1^n).$$

This is, of course, not unexpected as it represents the excess information in the product of marginal probability distributions as compared to the joint probability distribution. This has been termed the *multiinformation* [34] and equals the mutual information when  $n = 2$ .

- 2) While the resolution allocation  $\mathbf{K}$  amongst the  $n$  sources has a unique minimizing choice, there is some flexibility in rate allocations for the Slepian–Wolf encoder. Any point on the Slepian–Wolf joint-entropy boundary may be achieved with arbitrarily low probability of error.
- 3) The theorem seems to analytically separate correlations among sources from functional considerations, exploiting correlation even though the quantizers are regular. In reality, the binning introduced by Slepian–Wolf coding transforms the scalar quantizers of each source component into nonregular vector quantizers so as to remove redundancy between sources.



### E. Relationship to Locally-Quadratic Distortion Measures

Linder *et al.* consider the class of “locally-quadratic” distortion measures for variable-rate high-resolution quantization in [23]. They define locally-quadratic measures as those having the following two properties:

- 1) Let  $x$  be in  $\mathbb{R}^n$ . For  $y$  sufficiently close to  $x$  in the Euclidean metric, the distortion between  $x$  and  $y$  is well approximated by  $\sum_{i=1}^n M_i(x)|x_i - y_i|^2$ , where  $M_i(x)$  is a positive scaling factor. In other words, the distortion is a space-varying non-isotropically scaled MSE.
- 2) The distortion between two points is zero if and only if the points are identical.

For these distortion measures, the authors consider high-resolution variable-rate regular quantization, generalize Bucklew’s results [24] to non-functional distortion measures, and demonstrate the use of multidimensional companding functions to implement these quantizers. Of particular interest is the comparison they perform between joint vector quantization and separable scalar quantization. When Slepian–Wolf coding is employed for the latter, the scenario is similar to the developments of this section.

The source of this similarity is the implicit distortion measure we work with:  $d_g(x, y) = |g(x) - g(y)|^2$ . When  $x$  and  $y$  are very close to each other, Taylor approximation reduces this error to a quadratic form:

$$|g(x) - g(y)|^2 \approx \sum_{i=1}^n \left| \frac{\partial g(x_1^n)}{\partial x_i} \right|^2 |x_i - y_i|^2.$$

From this, one may obtain the same variable-rate Slepian–Wolf performance as (24) through the analysis in [23].

However, there are important differences between locally-quadratic distortion measures and the functional distortion measures we consider. First and foremost: a continuous scalar function of  $n$  variables,  $n > 1$ , is *guaranteed* to have an uncountable number of pairs  $x \neq y$  for which  $g(x) = g(y)$  and therefore that  $d_g(x, y) = 0$ . This violates the second condition of a locally-quadratic distortion measure, and the repercussions are felt most strikingly for non-monotonic functions—those for which regular quantizers are not necessarily optimal (see Section VI).

The second condition is also violated by functions that are not *strictly* monotonic in each variable; one finds that without strictness, variable-rate analysis of the centralized encoding problem is invalidated. Specifically, if the derivative vector

$$\left( \frac{\partial g(x_1^n)}{\partial x_1}, \frac{\partial g(x_1^n)}{\partial x_2}, \dots, \frac{\partial g(x_1^n)}{\partial x_n} \right)$$

has nonzero probability of possessing a zero component, the expected variable-rate distortion as derived by both Bucklew and Linder *et al.* is  $D = 0$ , regardless of rate. This answer arrives from the null derivative having violated the high-resolution approximation, and it implies that the distortion falls faster than  $2^{-2R/n}$ . In future work, generalizations of our results in Section VII may be able to address such deficiencies.

## V. EXAMPLES

Before moving on to extensions of the basic theory, we present a few examples to show how optimal ordinary scalar quantization and optimal DFSQ differ. We especially want to highlight a few simple examples in which performance scaling with respect to  $n$  differ greatly between ordinary and functionally-optimized quantization. To draw attention to this scaling, we define the *rate-per-source*  $\bar{R}$  as the sum-rate divided by the number of sources  $R/n$ , and hold this quantity constant as the number of sources increases.

*Example 3 (Linear function):* Consider the function  $g(x_1^n) = \sum_{j=1}^n a_j x_j$  where the  $a_j$ s are scalars. Then for any  $j$ ,  $\gamma_j(x) = |a_j|$ . Since  $\gamma_j(x)$  does not depend on  $x$ , it has no influence on the optimal point density for either the fixed- or variable-rate case; see (20) and (22).

Although  $\gamma_j(x)$  gives no information on which values of  $X_j$  are more important than others (or rather shows that they are all equally important) the set of  $\gamma_j$ s shows the relative importance of the components. This is reflected in the allocation of rate.  $\square$

*Example 4 (Maximum):* Let the set of sources  $X_1^n$  be uniformly distributed on  $[0, 1]^n$  and hence mutually independent. Consider the function

$$g(x_1^n) = \max(x_1, x_2, \dots, x_n).$$

Note that this function is differentiable outside the sets  $A_{i,j} = \{x_1^n : x_i = x_j\}$ , where  $i, j \in \{1, \dots, n\}$ . Each  $A_{i,j}$  is an  $(n - 1)$ -dimensional plane and therefore has Jordan measure zero, and since a finite union of Jordan-measure-zero sets has Jordan measure zero, condition MF1 is satisfied. Though very simple, this function is more interesting than a linear function because the derivative with respect to one variable depends sharply on all the others. The function is symmetric in its arguments, so for notational convenience consider only the design of the quantizer for  $X_1$ .

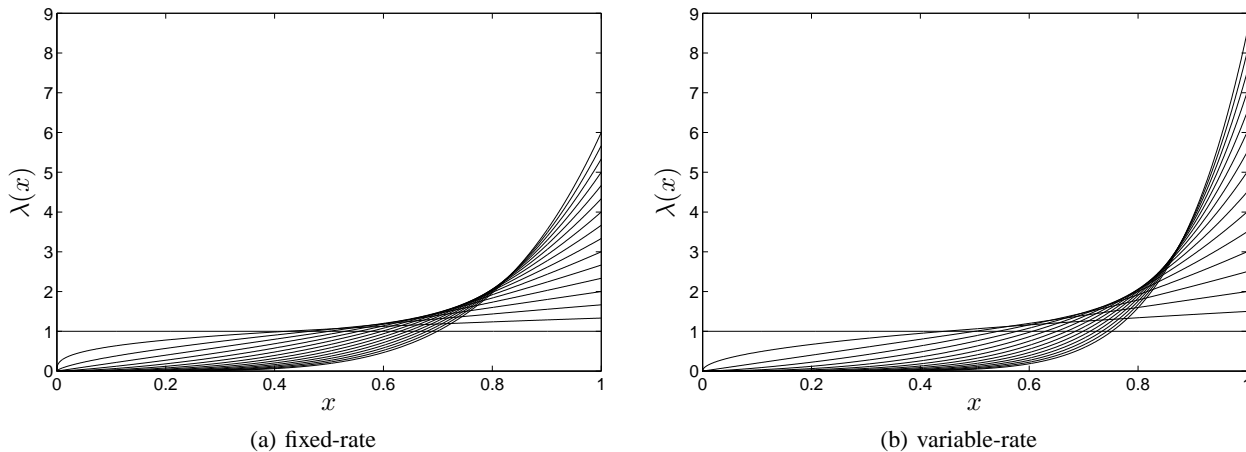


Fig. 4. Optimal point densities for Example 4 (maximum),  $n = 1, 2, \dots, 16$ . As  $n$  increases, the sensitivities  $\gamma_j(x)$  become more unbalanced toward large  $x$ ; this is reflected in the point densities, more so in the variable-rate case than in the fixed-rate case.

The partial derivative  $g_1(x_1^n)$  is 1 where the maximum is  $x_1$  and is 0 otherwise. Thus,

$$\begin{aligned} \gamma_1^2(x) &= \mathbf{E} [|g_1(X_1^n)|^2 | X_1 = x] \\ &= \mathbf{P}(\max(X_1^n) = X_1 | X_1 = x) \\ &= x^{n-1}, \end{aligned}$$

where the final step uses the probability of all  $n-1$  variables  $X_2^n$  being less than  $x$ .

The optimal point density for fixed-rate quantization is found by evaluating (20) to be

$$\lambda_1(x) = \frac{1}{3}(n+2)x^{(n-1)/3}.$$

The resulting distortion when each quantizer has rate  $\bar{R}$  (equal rate allocations) is found by evaluating (21) to be

$$\begin{aligned} D_{\text{fr}}^{\text{HR}}(n\bar{R}) &= \frac{n}{12} \|\gamma_1^2\|_{1/3} 2^{-2\bar{R}} = \frac{n}{12} \left(\frac{3}{n+2}\right)^3 2^{-2\bar{R}} \\ &= \frac{9n}{4(n+2)^3} 2^{-2\bar{R}}. \end{aligned}$$

The optimal point density for variable-rate quantization is found by evaluating (22) to be

$$\lambda_1(x) = \frac{1}{2}(n+1)x^{(n-1)/2}.$$

Substituting  $h(X_1) = 0$  and  $2^{2\mathbf{E}[\log \gamma_1(X_1)]} = e^{-n+1}$  into (23) gives

$$D_{\text{vr}}^{\text{HR}}(n\bar{R}) = \frac{n}{12} e^{-n+1} 2^{-2\bar{R}}.$$

The two computed distortions decrease sharply with  $n$ . This is in stark contrast to the results of ordinary quantization. When functional considerations are ignored, one optimally uses a uniform quantizer, resulting in  $\mathbf{E}[(X_j - \hat{X}_j)^2] \approx \frac{1}{12} 2^{-2R_j}$  for any component. Since the maximum is equal to one of the components, the functional distortion is  $D_{\text{ord}}^{\text{HR}}(n\bar{R}) = \frac{1}{12} 2^{-2\bar{R}}$ , unchanging with  $n$ .

The optimal point densities computed above are shown in Fig. 4. The distortions are presented along with the results of the following example in Fig. 5.  $\square$

*Example 5 (Median):* Let  $n = 2m + 1$ ,  $m \in \mathbb{N}$ , and again let the set of sources  $X_1^n$  be uniformly distributed on  $[0, 1]^n$ . The function

$$g(x_1^n) = \text{median}(x_1, x_2, \dots, x_n)$$

provides a similar but more complicated example. Note that, as in Example 4, this function is differentiable outside the zero-Jordan-measure sets  $A_{i,j}$ , and it therefore satisfies condition MF1.

The partial derivative  $g_1(x_1^n)$  is 1 where the median is  $x_1$  and is 0 otherwise. Thus,

$$\begin{aligned} \gamma_1^2(x) &= \mathbf{E} [|g_1(X_1^n)|^2 | X_1 = x] \\ &= \mathbf{P}(\text{median}(X_1^n) = X_1 | X_1 = x) \\ &= \binom{2m}{m} x^m (1-x)^m, \end{aligned}$$

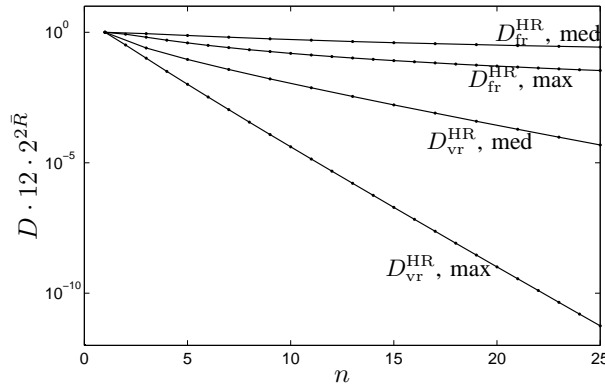


Fig. 5. Distortions of optimal fixed- and variable-rate functional quantizers for maximum and median functions from Examples 4 and 5. Shown is the dependence on the number of variables  $n$ ; by plotting  $D \cdot 12 \cdot 2^{2\bar{R}}$  we see the performance relative to ordinary quantization.

where the final step uses the binomial probability for the event of exactly  $m$  of the  $2m$  variables  $X_2^n$  exceeding  $x$ .

The optimal point density for fixed-rate quantization is found by evaluating (20) to be

$$\lambda_1(x) = \frac{x^{m/3}(1-x)^{m/3}}{B(m/3+1, m/3+1)}$$

where  $B$  is the beta function. The resulting distortion when each quantizer has rate  $\bar{R}$  is found by evaluating (21) to be

$$\begin{aligned} D_{\text{fr}}^{\text{HR}}(n\bar{R}) &= \frac{2m+1}{12} \|\gamma_1^2\|_{1/3} 2^{-2\bar{R}} \\ &= \frac{2m+1}{12} \binom{2m}{m} \left( B\left(\frac{m}{3}+1, \frac{m}{3}+1\right) \right)^3 2^{-2\bar{R}}. \end{aligned}$$

To understand the trend for large  $m$ , we can substitute in the Stirling approximations  $\binom{2m}{m} \sim (m\pi)^{-1/2} 2^{2m}$  and

$$B(m/3+1, m/3+1) \sim \sqrt{6\pi/m} 2^{-(2m/3+3/2)}$$

to obtain

$$D_{\text{fr}}^{\text{HR}}(n\bar{R}) \sim \frac{m}{6} \frac{2^{2m}}{\sqrt{m\pi}} \left(\frac{6\pi}{m}\right)^{3/2} 2^{-(2m+9/2)} 2^{-2\bar{R}} = \frac{\pi\sqrt{3}}{16m} 2^{-2\bar{R}}.$$

The optimal point density for variable-rate quantization is found by evaluating (22) to be

$$\lambda_1(x) = \frac{x^m(1-x)^m}{B(m+1, m+1)}.$$

To evaluate the resulting distortion, note that  $h(X_1) = 0$  and  $2^{2\mathbf{E}[\log \gamma_1(X_1)]} = \binom{2m}{m} e^{-2m}$ . Substituting into (23) gives

$$D_{\text{vr}}^{\text{HR}}(n\bar{R}) = \frac{2m+1}{12} \binom{2m}{m} e^{-2m} 2^{-2\bar{R}}.$$

Using the approximation above for the binomial factor we obtain

$$D_{\text{vr}}^{\text{HR}}(n\bar{R}) \sim \frac{m^{1/2}}{6\pi^{1/2}} \left(\frac{e}{2}\right)^{-2m} 2^{-2\bar{R}}.$$

The optimal point densities computed above are shown in Fig. 6. The distortions are presented along with the results of Example 4 in Fig. 5.

Note the following similarities to Example 4:  $D_{\text{ord}}^{\text{HR}}$  is constant with respect to  $n$ ,  $D_{\text{fr}}^{\text{HR}}$  decays polynomially with  $n$ , and  $D_{\text{vr}}^{\text{HR}}$  decays exponentially with  $n$ .  $\square$

The large performance improvement over ordinary quantization in these examples illustrates the potential benefits of functional quantization. Additional examples and details appear in [35].

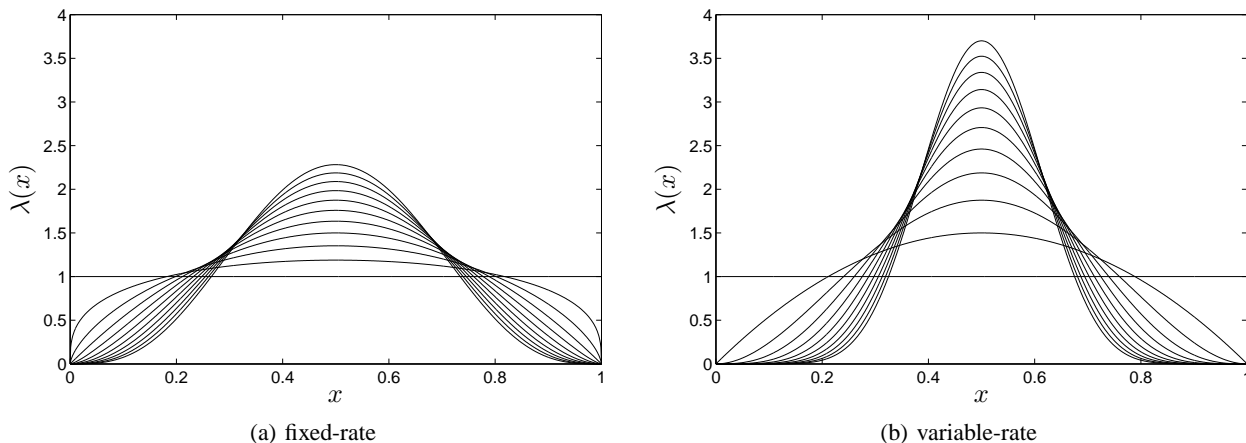


Fig. 6. Optimal point densities for Example 5 (median),  $n = 1, 3, \dots, 21$ . As  $n$  increases, the sensitivities  $\gamma_j(x)$  become more unbalanced toward  $x = 1/2$ ; this is reflected in the point densities, more so in the variable-rate case than in the fixed-rate case.

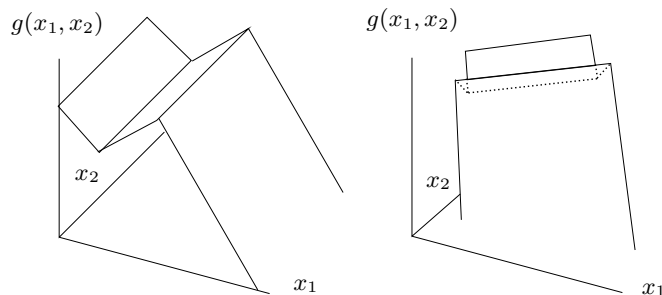


Fig. 7. Two functions of two variables are shown. The left function is separable and  $X_1$  is best quantized by a non-regular quantizer; for the right function (a rotated version of the left), a regular quantizer is asymptotically optimal. This is due to the right function being “equivalence-free.”

## VI. NON-MONOTONIC FUNCTIONS AND NON-REGULAR QUANTIZATION

The high-resolution approach to quantizer optimization is inherently limited to the design of regular quantizers. In particular, we have specified compander functions to be monotonic in Section II-A. The analysis of Section IV therefore gave us quantizer sequences within the class of regular quantizers.

In this section we explore less restrictive alternatives to the monotonicity requirement. Specifically, we introduce the concept of *equivalence-free* and show that if a function has this property, then non-regular companding quantizer sequences are asymptotically suboptimal.

Fig. 7 illustrates the concept. The function on the left is aligned with the axes in the sense that  $g(x_1, x_2)$  depends only on  $x_1$ . Since the dependence on  $x_1$  is not monotonic, there are pairs of distinct points  $(x_1^\dagger, x_1^\ddagger)$  where  $g(x_1^\dagger, x_2) = g(x_1^\ddagger, x_2)$  and thus the optimal quantizer at high enough resolution has  $Q_1(x_1^\dagger) = Q_1(x_1^\ddagger)$ , giving a non-regular quantizer. When the argument vector  $(x_1, x_2)$  of the function is rotated as shown on the right, the resulting function is still non-monotonic. However, there is no longer a clearly optimal non-regular quantization scheme. Specifically, for some fixed  $x_2$  there may be pairs  $(x_1^\dagger, x_1^\ddagger)$  such that  $g(x_1^\dagger, x_2) = g(x_1^\ddagger, x_2)$ , but the equality does not hold for all  $x_2$ . As we shall see, this results in the suboptimality of any compander that maps  $x_1^\dagger$  in the same way as  $x_1^\ddagger$ .

Our approach is to first create a model for high-resolution non-regular quantization, then to use this model to expand the class of functions for which regular quantization is optimal, and finally to construct asymptotically optimal non-regular quantizers when regularity is suboptimal.

### A. High-Resolution Non-Regular Quantization

To accommodate non-regular quantization, we extend the compander-based model of quantization. In Bennett’s development of optimal companding, reviewed in Sec. II, it is natural to require  $w$  to be both monotonic and have a bounded derivative everywhere; the derivative  $w'(x)$  is proportional to the quantizer point density  $\lambda(x)$  that has been central in our development thus far. Whether we look at  $\lambda$  or  $w$ , the role is to set the relative sizes of the quantization cells.

Since optimal functional quantizers are not necessarily regular, we adapt the conventional development to implement non-regular quantizers.

*Definition 4:* A function  $w : [0, 1] \rightarrow [0, 1]$  is a *generalized compander* if it is continuous, piecewise monotonic with a finite number of pieces, and has bounded derivative over each piece.

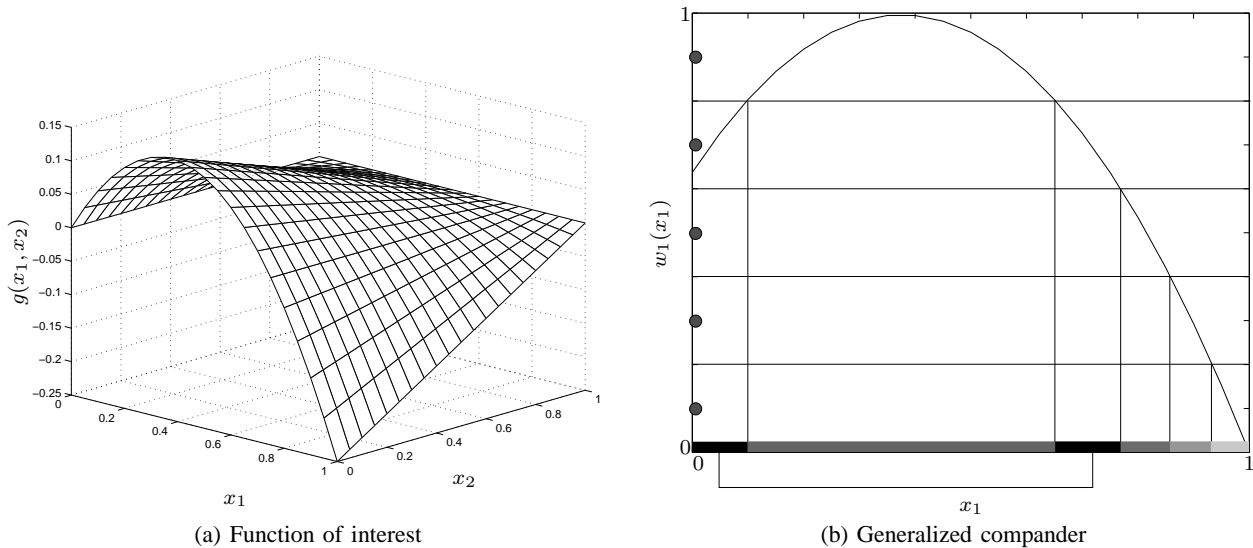


Fig. 8. Example of a generalized compander  $w_1(x_1)$  for a function  $g(x_1, x_2)$  and the partition resulting from uniform quantization of  $w_1(X_1)$ . Notice that the compander dictates both the relative sizes of cells and the binning of intervals of  $X$  values.

As in regular companding,  $w$  and  $w^{-1}$  are used along with a uniform quantizer  $Q_K^U$  as  $w^{-1}(Q_K^U(w(x)))$ . The restriction to a finite number of pieces is a limitation on the types of non-regular quantizers that can be captured with this model: those for which every quantizer cell is a finite union of intervals. Barring certain pathological situations, this restriction is reasonable.

Along with setting relative sizes of cells,  $w$  provides for non-regularity by allowing intervals to be binned together. To illustrate this, consider a simple example. Suppose that the pair  $(X_1, X_2)$  is uniformly distributed over  $[0, 1]^2$ , variable rate quantization is to be performed on both variables, and the function of interest is defined by

$$g(x_1, x_2) = x_1 \left( \frac{3}{4} - x_1 \right) (1 - x_2).$$

An optimal functional quantizer—a quantizer for  $X_1$  to minimize  $\mathbf{E}[(g(X_1, X_2) - g(\hat{X}_1, \hat{X}_2))^2]$ —should bin together  $X_1$  values that always yield the same  $g(X_1, X_2)$ . Furthermore, the magnitude of the slope of this quantizer should follow (22). The choice of

$$w_1(x_1) = \frac{64}{25} x_1 \left( \frac{3}{4} - x_1 \right) + \frac{16}{25}$$

can be shown to be optimal. Both  $w_1$  and the resulting quantizer at resolution  $K = 5$  are illustrated in Fig. 8b.

### B. Equivalence-Free Functions

We now define a broad class of functions for which regular quantization is optimal at sufficiently high resolutions. Consider the design of the  $j$ th quantizer in an  $n$ -dimensional distributed functional quantization setting.

We require a set of definitions:

*Definition 5:* For any  $s \neq t$  in the support of  $X_j$ , let

$$v_j(s, t) = \mathbf{E}[\text{var}(g(X_1^n) \mid X_j \in \{s, t\}, \{X_i\}_{i \neq j})].$$

If  $v_j(s, t) = 0$  then  $(s, t)$  is a *functional equivalence in the  $j$ th variable*. If  $g$  has no functional equivalences in any of its variables, we say it is *equivalence-free*.

The theorem below demonstrates that for DFSQ with an equivalence-free function, quantizer regularity is necessary for asymptotic optimality. Specifically, strictly non-regular quantization is shown to introduce a nonzero lower bound on the distortion, independent of rate. This is formalized with the aid of generalized companding. To simplify the proof somewhat, we assume that the marginal probability density  $f_j(X_j)$  is nonzero over  $[0, 1]$ . This assumption is without loss of generality, since one may consider the subset of  $[0, 1]$  where  $f_j(X_j)$  is nonzero.

*Theorem 14:* Let  $g$  be equivalence-free with respect to the pdf of  $X_1^n$  on  $[0, 1]^n$ . Suppose quantization of each  $X_j$  is performed as  $\hat{Y}_j = q(w_j(X_j))$  where  $w_j$  is a generalized compander and  $q$  is a uniform quantizer. If there is an index  $j$ , set  $S \subset [0, 1]$ , and function  $t: \mathbb{R} \rightarrow \mathbb{R}$  such that  $\mathbf{P}(X_j \in S) > 0$ , and, for every  $s \in S$ ,  $s \neq t(s)$  and  $w_j(s) = w_j(t(s))$ , then the distortion has a positive, resolution-independent lower bound.

*Proof:* See Appendix C. ■

The positive, rate-independent lower bound shows that the quantizer is suboptimal if the rate is sufficiently high; even naive uniform quantization will yield distortion with  $O(2^{-2R})$  dependence on rate and thus will eventually outperform the strictly non-regular quantizer.

When a function has equivalences, the best asymptotic quantization tactic is to design companders that bin all the equivalent values in each variable but are otherwise monotonic. In effect, this procedure losslessly converts the function into one that is equivalence-free. One might consider this a real-valued-source analogue of the functional compression procedure suggested by Doshi *et al.* [16].

## VII. DON'T-CARE INTERVALS AND RATE AMPLIFICATION

Ordinary high-resolution analysis produces point-density functions that reflect the source pdf in the sense that optimal quantizers never have zero point density where there is nonzero probability density. In fact, having zero point density where there is nonzero probability density invalidates high-resolution analysis. The situation is more complicated in the functional setting since the optimal point densities depend on both the functional sensitivity profiles and the source probability distribution. Having zero functional sensitivity where the probability density is nonzero changes the optimal quantizers in the variable-rate case.

The following example illustrates the potential for failure of the analysis of Section IV-C4. Note that the intricacies arise even with a univariate function.

*Example 6:* Let  $X$  have the uniform probability distribution over  $[0, 1]$ , and suppose the function of interest is  $g(X) = \min(X, 1/2)$ . It is clear that the optimal quantizer (for both fixed- and variable-rate) has uniform point density on  $[0, 1/2]$ . With the functional sensitivity profile given by

$$\gamma(x) = \begin{cases} 1, & \text{if } x < 1/2; \\ 0, & \text{otherwise,} \end{cases}$$

evaluating (12) and (14) is consistent with the intuitive result.

The distortion for the fixed-rate case obtained from (13) is  $(1/12)(1/2)^3 2^{-2R}$ . This is sensible since for half of the source values ( $X > 1/2$ ) there is zero distortion by having a single codeword at  $1/2$ , whereas for the other half of the source values ( $X < 1/2$ ),  $2^R - 1$  codewords quantize a random variable uniformly distributed over  $[0, 1/2]$ . However, assumption MF4 is not satisfied by this quantizer point density, so it is unclear whether this expression is an asymptotically valid approximation for the distortion-rate function.

The variable-rate case is also problematic. Since  $\mathbf{E}[\log \gamma(X)] = -\infty$ , evaluating (15) yields  $D_{\text{vr}}^{\text{HR}} = 0$ . Both the distortion-resolution and resolution-rate analyses fail because the quantization is not fine over the full support of  $f_X$ . However, if an alternative quantization structure is used, the distortion-rate performance can be accurately determined. In this alternative structure, the first representation bit specifies the event  $A = \{X < 1/2\}$  or its complement. Since additional bits are useful only when  $A$  occurs, one can spend  $2(R - 1)$  bits in those cases to have an average expenditure of  $R$  bits. The resulting distortion is

$$\begin{aligned} D_{\text{vr}}^{\text{HR}} &= \mathbf{P}(A) d_{g|A}^{\text{HR}} + \mathbf{P}(A^c) d_{g|A^c}^{\text{HR}} \\ &= \frac{1}{2} \cdot \frac{1}{12} \left(\frac{1}{2}\right)^2 2^{-2(2R-2)} + \frac{1}{2} \cdot 0 = \frac{1}{6} 2^{-4R}. \end{aligned}$$

Note that the exponent in the distortion-rate relationship is larger than it was in the fixed-rate case.  $\square$

In the example, there is an interval  $X \in [1/2, 1]$  of source values that need not be distinguished for function evaluation. Let us define a term for such intervals before discussing the example further.

*Definition 6:* An interval  $Z \subset [0, 1]$  is called a *don't-care interval* for the  $j$ th variable when the  $j$ th functional sensitivity  $\gamma_j$  is identically zero on  $Z$ , but the probability  $\mathbf{P}(X_j \in Z)$  is positive.

In univariate FSQ, at sufficiently high rates, each don't-care interval corresponding to a distinct value of the function should be allotted one codeword. This follows from reasoning similar to that given in Section VI-B and is illustrated by Example 6. In the fixed-rate case, the don't-care intervals simply occupy a few of the  $2^R$  codewords and have a limited effect. In the variable-rate case, however, the don't-care intervals produce a subset of source values that can be allotted very little rate. This gives more rate to be allotted outside the don't-care intervals and behavior we refer to as *rate amplification*.

We derive the high-resolution distortion-resolution function for this quantizer structure in Section VII-A, and in section VII-B the distortion-rate function is obtained.

### A. The Distortion-Resolution Function

In the following analysis we will assume that the  $j$ th variable has a finite number  $M_j$  of don't-care intervals  $\{Z_{j,1}, Z_{j,2}, \dots, Z_{j,M_j}\}$ . We also assume

$$\mathbf{P}(X_j \in Z_j) < 1 \quad \text{for } j = 1, 2, \dots, n, \quad (25)$$

where  $Z_j = \cup_{i=1}^{M_j} Z_{j,i}$  denotes the union of don't-care intervals for the  $j$ th variable. Without this, there is no improvement beyond  $M_j$  levels in representing  $X_j$ , so the high-resolution approach is wholly inappropriate. We will denote the event  $X_j \notin Z_j$  by  $A_j$ .

At sufficiently high rates, it is intuitive to allot a codeword of  $Q_j$  to each don't-care interval  $Z_{j,i}$ . The remaining  $K_j - M_j$  codewords are assigned optimally to  $[0, 1] \setminus Z_j$  according to the basic theory developed in Section IV. We refer to this quantizer structure as a *don't-care quantizer*.

*Theorem 15:* Suppose  $n$  sources  $X_1^n \in [0, 1]^n$  are quantized by a sequence of distributed don't-care quantizers  $Q_K^w$ . Further suppose that the sources, quantizers, and function  $g : [0, 1]^n \rightarrow \mathbb{R}$  satisfy assumptions MF1–MF3, and assumption MF4 is replaced by the following: The integrals

$$\int_{[0,1] \setminus Z_j} f(x_j) \mathbf{E} [|g_j(X^n)|^2 | X_j = x_j] \lambda_j(x_j)^{-2} dx_j$$

are finite for every  $j \in \{1, \dots, n\}$ . Finally, assume each source  $X_j$  has  $M_j$  don't-care intervals satisfying (25). Then the high-resolution distortion-resolution function is asymptotically accurate to the true distortion-resolution function:

$$\begin{aligned} d^{\text{HR}}(K; \boldsymbol{\lambda}) &= \frac{n}{12} \left( \prod_{j=1}^n \frac{\mathbf{P}(A_j)}{(K_j - M_j)^2} \mathbf{E} \left[ \left( \frac{\gamma_j(X_j)}{\lambda_j(X_j)} \right)^2 | A_j \right] \right)^{1/n}, \\ &\sim d(K; \boldsymbol{\lambda}). \end{aligned} \quad (26)$$

*Proof:* Follows from applying Theorem 9 to the region  $([0, 1] \setminus Z_1) \times \dots \times ([0, 1] \setminus Z_n)$ . ■

### B. The Distortion-Rate Functions

In the fixed-rate case, the high-resolution resolution-rate function is unchanged:  $K_{\text{fr}}^{\text{HR}}(R; \boldsymbol{\lambda}) = 2^R$ . Asymptotic validity is easily observed:  $\lim_{R \rightarrow \infty} K_{\text{fr}}(R; \boldsymbol{\lambda}) / K_{\text{fr}}^{\text{HR}}(R; \boldsymbol{\lambda}) = 1$ . Applying this to the distortion-resolution expression (26), we obtain the unoptimized high-resolution fixed-rate distortion-rate function:

$$\begin{aligned} D_{\text{fr}}(R; \boldsymbol{\lambda}) \sim D_{\text{fr}}^{\text{HR}}(R; \boldsymbol{\lambda}) &= \frac{n}{12} \left( \prod_{j=1}^n \frac{\mathbf{P}(A_j)}{(K_j - M_j)^2} \mathbf{E} \left[ \left( \frac{\gamma_j(X_j)}{\lambda_j(X_j)} \right)^2 | A_j \right] \right)^{1/n} \\ &\sim \frac{n}{12K^{2/n}} \left( \prod_{j=1}^n \mathbf{P}(A_j) \mathbf{E} \left[ \left( \frac{\gamma_j(X_j)}{\lambda_j(X_j)} \right)^2 | A_j \right] \right)^{1/n}. \end{aligned}$$

The optimal point densities for fixed-rate quantization are given by (20) outside of the don't-care intervals. These point densities yield an optimized high-resolution fixed-rate distortion-rate function

$$D_{\text{fr}}^{\text{HR}}(R) = \frac{n}{12} \left( \prod_{j=1}^n \|\gamma_j^2 f_{X_j}\|_{1/3} \right)^{1/n} 2^{-2R/n}. \quad (27)$$

The variable-rate case is a bit more involved. To formalize the analysis, we define discrete random variables to represent the events of source variables lying in don't-care intervals.

*Definition 7:* The random variable

$$I_j = \begin{cases} i, & \text{if } X_j \in Z_{j,i} \text{ for } i \in \{1, 2, \dots, M_j\}; \\ 0, & \text{otherwise} \end{cases}$$

is called the  $j$ th *don't-care variable*. The previously-defined event  $A_j$  can be expressed as  $\{I_j = 0\}$ .

At sufficiently high rate, the  $j$ th encoder communicates  $I_j$  and in addition, *only when*  $I_j = 0$ , a fine quantization of  $X_j$ . The resulting performance is summarized by the following theorem.

*Theorem 16:* Under the conditions of Theorem 15, the optimal point densities for variable-rate quantization follow (22) and yield

$$\begin{aligned} D_{\text{vr}}^{\text{HR}}(R) &= \frac{n}{12} \left( \prod_{j=1}^n \rho_j^{-1} \right. \\ &\quad \left. \times 2^{-2(\rho_j(R - H(I_j)) + 2h(X_j | A_j) + 2\mathbf{E}[\log(\gamma_j(X_j)) | A_j])} \right)^{1/n}. \end{aligned} \quad (28)$$

where  $\rho_j = 1/\mathbf{P}(A_j)$  is the *amplification* of  $R_j$ .

*Proof:* See Appendix D. ■

Some remarks:

- 1) The quantity  $H(I_j)$  may be identified as the cost of communicating the indicator information to the decoder. The remaining rate,  $R_j - H(I_j)$ , is amplified by factor  $\rho_j$  because additional description of  $X_j$  is useful only when  $X_j \notin Z_j$ .

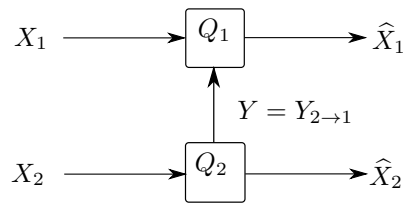


Fig. 9. Suppose the encoder for  $X_2$  could send a bit to the encoder for  $X_1$ . Is there any benefit? How does it compare to sending an additional bit to the decoder?

The amplification shows that the standard  $-6$  dB/bit-per-source distortion decay may be exceeded in the presence of don't-care regions.

- 2) At moderate rates, it may not be optimal to communicate  $I_j$  losslessly, and it may be beneficial to include  $X_j$  values with small but positive  $\gamma_j$  in don't-care intervals. Study of this topic is left for specific applications.
- 3) The rate amplification we have seen in the variable-rate case and the relative lack of importance of don't-care intervals in the fixed-rate case have a close analogy in ordinary lossy source coding. Suppose a source  $X$  is a mixed random variable with an  $M$ -valued discrete component and a continuous component. High-resolution quantization of  $X$  will allocate one level to each discrete value and the remaining levels to the continuous component. The discrete component changes the constant factor in  $\Theta(2^{-2R})$  fixed-rate operational distortion–rate performance while it changes the decay rate in the variable-rate case. See [36] for related Shannon-theoretic (rather than high-resolution quantization) results.

### VIII. CHATTING ENCODERS

Our final variation on the basic theory of distributed functional scalar quantization is to allow limited communication between the encoders. How much can the distortion be reduced via this communication? Echoing the results of the previous section, we will find dramatically different answers in the fixed- and variable-rate cases.

For notational convenience, we will fix the communication to be from encoder 2 to encoder 1 though the number of source variables  $n$  remains general. In accordance with the block diagram of Fig. 9, the information  $Y = Y_{2 \rightarrow 1}$  must be conditionally independent of  $X_1$  given  $X_2$ . We consider only the case where  $Y$  is a single bit; this suffices to illustrate the key ideas.

In this section, we express the high-resolution distortion as

$$D^{\text{HR}} = \frac{n}{12} 2^{-2R/n} \left( \prod_{j=1}^n D_j \right)^{1/n},$$

where various expressions for  $D_j$  have been found for different scenarios, including for fixed-rate (21) and variable-rate (23) quantization. At issue is how  $D_1$  is affected by  $Y$ ; the other  $D_j$ s are not affected.

#### A. Fixed-Rate Quantization

In general, the availability of a single bit  $Y$  causes one to choose between two potentially-different quantizers  $Q_{1|Y=0}$  and  $Q_{1|Y=1}$  in the quantization of  $X_1$ . We express the optimal quantizers and the resulting distortion contribution  $D_1$  by way of the following concept.

*Definition 8:* The  $j$ th conditional functional sensitivity profile of  $g$  given  $Y = y$  is defined as

$$\gamma_{j|Y}(x|y) = \left( \mathbf{E} \left[ |g_j(X_1^n)|^2 \mid X_j = x, Y = y \right] \right)^{1/2}.$$

Now several results follow by analogy with Theorem 13. For the case of  $Y = y$ , the optimal point density is given by

$$\lambda_{1|Y}(x|y) = \frac{\left( \gamma_{1|Y}^2(x|y) f_{X_1|Y}(x|y) \right)^{1/3}}{\int_0^1 \left( \gamma_{1|Y}^2(t|y) f_{X_1|Y}(t|y) \right)^{1/3} dt}$$

resulting in conditional distortion contribution

$$\frac{1}{12K_1^2} \left\| \gamma_{1|Y=y}^2 f_{X_1|Y=y} \right\|_{1/3}.$$

Combining the two possibilities for  $Y$  via total expectation gives

$$D_1 = \sum_{y=0}^1 \mathbf{P}(Y = y) \left\| \gamma_{1|Y=y}^2 f_{X_1|Y=y} \right\|_{1/3}. \quad (29)$$



From this expression we reach an important conclusion on the effect of the chatting bit  $Y$ .

*Theorem 17:* For fixed-rate quantization, communication of one bit of information from decoder 2 to decoder 1 will asymptotically reduce  $D_1$  by at most a factor of 4.

*Proof:* From Theorem 13, the distortion contribution analogous to (29) without the chatting bit  $Y$  is  $\|\gamma_1^2 f_{X_1}\|_{1/3}$ . Thus the fact we wish to prove is a statement about  $\mathcal{L}^{1/3}$  quasinorms of weighted densities and their conditional forms.

We proceed as follows:

$$\begin{aligned}
D_1 &= \sum_{y=0}^1 \left\| \mathbf{P}(Y=y) \gamma_{1|Y}^2(x|y) f_{X_1|Y}(x|y) \right\|_{1/3} \\
&\stackrel{(a)}{\geq} \frac{1}{4} \left\| \sum_{y=0}^1 \mathbf{P}(Y=y) \gamma_{1|Y}^2(x|y) f_{X_1|Y}(x|y) \right\|_{1/3} \\
&= \frac{1}{4} \left\| f_{X_1}(x) \sum_{y=0}^1 \frac{\mathbf{P}(Y=y) f_{X_1|Y}(x|y)}{f_{X_1}(x)} \gamma_{1|Y}^2(x|y) \right\|_{1/3} \\
&\stackrel{(b)}{=} \frac{1}{4} \left\| f_{X_1}(x) \sum_{y=0}^1 \mathbf{P}(Y=y | X_1=x) \gamma_{1|Y}^2(x|y) \right\|_{1/3} \\
&\stackrel{(c)}{=} \frac{1}{4} \|f_{X_1}(x) \gamma_1^2(x)\|_{1/3},
\end{aligned}$$

where (a) uses a quasi-triangle inequality that may be established via well-known inequalities (see Appendix E for a statement and proof); (b) is an application of Bayes's Rule; and (c) is based on an evaluation of the (unconditional) functional sensitivity via the total expectation theorem with conditioning on  $Y$ . This proves the theorem.  $\blacksquare$

Note that while the bit  $Y$  leads a reduction of  $D_1$  by at most a factor of 4 and therefore a reduction of  $D^{\text{HR}}$  by at most a factor of  $4^{1/n}$ , an identical reduction in distortion is achieved simply by increasing the rate  $R$  to the centralized decoder by one bit. Generalizing to any number of chatting bits, we obtain the following corollary.

*Corollary 18:* For fixed-rate functional quantization, communication of some number of bits from encoder  $j$  to encoder  $k$  performs *at best* as well as increasing the communication to the centralized decoder by the same number of bits.

In general, the idea that bits from encoder 2 to encoder 1 are as good as bits from encoder 1 to the decoder is optimistic. In particular, if  $\mathbf{E}[\gamma_1^2(X_1)] > 0$ , then  $D_1$  is bounded away from zero for any amount of communication from encoder 2 to encoder 1.

## B. Variable-Rate Quantization

In a variable-rate scenario, the rate could be made to depend on the chatting bit  $Y$ , introducing a bit allocation problem between the cases of  $Y = 0$  and  $Y = 1$ . Even without such dependence, we can demonstrate that the bit  $Y$  can reduce the first variable's contribution to the functional distortion by an arbitrary factor.

Analogous to (29),

$$D_1 = \sum_{y=0}^1 \mathbf{P}(Y=y) 2^{2h(X_1|Y=y)+2\mathbf{E}[\log \gamma_{1|Y=y}(X_1)]} \quad (30)$$

by comparison with (23). In contrast to the  $\mathcal{L}^{1/3}$  quasinorms in (29), this linear combination can be arbitrarily smaller than

$$2^{2h(X_1)+2\mathbf{E}[\log \gamma_1(X_1)]}.$$

We demonstrate this through a simple example.

*Example 7:* Let sources  $X_1$  and  $X_2$  be uniformly distributed on  $[0, 1]^2$ . We specify the function of interest  $g$  through its partial derivatives. Let  $g_2(x_1, x_2) = 1$  for all  $(x_1, x_2)$  and let  $g_1(x_1, x_2)$  be piecewise constant as shown in Fig. 10, where  $L$  is a positive constant.

While  $g(x_1, x_2)$  is not continuous everywhere and condition MF1 is therefore not strictly satisfied, the points of discontinuity fall along the line  $x_2 = 1/2$ . As observed following the proof of Corollary 8, this variety of discontinuity can be easily and intuitively merged with high-resolution analysis: one simply places an extra quantizer cell boundary at  $x_2 = 1/2$  for every quantizer in the sequence being considered. This increases the resolution  $K$  by 1, but has negligible impact on the rate in the limit  $K \rightarrow \infty$ .

We can easily derive the first functional sensitivity profile of  $g$  to be

$$\gamma_1(x) = \sqrt{\frac{1}{2}(L^2 + 1)}.$$

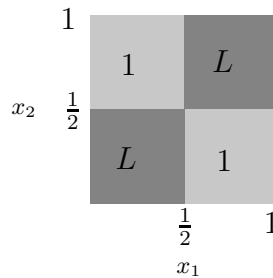


Fig. 10. Illustration for Example 7. Shown is the unit square  $[0, 1]^2$  with quadrants marked with the value of  $g_1(x_1, x_2)$ , the derivative of  $g$  with respect to  $x_1$ .

This also allows us to find the distortion contribution factor  $D_1$  without chatting to be

$$D_1 = \frac{1}{4}(L^2 + 1)^2.$$

In this example, one bit about  $X_2$  is enough to allow the encoder for  $X_1$  to perfectly tailor its point density to match the sensitivity of  $g$  at  $(X_1, X_2)$ . Of course, the chatting bit should simply be

$$Y = \begin{cases} 0, & \text{if } X_2 > 1/2; \\ 1, & \text{otherwise.} \end{cases}$$

The first conditional functional sensitivity profiles for  $g$  are then

$$\gamma_{1|Y}(x | y) = \begin{cases} 1, & \text{for } Y = 0 \text{ and } X_1 \leq 1/2 \\ & \text{or } Y = 1 \text{ and } X_1 > 1/2; \\ L, & \text{otherwise.} \end{cases}$$

Now for either value of  $y$ , we have  $\int_0^1 \gamma_{1|Y}(x | y) dx = \frac{1}{2}(L + 1)$  and  $\mathbf{E}[\log \gamma_{1|Y=y}(X_1)] = \frac{1}{2} \log L$ . Thus, evaluating (30) gives

$$D_1 = \frac{1}{4}(L + 1)^2 L.$$

This is smaller than the  $D_1$  with no chatting by about a factor of  $L$ . The performance gap can be made arbitrarily large by increasing  $L$ —all from just one bit of information communicated between encoders per sample.  $\square$

### C. Comparison with Ordinary Source Coding

The results of this section are strikingly different from those of ordinary source coding. Consider first the discrete scenario in which we wish to recreate  $X_1^n$  perfectly at the decoder. Can communication between encoders enable a reduction in the rate of communication to the decoder? According to Slepian and Wolf, the answer is a resounding “no.” Even in the case of unlimited collaboration via fused encoders, the minimum sum rate to the decoder remains unchanged.

How about in lossy source coding? If quantization is variable-rate and Slepian–Wolf coding is employed on the quantization indices, no gains are possible from encoder interactions. This is a consequence of the work of Rebollo-Monedero *et al.* [8] on high-resolution Wyner–Ziv coding, where it is shown that there is no gain from supplying the source encoder with the decoder side information.

## IX. SUMMARY

We have developed asymptotically-optimal companding designs of functional quantizers using high-resolution quantization theory. This has shown that accounting for a function while quantizing a source can lead to arbitrarily large improvements in distortion. In certain scenarios (Section V), this improvement can grow exponentially with the number of sources. In others (Section VII), it can grow exponentially with rate.

Additionally, our study of functional quantization has highlighted some striking distinctions between fixed- and variable-rate cases:

- 1) For certain simple functions of order statistics, distortion relative to ordinary quantization falls polynomially with the number of sources in the fixed-rate case, whereas in the variable-rate case it falls exponentially.
- 2) The distortion associated with fixed-rate quantizers will always exhibit  $-6$  dB/bit rate dependence at high rates, whereas the decay of distortion can be faster in some variable-rate cases.
- 3) Information sent from encoder-to-encoder can lead to arbitrarily-large improvements in distortion for variable-rate, whereas for fixed-rate this information can be no more useful than an equal amount of information sent to the decoder.

The second and third of these have extensions or analogues beyond functional quantization. Rate amplification is a feature of quantizing sources with mixed probability distributions, and the results on chatting encoders continue to hold when the function  $g$  is the identity operation.

APPENDIX A  
PROOF OF THEOREM 9

*Lemma 19:* Let  $X$  be a real-valued random variable distributed over a bounded interval  $S$ , and let  $g(x)$  be an absolutely continuous, real-valued function on  $S$  with bounded derivative  $g'(x)$ . If  $g'(x)$  is defined for *almost every*  $x \in S$ , then  $\text{var}(g(X)) \leq \text{var}(bX)$ , where  $b = \sup_{x \in S} |g'(x)|$ . If  $g'(x)$  is furthermore defined for *every*  $x \in S$ , then  $\text{var}(g(X)) \geq \text{var}(aX)$ , where  $a = \inf_{x \in S} |g'(x)|$ .

*Proof:* Draw  $X_1$  and  $X_2$  i.i.d. according to the distribution of  $X$ , and define the following functions:

$$\begin{aligned} D_a(X_1, X_2) &= a(X_1 - X_2), \\ D_g(X_1, X_2) &= g(X_1) - g(X_2), \\ D_b(X_1, X_2) &= b(X_1 - X_2). \end{aligned}$$

To prove the first part of the lemma, assume the derivative of  $g$  is defined for *almost every*  $x \in S$ . By the absolute continuity of  $g$ ,

$$D_g(X_1, X_2) = g(X_2) - g(X_1) = \int_{X_1}^{X_2} g'(x) dx.$$

Therefore, the magnitude  $|D_g(X_1, X_2)|$  can be bounded above as follows:

$$|D_g(X_1, X_2)| \leq \int_{X_1}^{X_2} |g'(x)| dx \leq b|X_1 - X_2| = |D_b(X_1, X_2)|.$$

This then implies that  $\mathbf{E}[D_g^2] \leq \mathbf{E}[D_b^2]$  and, since  $\mathbf{E}[D_g] = \mathbf{E}[D_b] = 0$ , that

$$\text{var}(D_g) \leq \text{var}(D_b). \quad (31)$$

Since each  $D_b$  and  $D_g$  is a sum of i.i.d. variables,  $\text{var}(D_b) = 2 \text{var}(bX)$  and  $\text{var}(D_g) = 2 \text{var}(g(X))$ . Inserting these into (31) and dividing by 2 proves the first part of the lemma.

Now to prove the second part of the lemma, assume further that  $g'(x)$  is defined for *every*  $x \in S$ . By the mean value theorem, there exists  $X_0$  between  $X_1$  and  $X_2$  such that

$$|D_g(X_1, X_2)| = \left| \int_{X_1}^{X_2} g'(x) dx \right| = |g'(X_0)(X_1 - X_2)| \geq a|X_1 - X_2| = |D_a(X_1, X_2)|.$$

As before, this implies that

$$\text{var}(D_g) \geq \text{var}(D_a). \quad (32)$$

Since  $\text{var}(D_a) = 2 \text{var}(aX)$  and  $\text{var}(D_g) = 2 \text{var}(g(X))$ , substituting into (32) and dividing by 2 proves the second part of the lemma. ■

We now define a function  $\tilde{g}_{\tilde{x}_1^{j-1}}(x)$  that will appear in the proof of the theorem after we establish properties of the function in a lemma.

*Definition 9:* Suppose  $\tilde{X}_1^n$  is uniformly distributed over a rectangular region  $S$ . The  $j$ th *reduced-dimension function* (with parameter vector  $\tilde{x}_1^{j-1}$ ) is defined as

$$\tilde{g}_{\tilde{x}_1^{j-1}}(x) = \mathbf{E} \left[ g(\tilde{X}_1^n) \mid \tilde{X}_1^{j-1} = \tilde{x}_1^{j-1}, \tilde{X}_j = x \right].$$

*Lemma 20:* Let  $\tilde{X}_1^n$  be uniformly distributed over a rectangular region  $S = S_1 \times S_2 \times \cdots \times S_n$ , and let  $g$  be Lipschitz continuous over  $S$ . If the first and second derivatives of  $g$  are defined and bounded *almost everywhere* in  $S$ , then:

- 1)  $\tilde{g}_{\tilde{x}_1^{j-1}}(x)$  is Lipschitz continuous in  $x$ .
- 2) Where defined,  $|\tilde{g}'_{\tilde{x}_1^{j-1}}(x)| \leq b_j = \sup_{x_1^n \in S} |g_j(x_1^n)|$ .

If the first and second derivatives of  $g$  are furthermore defined and bounded *everywhere* in  $S$ , then:

- 3) For all  $\tilde{x}_1^{j-1} \in S_1 \times S_2 \times \cdots \times S_{j-1}$ , the derivative  $\tilde{g}'_{\tilde{x}_1^{j-1}}(x)$  is defined for all  $x \in S_j$ .
- 4) The magnitude of derivative may be lower bounded:  $|\tilde{g}'_{\tilde{x}_1^{j-1}}(x)| \geq a_j = \inf_{x_1^n \in S} |g_j(x_1^n)|$ .

*Proof:* First, assume that the first and second derivatives of  $g$  are defined and bounded *almost everywhere* in  $S$ .

- 1) Since  $\tilde{g}_{\tilde{x}_1^{j-1}}(x)$  is an average of functions with a common Lipschitz constant, it too is Lipschitz with this constant.
- 2) Where  $\tilde{g}'_{\tilde{x}_1^{j-1}}(x)$  is defined, we have

$$\begin{aligned} \left| \tilde{g}'_{\tilde{x}_1^{j-1}}(x_j) \right| &= \left| \mathbf{E} \left[ g_j(X_1^n) \mid \tilde{X}_1^j = \tilde{x}_1^j \right] \right| \\ &\leq \mathbf{E} \left[ |g_j(X_1^n)| \mid \tilde{X}_1^j = \tilde{x}_1^j \right] \\ &\leq b_j. \end{aligned}$$

Now assume that furthermore the first and second derivatives of  $g$  are defined and bounded *everywhere* in  $S$ .

- 3) Since  $g_j(\tilde{x}_1^n)$  is defined for all  $\tilde{x}_1^n \in S$ , the average derivative  $\tilde{g}'_{\tilde{x}_1^{j-1}}(\tilde{x}_j)$  is defined for all  $\tilde{x}_j \in S_j$ .
- 4) We now obtain a lower bound on the derivative. As before, we note that  $\tilde{g}'_{\tilde{x}_1^{j-1}}(x_j) = \mathbf{E}[g_j(X_1^n) \mid \tilde{X}_1^j = \tilde{x}_1^j]$ . Because the derivatives of  $g_j$  are defined everywhere in  $S$ , and because the expectation under a uniform distribution is just an average, the mean value theorem guarantees the existence of an  $e_{\tilde{x}_1^j} \in S$  such that  $g_j(e_{\tilde{x}_1^j}) = \mathbf{E}[g_j(X_1^n) \mid \tilde{X}_1^j = \tilde{x}_1^j]$ . Finally, since  $e_{\tilde{x}_1^j} \in S$ , we have  $|g_j(e_{\tilde{x}_1^j})| \geq a_j$ . To summarize:

$$\begin{aligned} \left| \tilde{g}'_{\tilde{x}_1^{j-1}}(x_j) \right| &= \left| \mathbf{E} \left[ g_j(X_1^n) \mid \tilde{X}_1^j = \tilde{x}_1^j \right] \right| \\ &= \left| g_j(e_{\tilde{x}_1^j}) \right| \\ &\geq a_j. \end{aligned}$$

■

We now bound the variance of the function within a rectangular cell, assuming the source is uniformly distributed. This will later be adapted to the case where the source is nonuniformly distributed.

*Lemma 21:* Let  $g(x_1^n)$  be a Lipschitz continuous function defined over a rectangular cell  $S = S_1 \times \dots \times S_n$  with edge lengths  $\Delta_1, \dots, \Delta_n$ , let  $a_j$  and  $b_j$  be lower and upper bounds to  $|g_j(x_1^n)|$ , when it exists, and let  $\bar{g}$  denote the average value of  $g$  within  $S$ :

$$\bar{g} = \frac{1}{\prod_{j=1}^n \Delta_j} \int_S g(x_1^n) dx_1^n.$$

If the first and second derivatives of  $g$  are defined *almost everywhere* in  $S$ , then

$$\left( \prod_{j=1}^n \frac{1}{\Delta_j} \right) \int_S |g(x_1^n) - \bar{g}|^2 dx_1^n \leq \sum_{j=1}^n \frac{b_j^2 \Delta_j^2}{12}.$$

If the first and second derivatives of  $g$  are furthermore defined *everywhere* in  $S$ , then

$$\sum_{j=1}^n \frac{a_j^2 \Delta_j^2}{12} \leq \left( \prod_{j=1}^n \frac{1}{\Delta_j} \right) \int_S |g(x_1^n) - \bar{g}|^2 dx_1^n.$$

*Proof:* Since  $\tilde{X}_1^n$  is uniformly distributed over  $S$ ,

$$\frac{1}{\prod_{j=1}^n \Delta_j} \int_S |g(x_1^n) - \bar{g}|^2 dx_1^n = \text{var}(g(\tilde{X}_1^n)).$$

This may be expanded by repeated application of the law of total variance:

$$\begin{aligned} \text{var}(g(\tilde{X}_1^n)) &\stackrel{(a)}{=} \mathbf{E} \left[ \text{var} \left( g(\tilde{X}_1^n) \mid \tilde{X}_1 \right) \right] + \text{var} \left( \mathbf{E} \left[ g(\tilde{X}_1^n) \mid \tilde{X}_1 \right] \right) \\ &\stackrel{(b)}{=} \mathbf{E} \left[ \mathbf{E} \left[ \text{var} \left( g(\tilde{X}_1^n) \mid \tilde{X}_1, \tilde{X}_2 \right) \mid \tilde{X}_1 \right] + \text{var} \left( \mathbf{E} \left[ g(\tilde{X}_1^n) \mid \tilde{X}_1, \tilde{X}_2 \right] \mid \tilde{X}_1 \right) \right] + \text{var} \left( \mathbf{E} \left[ g(\tilde{X}_1^n) \mid \tilde{X}_1 \right] \right) \\ &\stackrel{(c)}{=} \mathbf{E} \left[ \text{var} \left( g(\tilde{X}_1^n) \mid \tilde{X}_1, \tilde{X}_2 \right) \right] + \mathbf{E} \left[ \text{var} \left( \mathbf{E} \left[ g(\tilde{X}_1^n) \mid \tilde{X}_1, \tilde{X}_2 \right] \mid \tilde{X}_1 \right) \right] + \text{var} \left( \mathbf{E} \left[ g(\tilde{X}_1^n) \mid \tilde{X}_1 \right] \right) \\ &\stackrel{(d)}{=} \mathbf{E} \left[ \text{var} \left( g(\tilde{X}_1^n) \mid \tilde{X}_1^{n-1} \right) \right] + \sum_{j=1}^n \mathbf{E} \left[ \text{var} \left( \mathbf{E} \left[ g(\tilde{X}_1^n) \mid \tilde{X}_1^j \right] \mid \tilde{X}_1^{j-1} \right) \right] \\ &= \sum_{j=1}^n \mathbf{E} \left[ \text{var} \left( \mathbf{E} \left[ g(\tilde{X}_1^n) \mid \tilde{X}_1^j \right] \mid \tilde{X}_1^{j-1} \right) \right] \\ &= \sum_{j=1}^n \mathbf{E} \left[ \text{var} \left( \tilde{g}_{\tilde{X}_1^{j-1}}(\tilde{X}_j) \mid \tilde{X}_1^{j-1} \right) \right], \end{aligned} \tag{33}$$

where (a) follows from the law of total variance with conditioning on  $\tilde{X}_1$ ; (b) uses the law of total variance applied to the variance within the expectation in the first term, with conditioning performed on  $\tilde{X}_2$ ; (c) simplifies the first term using iterated expectation; and (d) applies the law of total variance repeatedly to the variance within the first expectation, as in step (b), with conditioning on  $\tilde{X}_j$  during the  $j$ th iteration.

*Upper bound.* Let  $A(\tilde{x}_1^{j-1}) \in S_j$  be the set of points  $x_j \in S_j$  where the derivative  $\tilde{g}'_{\tilde{x}_1^{j-1}}(x_j)$  is undefined. By Lemma 20,  $\tilde{g}_{\tilde{x}_1^{j-1}}(x_j)$  is Lipschitz continuous and therefore  $A(\tilde{x}_1^{j-1})$  is of measure zero. As such, for every value of  $\tilde{X}_1^{j-1}$  considered

within the expectation,  $\text{var} \left( \tilde{g}_{\tilde{X}_1^{j-1}}(\tilde{X}_j) \mid \tilde{X}_1^{j-1} \right)$  is now the variance of a function satisfying the upper bound conditions for Lemma 19. Applying this upper bound within the expectation, we have

$$\begin{aligned} \sum_{j=1}^n \mathbf{E} \left[ \text{var} \left( \tilde{g}_{\tilde{X}_1^{j-1}}(\tilde{X}_j) \mid \tilde{X}_1^{j-1} \right) \right] &\leq \sum_{j=1}^n \mathbf{E} \left[ \text{var} \left( b_j \tilde{X}_j \mid \tilde{X}_1^{j-1} \right) \mid \mu \left( A(\tilde{X}_1^{j-1}) \right) = 0 \right] \\ &= \sum_{j=1}^n \frac{b_j^2 \Delta_j^2}{12}, \end{aligned}$$

which proves the first half of the lemma.

*Lower bound.* We return to (33), now assuming that the first and second derivatives of  $g$  are defined everywhere in the cell  $S$ . By Lemma 20, for any choice of  $\tilde{x}_1^j \in S_1 \times \cdots \times S_j$  the function  $\tilde{g}_{\tilde{x}_1^{j-1}}(x_j)$  satisfies the conditions for the lower bound in Lemma 19. Inserting this lower bound into the expectation, we obtain

$$\begin{aligned} \sum_{j=1}^n \mathbf{E} \left[ \text{var} \left( \tilde{g}_{\tilde{X}_1^{j-1}}(\tilde{X}_j) \mid \tilde{X}_1^{j-1} \right) \right] &\geq \sum_{j=1}^n \mathbf{E} \left[ \text{var} \left( a_j(\tilde{X}_j) \mid \tilde{X}_1^{j-1} \right) \right] \\ &= \sum_{j=1}^n \frac{a_j^2 \Delta_j^2}{12}, \end{aligned}$$

which proves the second half of the lemma. ■

Armed with this lemma, we may now determine upper and lower bounds to the distortion of  $g(x_1^n)$  within a single quantizer cell.

*Lemma 22:* Suppose that over a rectangular cell  $S \subset [0, 1]^n$  the function  $g(x_1^n)$  is Lipschitz continuous and the probability density  $f(x_1^n)$  is continuous, and suppose  $g(x_1^n)$  has bounded first and second derivatives almost-everywhere in  $S$ . Let  $A_S$  denote the subset of  $S$  where the first and second derivatives of  $g(x_1^n)$  are defined. Then, defining  $a_j = \inf_{x_1^n \in S} \mathbf{1}_{A_S}(x_1^n) |g_j(x_1^n)|$  and  $b_j = \sup_{x_1^n \in S} |g_j(x_1^n)|$ ,

$$f(\chi \mid X_1^n \in S) \sum_{j=1}^n \frac{a_j^2 \Delta_j^2}{12} \prod_{i=1}^n \Delta_i \leq \text{var} \left( g(X_1^n) \mid X_1^n \in S \right) \leq f(\xi \mid X_1^n \in S) \sum_{j=1}^n \frac{b_j^2 \Delta_j^2}{12} \prod_{i=1}^n \Delta_i$$

for some  $\chi, \xi \in S$ .

*Proof:* We first prove the lower bound. If  $A_S$  is nonempty,  $a_j = 0$  for every  $j$  so the lower bound is trivially true. Now suppose  $A_S$  is empty and therefore that the first and second derivatives of  $g(x_1^n)$  are defined everywhere in  $S$ . In this case,

$$\begin{aligned} \text{var} \left( g(X_1^n) \mid X_1^n \in S \right) &= \int_S f(x_1^n \mid X_1^n \in S) \left( g(X_1^n) - \mathbf{E} \left[ g(X_1^n) \mid X_1^n \in S \right] \right)^2 dx_1^n \\ &\stackrel{(a)}{=} f(\chi \mid X_1^n \in S) \int_S \left( g(X_1^n) - \mathbf{E} \left[ g(X_1^n) \mid X_1^n \in S \right] \right)^2 dx_1^n \\ &= f(\chi \mid X_1^n \in S) \left( \prod_{i=1}^n \Delta_i \right) \int_S \frac{1}{\left( \prod_{i=1}^n \Delta_i \right)} \left( g(X_1^n) - \mathbf{E} \left[ g(X_1^n) \mid X_1^n \in S \right] \right)^2 dx_1^n \\ &\stackrel{(b)}{=} f(\chi \mid X_1^n \in S) \left( \prod_{i=1}^n \Delta_i \right) \mathbf{E} \left[ \left( g(\tilde{X}_1^n) - \mathbf{E} \left[ g(X_1^n) \mid X_1^n \in S \right] \right)^2 \right] \\ &\stackrel{(c)}{\geq} f(\chi \mid X_1^n \in S) \left( \prod_{i=1}^n \Delta_i \right) \text{var} \left( g(\tilde{X}_1^n) \right) \\ &\stackrel{(d)}{\geq} f(\chi \mid X_1^n \in S) \left( \prod_{i=1}^n \Delta_i \right) \sum_{j=1}^n \frac{a_j^2 \Delta_j^2}{12}, \end{aligned}$$

where (a) follows from the first mean value theorem for integration; (b) introduces the random vector  $\tilde{X}_1^n$  that is uniform over  $S$ ; (c) is true because the variance is the smallest possible mean squared error from a constant estimate; and (d) follows from the lower bound in Lemma 21.

For the upper bound, we proceed in a similar manner:

$$\begin{aligned}
\text{var}(g(X_1^n) | X_1^n \in S) &= \int_S f(x_1^n | X_1^n \in S) (g(X_1^n) - \mathbf{E}[g(X_1^n) | X_1^n \in S])^2 dx_1^n \\
&\stackrel{(a)}{\leq} \int_S f(x_1^n | X_1^n \in S) (g(X_1^n) - \bar{g}_S)^2 dx_1^n \\
&\stackrel{(b)}{=} f(\xi | X_1^n \in S) \int_S (g(X_1^n) - \bar{g}_S)^2 dx_1^n \\
&= f(\xi | X_1^n \in S) \left(\prod_{i=1}^n \Delta_i\right) \int_S \frac{1}{\prod_{i=1}^n \Delta_i} (g(X_1^n) - \bar{g}_S)^2 dx_1^n \\
&\stackrel{(c)}{\leq} f(\xi | X_1^n \in S) \left(\prod_{i=1}^n \Delta_i\right) \sum_{j=1}^n \frac{b_j^2 \Delta_j^2}{12},
\end{aligned}$$

where in (a) we reintroduce the notation  $\bar{g} = \int_S g(X_1^n) \frac{1}{\prod_{i=1}^n \Delta_i}$  for the average value of  $g$  with respect to a uniform distribution, and the inequality is valid because the expected value of a random variable minimizes the mean-squared error of the estimate; (b) is due to the first mean value theorem for integration; and (c) follows from the upper bound in Lemma 21.  $\blacksquare$

At this point, we provide a proof of the theorem.

*Proof:* The distortion  $d(\mathbf{K}; \boldsymbol{\lambda})$  is given by

$$d_g = \sum_{i_1^n} \mathbf{P}(X_1^n \in S_{i_1^n}) \text{var}(g(X_1^n) | X_1^n \in S_{i_1^n}).$$

Let  $A \in [0, 1]^n$  denote the set of points  $x_1^n$  where both the first and second derivatives of  $g(x_1^n)$  are defined. By assumption MF1,  $[0, 1]^n \setminus A$  has both Jordan and Lebesgue measure zero. Defining  $a_{i_1^n, j} = \inf_{x_1^n \in S_{i_1^n}} \mathbf{1}_A(x_1^n) |g_j(x_1^n)|$  and  $b_{i_1^n, j} = \sup_{x_1^n \in S_{i_1^n}} |g_j(x_1^n)|$ , we may obtain lower and upper bounds to  $d_g$  by applying Lemma 22 to each term within the summation:

$$\begin{aligned}
&\sum_{i_1^n} \left( \prod_{j=1}^n \Delta_{i_1^n, j} \right) \mathbf{P}(X_1^n \in S_{i_1^n}) f(\chi_{i_1^n} | X_1^n \in S_{i_1^n}) \sum_{j=1}^n \frac{a_{i_1^n, j}^2 \Delta_{i_1^n, j}^2}{12} \leq d_g \\
&\leq \sum_{i_1^n} \left( \prod_{j=1}^n \Delta_{i_1^n, j} \right) \mathbf{P}(X_1^n \in S_{i_1^n}) f(\xi_{i_1^n} | X_1^n \in S_{i_1^n}) \sum_{j=1}^n \frac{b_{i_1^n, j}^2 \Delta_{i_1^n, j}^2}{12}.
\end{aligned}$$

Let  $K_j$  be the number of cells in the quantizer for  $X_j$ . For any cell  $S$  of this quantizer,  $\int_S \lambda_j(x_j) dx_j = 1/K_j$ . By continuity of  $\lambda_j$  and the first mean value theorem, this implies that the length of interval  $S$  is given by  $(K_j \lambda_j(\eta))^{-1}$  for some  $\eta \in S$ . Therefore,  $\Delta_{i_1^n, j}$  in the above expression may be replaced by  $(K_j \lambda_j(\eta_{i_1^n}))$  for some  $\eta_{i_1^n} \in S_{i_1^n}$ :

$$\begin{aligned}
&\sum_{i_1^n} \left( \prod_{j=1}^n \Delta_{i_1^n, j} \right) \mathbf{P}(X_1^n \in S_{i_1^n}) f(\chi_{i_1^n} | X_1^n \in S_{i_1^n}) \sum_{j=1}^n \frac{a_{i_1^n, j}^2}{12 K_j^2 \lambda_j(\eta_{i_1^n})^2} \leq d_g \\
&\leq \sum_{i_1^n} \left( \prod_{j=1}^n \Delta_{i_1^n, j} \right) \mathbf{P}(X_1^n \in S_{i_1^n}) f(\xi_{i_1^n} | X_1^n \in S_{i_1^n}) \sum_{j=1}^n \frac{b_{i_1^n, j}^2}{12 K_j^2 \lambda_j(\eta_{i_1^n})^2}.
\end{aligned}$$

Furthermore, we may recognize that  $\mathbf{P}(X_1^n \in S_{i_1^n}) f(\xi_{i_1^n} | X_1^n \in S_{i_1^n}) = f(\xi_{i_1^n})$ , simplifying the bounds further:

$$\sum_{i_1^n} f(\chi_{i_1^n}) \sum_{j=1}^n \frac{a_{i_1^n, j}^2}{12 K_j^2 \lambda_j(\eta_{i_1^n})^2} \prod_{j=1}^n \Delta_{i_1^n, j} \leq d_g \leq \sum_{i_1^n} f(\xi_{i_1^n}) \sum_{j=1}^n \frac{b_{i_1^n, j}^2}{12 K_j^2 \lambda_j(\eta_{i_1^n})^2} \prod_{j=1}^n \Delta_{i_1^n, j}.$$

Consider the  $j$ th term in the lower-bound summation,

$$\sum_{i_1^n} f(\chi_{i_1^n}) \frac{a_{i_1^n, j}^2}{12 \lambda_j(\eta_{i_1^n})} \prod_{i=1}^n \Delta_{i_1^n, i}.$$

One may observe that this expression approaches a Riemann integral:

- 1) By Lemma 22,  $\chi_{i_1^n} \in S_{i_1^n}$ .
- 2) By definition,  $a_{i_1^n, j}$  is the minimal value of  $\mathbf{1}_A(x_1^n) |g_j(x_1^n)|$  within the cell  $S_{i_1^n}$ .
- 3) By its definition,  $\eta_{i_1^n}$  is also an element in  $S_{i_1^n}$ .

- 4) The product  $\prod_{j=1}^n \Delta_{i_1^n, j}$  is the size of the cell, and because the largest quantizer cell size goes to zero as every element of the vector  $\mathbf{K}$  grows, the mesh of this summation also goes to zero.

Since  $A$  is Jordan-measurable,  $\mathbf{1}_A(x_1^n)$  is Riemann integrable. By assumption MF2, the expression  $f(x_1^n)g_j(x_1^n)^2 / (12\lambda_j(x_j)^2)$  is Riemann integrable. Since the product of two integrable functions is integrable, the Riemann integral of  $\mathbf{1}_A(x_1^n)f(x_1^n)g_j(x_1^n)^2 / (12\lambda_j(x_j)^2)$  is defined and

$$\begin{aligned} \sum_{i_1^n} f(\chi_{i_1^n}) \frac{a_{i_1^n, j}^2}{12\lambda_j(\eta_{i_1^n})} \prod_{i=1}^n \Delta_{i_1^n, j} &\sim \int_{[0,1]^n} f(x_1^n) \mathbf{1}_A(x_1^n) \frac{g_j(x_1^n)^2}{12\lambda_j(x_j)^2} dx_1^n \\ &= \int_A f(x_1^n) \frac{g_j(x_1^n)^2}{12\lambda_j(x_j)^2} dx_1^n \\ &\stackrel{(a)}{=} \int_{[0,1]^n} f(x_1^n) \frac{g_j(x_1^n)^2}{12\lambda_j(x_j)^2} dx_1^n \\ &\stackrel{(b)}{=} \int_{[0,1]} f(x_j) \frac{\gamma_j(x_j)^2}{12\lambda_j(x_j)^2} dx_j, \end{aligned}$$

where (a) follows from the Jordan measure of  $[0, 1]^n \setminus A$  being zero; and (b) is the result of integrating over  $x_1^{j-1}$  and  $x_{j+1}^n$ . This relation then yields

$$\sum_{i_1^n} f(\chi_{i_1^n}) \frac{a_{i_1^n, j}^2}{12K_j^2 \lambda_j(\eta_{i_1^n})} \prod_{i=1}^n \Delta_{i_1^n, j} \sim \int_{[0,1]} f(x_j) \frac{\gamma_j(x_j)^2}{12K_j^2 \lambda_j(x_j)^2} dx_j = \frac{1}{12K_j^2} \mathbf{E} \left[ \left( \frac{\gamma_j(X_j)}{\lambda_j(X_j)} \right)^2 \right].$$

Since this holds for any  $j \in \{1, \dots, n\}$ , it holds for the sum over  $j$  as well:

$$\sum_{i_1^n} f(\chi_{i_1^n}) \sum_{j=1}^n \frac{a_{i_1^n, j}^2}{12K_j^2 \lambda_j(\eta_{i_1^n})} \prod_{i=1}^n \Delta_{i_1^n, j} \sim \sum_{j=1}^n \frac{1}{12K_j^2} \mathbf{E} \left[ \left( \frac{\gamma_j(X_j)}{\lambda_j(X_j)} \right)^2 \right].$$

Similarly,

$$\sum_{i_1^n} f(\xi_{i_1^n}) \sum_{j=1}^n \frac{b_{i_1^n, j}^2}{12K_j^2 \lambda_j(\eta_{i_1^n})} \prod_{i=1}^n \Delta_{i_1^n, j} \sim \sum_{j=1}^n \frac{1}{12K_j^2} \mathbf{E} \left[ \left( \frac{\gamma_j(X_j)}{\lambda_j(X_j)} \right)^2 \right].$$

Since  $d_g$  is bounded between these two quantities, this proves the theorem.  $\blacksquare$

## APPENDIX B PROOF OF LEMMA 12

We start by defining the distortion-resolution optimization function  $\mathbf{K}_{\text{fr}}(R; \lambda)$  as the resolution vector that minimizes  $d_{\text{fr}}(\mathbf{K}_{\text{fr}}(R; \lambda); \lambda)$ , the distortion subject to a fixed rate constraint. We define  $\mathbf{K}_{\text{vr}}(R; \lambda)$  and  $\mathbf{K}_{\text{sw}}(R; \lambda)$  analogously, and we write  $\mathbf{K}_{\text{fr, vr, sw}}$  when we can combine all three cases to be handled identically. We similarly define the high-resolution distortion-resolution optimizing function  $\mathbf{K}_{\text{fr, vr, sw}}^{\text{HR}}(R; \lambda)$  as the resolution vector that minimizes  $d_{\text{fr, vr, sw}}^{\text{HR}}(\mathbf{K}_{\text{fr, vr, sw}}^{\text{HR}}(R; \lambda); \lambda)$  under a rate constraint. Note that by definition  $d_{\text{fr, vr, sw}}(\mathbf{K}_{\text{fr, vr, sw}}(R; \lambda); \lambda) = D_{\text{fr, vr, sw}}(R; \lambda)$ .

*Lemma 23:* Under assumptions MF1–4, every component of the vectors  $\mathbf{K}_{\text{fr, vr, sw}}(R; \lambda)$  and  $\mathbf{K}_{\text{fr, vr, sw}}^{\text{HR}}(R; \lambda)$  diverges with increasing  $R$ .

*Proof:* For a function, source, and quantizer point density that together satisfy conditions MF1–4, we demonstrate that every component of both  $\mathbf{K}_{\text{fr, vr, sw}}(R; \lambda)$  and  $\mathbf{K}_{\text{fr, vr, sw}}^{\text{HR}}(R; \lambda)$  diverges. Suppose first that the  $j$ th element of  $\mathbf{K}_{\text{fr, vr, sw}}(R; \lambda)$  is bounded by a finite value  $K$  for any  $R$ . Then the quantizer  $Q_K^{\lambda_j}$  is a sufficient description of  $X_j$  for achieving arbitrarily small distortion for the function  $g(X_1^n)$ . More precisely, there exists a reconstruction function  $\hat{g}$  such that  $g(X^n) = \hat{g}(X_1^{j-1}, Q_K^{\lambda_j}(X_j), X_{j+1}^n)$  with probability one. This then implies that  $\gamma_j(X_j)$  is zero with probability one, but this violates condition MF4 and thus every component of  $\mathbf{K}_{\text{fr, vr, sw}}(R; \lambda)$  diverges with  $R$ .

If the  $j$ th component of  $\mathbf{K}_{\text{fr, vr, sw}}^{\text{HR}}$  has a finite upper bound  $K$ , then the high-resolution distortion is lower bounded by

$$d_{\text{fr, vr, sw}}^{\text{HR}}(\mathbf{K}; \lambda) \geq \frac{1}{12K^2} \mathbf{E} \left[ \left( \frac{\gamma_j(X_j)}{\lambda_j(X_j)} \right)^2 \right].$$

By condition MF4, this lower bound is strictly positive, and therefore this choice of  $\mathbf{K}_{\text{fr, vr, sw}}^{\text{HR}}$  is suboptimal.  $\blacksquare$

Using this lemma, we are able to connect the distortion-rate function to the high-resolution distortion-resolution function.

*Lemma 24:* The distortion-rate function is asymptotically equal to the optimized high-resolution distortion-resolution function:  $D_{\text{fr, vr, sw}}(R; \lambda) \sim d_{\text{fr, vr, sw}}^{\text{HR}}(\mathbf{K}_{\text{fr, vr, sw}}^{\text{HR}}(R; \lambda); \lambda)$ .

*Proof:* Since by Lemma 23 both  $\mathbf{K}_{\text{fr, vr, sw}}$  and  $\mathbf{K}_{\text{fr, vr, sw}}^{\text{HR}}$  diverge in every component, Theorem 9 tells us that

$$d_{\text{fr, vr, sw}}(\mathbf{K}_{\text{fr, vr, sw}}(R; \boldsymbol{\lambda}); \boldsymbol{\lambda}) \sim d_{\text{fr, vr, sw}}^{\text{HR}}(\mathbf{K}_{\text{fr, vr, sw}}(R; \boldsymbol{\lambda}); \boldsymbol{\lambda})$$

and

$$d_{\text{fr, vr, sw}}(\mathbf{K}_{\text{fr, vr, sw}}^{\text{HR}}(R; \boldsymbol{\lambda}); \boldsymbol{\lambda}) \sim d_{\text{fr, vr, sw}}^{\text{HR}}(\mathbf{K}_{\text{fr, vr, sw}}^{\text{HR}}(R; \boldsymbol{\lambda}); \boldsymbol{\lambda}).$$

Furthermore, by definition we have that

$$d_{\text{fr, vr, sw}}(\mathbf{K}_{\text{fr, vr, sw}}(R; \boldsymbol{\lambda}); \boldsymbol{\lambda}) \leq d_{\text{fr, vr, sw}}(\mathbf{K}_{\text{fr, vr, sw}}^{\text{HR}}(R; \boldsymbol{\lambda}); \boldsymbol{\lambda})$$

and

$$d_{\text{fr, vr, sw}}^{\text{HR}}(\mathbf{K}_{\text{fr, vr, sw}}^{\text{HR}}(R; \boldsymbol{\lambda}); \boldsymbol{\lambda}) \leq d_{\text{fr, vr, sw}}^{\text{HR}}(\mathbf{K}_{\text{fr, vr, sw}}(R; \boldsymbol{\lambda}); \boldsymbol{\lambda}).$$

Therefore,

$$D_{\text{fr, vr, sw}}(R; \boldsymbol{\lambda}) = d_{\text{fr, vr, sw}}(\mathbf{K}_{\text{fr, vr, sw}}(R; \boldsymbol{\lambda}); \boldsymbol{\lambda}) \sim d_{\text{fr, vr, sw}}^{\text{HR}}(\mathbf{K}_{\text{fr, vr, sw}}^{\text{HR}}(R; \boldsymbol{\lambda}); \boldsymbol{\lambda}).$$

■

Before proceeding with the proof, we define three countably infinite subsets of  $\mathbb{R}^n$  that describe the rate vectors achievable by a certain choice of point densities  $\boldsymbol{\lambda}$ :

$$\mathcal{R}_{\text{fr}} = \{(\log K_1, \log K_2, \dots, \log K_n) : \mathbf{K} \in \mathbb{N}^n\}, \quad (34)$$

$$\mathcal{R}_{\text{vr}} = \{(H(Q_{K_1}^{\lambda_1}(X_1)), H(Q_{K_2}^{\lambda_2}(X_2)), \dots, H(Q_{K_n}^{\lambda_n}(X_n))) : \mathbf{K} \in \mathbb{N}^n\}, \quad (35)$$

$$\mathcal{R}_{\text{sw}} = \{(H(Q_{K_1}^{\lambda_1}(X_1)), H(Q_{K_2}^{\lambda_2}(X_2) | Q_{K_1}^{\lambda_1}(X_1)), \dots, H(Q_{K_n}^{\lambda_n}(X_n) | Q_{K_{n-1}}^{\lambda_{n-1}}(X_{n-1}))) : \mathbf{K} \in \mathbb{N}^n\}. \quad (36)$$

Using these definitions and Lemma 24, we may rephrase the distortion-rate functions somewhat:

$$D_{\text{fr}}(R; \boldsymbol{\lambda}) \sim d_{\text{fr}}^{\text{HR}}(\mathbf{K}_{\text{fr}}^{\text{HR}}(R; \boldsymbol{\lambda}); \boldsymbol{\lambda}) = \min_{\mathbf{R} \in \mathcal{R}_{\text{fr}}: \sum R_j \leq R} \sum_{j=1}^n \frac{1}{12} 2^{-2R_j} \mathbf{E} \left[ \left( \frac{\gamma_j(X_j)}{\lambda_j(X_j)} \right)^2 \right] \quad (37)$$

$$D_{\text{vr}}(R; \boldsymbol{\lambda}) \sim d_{\text{vr}}^{\text{HR}}(\mathbf{K}_{\text{vr}}^{\text{HR}}(R; \boldsymbol{\lambda}); \boldsymbol{\lambda}) = \min_{\mathbf{R} \in \mathcal{R}_{\text{vr}}: \sum R_j \leq R} \sum_{j=1}^n \frac{1}{12} 2^{-2R_j + 2h(X_j) + 2\mathbf{E}[\log \lambda_j(X_j)]} \mathbf{E} \left[ \left( \frac{\gamma_j(X_j)}{\lambda_j(X_j)} \right)^2 \right] \quad (38)$$

$$D_{\text{sw}}(R; \boldsymbol{\lambda}) \sim d_{\text{sw}}^{\text{HR}}(\mathbf{K}_{\text{sw}}^{\text{HR}}(R; \boldsymbol{\lambda}); \boldsymbol{\lambda}) = \min_{\mathbf{R} \in \mathcal{R}_{\text{sw}}: \sum R_j \leq R} \sum_{j=1}^n \frac{1}{12} 2^{-2R_j + 2h(X_j | X^{j-1}) + 2\mathbf{E}[\log \lambda_j(X_j)]} \mathbf{E} \left[ \left( \frac{\gamma_j(X_j)}{\lambda_j(X_j)} \right)^2 \right] \quad (39)$$

Additionally, we introduce the concept of *increasing granularity*:

*Definition 10:* A countably infinite set  $\mathcal{R} \subset \mathbb{R}^n$  is said to be *increasingly granular* if for any  $r \geq 0$  there exists a vanishing nonnegative function  $\delta(r) : [0, \infty) \rightarrow [0, \infty)$  such that for any  $\mathbf{R} \in \mathbb{R}^n$  whose components are each greater than  $r$ , there exists a point  $\bar{\mathbf{R}} \in \mathcal{R}$  within  $\delta(r)$  of each component of  $\mathbf{R}$ :  $\max_j |R_j - \bar{R}_j| \leq \delta(r)$ . The function  $\delta(r)$  is called the *granularity function* of the set  $\mathcal{R}$ .

*Lemma 25:* The sets  $\mathcal{R}_{\text{fr}}$ ,  $\mathcal{R}_{\text{vr}}$ , and  $\mathcal{R}_{\text{sw}}$  are increasingly granular.

*Proof:* Let  $r > 0$ , and let every component of  $\mathbf{R} \in \mathbb{R}^n$  be greater than  $r$ . We prove the granularity of each of the three sets in turn.

$\mathcal{R}_{\text{fr}}$ : Define the point  $\bar{\mathbf{R}} \in \mathbb{R}^n$  so that  $\bar{R}_j = \log \lfloor 2^{R_j} \rfloor$ . This point is clearly a member of  $\mathcal{R}_{\text{fr}}$ . Furthermore, we can easily bound the distance between  $R_j$  and  $\bar{R}_j$ :

$$|R_j - \bar{R}_j| \leq \log \frac{2^{R_j}}{2^{\bar{R}_j} - 1} \leq \log \frac{2^r}{2^r - 1}.$$

Defining  $\delta(r) = \log \left( \frac{2^r}{2^r - 1} \right) \rightarrow 0$ , we have shown that  $\mathcal{R}_{\text{fr}}$  is increasingly granular.

$\mathcal{R}_{\text{vr}}$ : Define  $\bar{\mathbf{R}}$  so that  $\bar{R}_j = H(Q_{K_j}^{\lambda_j}(X_j))$  where  $K_j$  is chosen according to

$$K_j = \arg \min_K |R_j - h(X_j) - \mathbf{E}[\log \lambda_j] - \log K|.$$

We may then bound the distance between  $R_j$  and  $\bar{R}_j$ :

$$\begin{aligned} |R_j - \bar{R}_j| &\leq |\bar{R}_j - h(X_j) - \mathbf{E}[\log \lambda_j] - \log K_j| + |h(X_j) + \mathbf{E}[\log \lambda_j] + \log K_j - R_j| \\ &\leq |\bar{R}_j - h(X_j) - \mathbf{E}[\log \lambda_j] - \log K_j| + \log \frac{K_j}{K_j - 1} \\ &\rightarrow 0, \end{aligned}$$

where the first term goes to zero by Lemma 10 and the second by Lemma 23.



$\mathcal{R}_{\text{sw}}$ : Define  $\bar{\mathbf{R}}$  so that  $\bar{R}_j = H(Q_{K_j}^{\lambda_j}(X_j)|Q_{K_j}^{\lambda_j}(X^{j-1}))$  where  $K_j$  is chosen according to

$$K_j = \arg \min_K |R_j - h(X_j | X^{j-1}) - \mathbf{E}[\log \lambda_j] - \log K|.$$

The distance between  $R_j$  and  $\bar{R}_j$  may then be bounded in the following manner:

$$\begin{aligned} |R_j - \bar{R}_j| &\leq |\bar{R}_j - h(X_j|X^{j-1}) - \mathbf{E}[\log \lambda_j] - \log K_j| + |h(X_j|X^{j-1}) + \mathbf{E}[\log \lambda_j] + \log K_j - R_j| \\ &\leq |\bar{R}_j - h(X_j|X^{j-1}) - \mathbf{E}[\log \lambda_j] - \log K_j| + \log \frac{K_j}{K_j - 1} \\ &\rightarrow 0. \end{aligned}$$

To show that the first term goes to zero, we invoke Lemma 10 to state that

$$H(Q_{K_j}^{\lambda_j}(X_j), Q_{K_j}^{\lambda_j}(X^{j-1})) - \sum_{i=1}^j \log K_i \rightarrow h(X^j) + \mathbf{E} \left[ \sum_{i=1}^j \log \lambda_i(X_i) \right].$$

Subtracting from this the similar expression (also obtained from Lemma 10)

$$H(Q_{K_j}^{\lambda_j}(X^{j-1})) - \sum_{i=1}^{j-1} \log K_i \rightarrow h(X^{j-1}) + \mathbf{E} \left[ \sum_{i=1}^{j-1} \log \lambda_i(X_i) \right],$$

yields that

$$\bar{R}_j - h(X_j | X^{j-1}) - \mathbf{E}[\log \lambda_j] - \log K_j = H(Q_{K_j}^{\lambda_j}(X_j)|Q_{K_j}^{\lambda_j}(X^{j-1})) - h(X_j | X^{j-1}) - \mathbf{E}[\log \lambda_j] - \log K_j \rightarrow 0. \quad \blacksquare$$

We now establish an important property of increasingly granular sets.

*Lemma 26:* Suppose

$$f(R) = \min_{\mathbf{R} \in \mathcal{R}: \sum_{j=1}^n R_j \leq R} \sum_{j=1}^n \alpha_j 2^{-R_j},$$

and

$$\tilde{f}(R) = \min_{\mathbf{R} \in \mathbb{R}^n: \sum_{j=1}^n R_j \leq R} \sum_{j=1}^n \alpha_j 2^{-R_j},$$

where  $\alpha_j > 0$  for all  $j$  and  $\mathcal{R}$  is an increasingly granular subset of  $\mathbb{R}^n$ . Then  $f(R) \sim \tilde{f}(R)$ .

*Proof:* Since  $\mathcal{R} \subset \mathbb{R}^n$ , we have that

$$\tilde{f}(R) \leq f(R).$$

Let  $\mathbf{R}^* = \arg \min_{\mathbf{R} \in \mathbb{R}^n: \sum_{j=1}^n R_j \leq R} \sum_{j=1}^n \alpha_j 2^{-R_j}$ , let  $R_{\text{inf}}^*$  indicate the smallest element of  $\mathbf{R}^*$ , and let  $\delta(r)$  be the granularity function of  $\mathcal{R}$ . Since there must be an element of  $\mathcal{R}$  within distance  $\delta_{R_{\text{inf}}^*}$  of each of the coordinates of  $\mathbf{R}^*$ , we may create a bound in the opposite direction:

$$f(R) \leq \sum_{j=1}^n \alpha_j 2^{-(R_j^* - \delta_{R_{\text{inf}}^*})} = 2^{\delta_{R_{\text{inf}}^*}} \tilde{f}(R).$$

Because  $\alpha_j > 0$  for  $j \in \{1, \dots, n\}$ ,  $R_{\text{inf}}^*$  diverges with  $R$  and  $\delta_{R_{\text{inf}}^*}$  vanishes. Combining the two bounds, we have that  $\tilde{f}(R) \sim f(R)$ , which proves the lemma.  $\blacksquare$

Applying Lemmas 25 and 26 to (37), (38), and (39), we may widen the optimization to occur over any positive real-valued rate vector  $\mathbf{R}$ :

$$D_{\text{fr}}(R; \boldsymbol{\lambda}) \sim d_{\text{fr}}^{\text{HR}}(\mathbf{K}_{\text{fr}}^{\text{HR}}(R; \boldsymbol{\lambda}); \boldsymbol{\lambda}) \sim \min_{\mathbf{R} \in \mathbb{R}^n: \sum R_j \leq R} \sum_{j=1}^n \frac{1}{12} 2^{-2R_j} \mathbf{E} \left[ \left( \frac{\gamma_j(X_j)}{\lambda_j(X_j)} \right)^2 \right] \quad (40)$$

$$D_{\text{vr}}(R; \boldsymbol{\lambda}) \sim d_{\text{vr}}^{\text{HR}}(\mathbf{K}_{\text{vr}}^{\text{HR}}(R; \boldsymbol{\lambda}); \boldsymbol{\lambda}) \sim \min_{\mathbf{R} \in \mathbb{R}^n: \sum R_j \leq R} \sum_{j=1}^n \frac{1}{12} 2^{-2R_j + 2h(X_j) + 2\mathbf{E}[\log \lambda_j(X_j)]} \mathbf{E} \left[ \left( \frac{\gamma_j(X_j)}{\lambda_j(X_j)} \right)^2 \right] \quad (41)$$

$$D_{\text{sw}}(R; \boldsymbol{\lambda}) \sim d_{\text{sw}}^{\text{HR}}(\mathbf{K}_{\text{sw}}^{\text{HR}}(R; \boldsymbol{\lambda}); \boldsymbol{\lambda}) \sim \min_{\mathbf{R} \in \mathbb{R}^n: \sum R_j \leq R} \sum_{j=1}^n \frac{1}{12} 2^{-2R_j + 2h(X_j|X^{j-1}) + 2\mathbf{E}[\log \lambda_j(X_j)]} \mathbf{E} \left[ \left( \frac{\gamma_j(X_j)}{\lambda_j(X_j)} \right)^2 \right] \quad (42)$$

The proof is completed by a straightforward application of Lemma 4 to optimize the rate allocation in each of these expressions.

APPENDIX C  
PROOF OF THEOREM 14

The theorem asserts that when the function is equivalence-free,  $w_j$  failing to be one-to-one on the support of  $X_j$  creates a component of the distortion that cannot be eliminated by quantizing more finely. The proof here lower-bounds the distortion by focusing on the contribution from just the  $j$ th variable. The bound is especially crude because it is based on observing  $\{X_i\}_{i \neq j}$  and  $w_j(X_j)$  without quantization and it uses only the contribution from  $X_j \in S \cup t(S)$ .

We wish to first bound the functional distortion in terms of a contribution from the  $j$ th variable:

$$\begin{aligned}
d_g(K; \mathbf{w}) &\stackrel{(a)}{\geq} \mathbf{E} \left[ \text{var}(g(X_1^n) \mid \widehat{Y}_1^n) \right] \\
&\stackrel{(b)}{\geq} \mathbf{E} \left[ \text{var}(g(X_1^n) \mid \widehat{Y}_j, \{X_i\}_{i \neq j}) \right] \\
&\stackrel{(c)}{\geq} \mathbf{E} \left[ \text{var}(g(X_1^n) \mid w_j(X_j), \{X_i\}_{i \neq j}) \right] \\
&\stackrel{(d)}{=} \mathbf{E} \left[ \text{var}(g(X_1^n) \mid w_j(X_j), \{X_i\}_{i \neq j} \mid A) \mathbf{P}(A) \right. \\
&\quad \left. + \text{var}(g(X_1^n) \mid w_j(X_j), \{X_i\}_{i \neq j} \mid A^c) \mathbf{P}(A^c) \right] \\
&\stackrel{(e)}{\geq} \mathbf{E} \left[ \text{var}(g(X_1^n) \mid w_j(X_j), \{X_i\}_{i \neq j} \mid A) \mathbf{P}(A) \right] \\
&= \mathbf{E} \left[ \text{var}(g(X_1^n) \mid X_j \in w_j^{-1}(X_j), \{X_i\}_{i \neq j} \mid A) \mathbf{P}(A) \right] \\
&\stackrel{(f)}{=} \int_{x \in S \cup t(S)} \mathbf{E} \left[ \text{var}(g(X_1^n) \mid X_j \in w_j^{-1}(x), \{X_i\}_{i \neq j}) \right] dx, \tag{43}
\end{aligned}$$

where  $A$  is the event  $X_j \in S \cup t(S)$ . Step (a) will hold with equality when the optimal estimate (the conditional expectation of  $g(X_1^n)$  given the quantized values) is used; (b) holds because, for each  $i \neq j$ ,  $\widehat{Y}_i$  is a function of  $X_i$ ; (c) holds because  $\widehat{Y}_j$  is a function of  $w_j(X_j)$ ; (d) is an application of the law of total expectation; (e) holds because the discarded term is nonnegative; and (f) converts the expectation over  $A$  into integral form. It remains to use the hypotheses of the theorem to bound the conditional variance in the final expression.

Since the function is equivalence free, for every set  $B \subset [0, 1]$  of cardinality greater than one,

$$\mathbf{E} \left[ \text{var}(g(X_1^n) \mid X_j \in B, \{X_i\}_{i \neq j}) \right] > 0.$$

Since  $w_j(s) = w_j(t(s))$  for any  $s \in S$ , the set  $w_j^{-1}(x_j)$  is of cardinality greater than one for any  $x_j$  in  $S \cup t(S)$ . Therefore for any  $x \in S \cup t(S)$ ,

$$\mathbf{E} \left[ \text{var}(g(X_1^n) \mid X_j \in w_j^{-1}(x), \{X_i\}_{i \neq j}) \right] > 0,$$

and (43) is therefore greater than zero and independent of rate.

APPENDIX D  
PROOF OF THEOREM 16

It is already shown in Theorem 15 that the distortion-resolution expression (26) holds when a codeword is allocated to each of the don't-care intervals. After an appropriate rate analysis, we will optimize the point densities outside of the don't-care intervals.

The key technical problem is that the rate analysis (6) does not hold when there are intervals where  $f_X$  is positive but  $\lambda$  is not. This is easily remedied by only applying (6) conditioned on  $A_j$ :

$$\lim_{K_j \rightarrow \infty} \left[ H(\widehat{X}_j \mid A_j) - \log(K_j - M_j) \right] = h(X_j \mid A_j) + \mathbf{E} \left[ \log \lambda_j(X_j) \mid A_j \right]. \tag{44}$$

Note that this approximation can be shown to be asymptotically valid in the same manner as in Lemmas 1 and 2. Now conditioned on  $A_j$ , the dependence of distortion and rate on  $\lambda_j$  is precisely in the standard form of Section IV. Thus, following Theorem 13, the optimal point density outside of  $Z_j$  is given by (22).

Since the previous results now give the distortion in terms of the conditional entropies  $H(\widehat{X}_j \mid A_j)$ , what remains is to relate these to the rates:

$$\begin{aligned}
R_j &= H(\widehat{X}_j) \\
&\stackrel{(a)}{=} H(\widehat{X}_j, I_j) \\
&= H(I_j) + H(\widehat{X}_j \mid I_j) \\
&\stackrel{(b)}{=} H(I_j) + \mathbf{P}(A_j) H(\widehat{X}_j \mid A_j),
\end{aligned}$$

where (a) uses that  $I_j$  is a deterministic function of  $\widehat{X}_j$ ; and (b) uses that specifying any  $I_j \neq 0$  determines  $\widehat{X}_j$  uniquely. In anticipation of evaluating (26), we define the high-rate resolution-rate function as before:

$$\log(K_j^{\text{HR}}(R_j; \lambda_j) - M_j) \sim (\mathbf{P}(A_j))^{-1} (R_j - H(I_j)) - h(X_j | A_j) - \mathbf{E}[\log(\lambda_j(X_j) | A_j)].$$

Asymptotic accuracy of this approximation follows from (44). As before, one may insert this into the high-resolution distortion-resolution expression (26) and bound the effect of the approximation as a multiplying factor that goes to one. Now evaluating (26) with optimal point densities (22) gives (28).

#### APPENDIX E A QUASI-TRIANGLE INEQUALITY

*Lemma 27:* The  $\mathcal{L}^{1/3}$  “norm” is a quasinorm with constant 4. Equivalently, letting  $x$  and  $y$  be functions  $\mathbb{R} \rightarrow \mathbb{R}^+$  with finite  $\mathcal{L}^{1/3}$  quasinorms,

$$\|x + y\|_{1/3} \leq 4 (\|x\|_{1/3} + \|y\|_{1/3}).$$

*Proof:* First, we prove the relation  $4(a^3 + b^3) \geq (a + b)^3$  for positive real numbers  $a$  and  $b$ :

$$\begin{aligned} 4(a^3 + b^3) - (a + b)^3 &= 4a^3 + 4b^3 - a^3 - b^3 - 3a^2b - 3ab^2 \\ &= 3(a + b)(a - b)^2 \geq 0. \end{aligned}$$

Now by this relation, with  $a = \int x(t)^{1/3} dt$  and  $b = \int y(t)^{1/3} dt$ :

$$\begin{aligned} \|x\|_{1/3} + \|y\|_{1/3} &= \left( \int x(t)^{1/3} dt \right)^3 + \left( \int y(t)^{1/3} dt \right)^3 \\ &\geq \frac{1}{4} \left( \int (x(t)^{1/3} + y(t)^{1/3}) dt \right)^3 \\ &\geq \frac{1}{4} \left( \int ((x(t) + y(t))^{1/3}) dt \right)^3 \\ &= \frac{1}{4} \|x + y\|_{1/3}, \end{aligned}$$

where the second inequality uses, pointwise over  $t$ , the concavity of the cube-root function on  $[0, \infty)$ . ■

#### ACKNOWLEDGMENTS

The authors would like to thank John Sun for his insights and suggestions about both the problem and the manuscript. They would also like to thank both the reviewers and the Associate Editor, Erik Ordentlich, for helping to significantly improve the clarity and rigor of the results presented here.

#### REFERENCES

- [1] V. Doshi, D. Shah, and M. Médard, “Source coding with distortion through graph coloring,” in *Proc. IEEE Int. Symp. Inform. Theory (ISIT 2007)*, Jun. 2007, pp. 1501–1505.
- [2] D. Slepian and J. K. Wolf, “Noiseless coding of correlated information sources,” *IEEE Trans. Inf. Theory*, vol. IT-19, no. 4, pp. 471–480, Jul. 1973.
- [3] S. D. Servetto, “Achievable rates for multiterminal source coding with scalar quantizers,” in *Conf. Rec. 39th Asilomar Conf. Signals, Syst. Comput.*, Oct. 2005, pp. 1762–1766.
- [4] A. B. Wagner, S. Tavildar, and P. Viswanath, “Rate region of the quadratic Gaussian two-terminal source-coding problem,” *IEEE Trans. Inf. Theory*, vol. 54, no. 5, pp. 1938–1961, May 2008.
- [5] R. Zamir and T. Berger, “Multiterminal source coding with high resolution,” *IEEE Trans. Inf. Theory*, vol. 45, no. 1, pp. 106–117, Jan. 1999.
- [6] A. D. Wyner and J. Ziv, “The rate-distortion function for source coding with side information at the decoder,” *IEEE Trans. Inf. Theory*, vol. IT-22, no. 1, pp. 1–10, Jan. 1976.
- [7] R. Zamir, “The rate loss in the Wyner-Ziv problem,” *IEEE Trans. Inf. Theory*, vol. 42, no. 6, pp. 2073–2084, Nov. 1996.
- [8] D. Rebollo-Monedero, S. Rane, A. Aaron, and B. Girod, “High-rate quantization and transform coding with side information at the decoder,” *Signal Process.*, vol. 86, no. 11, pp. 3160–3179, Nov. 2006.
- [9] Z. Liu, S. Cheng, A. D. Liveris, and Z. Xiong, “Slepian-Wolf coded nested lattice quantization for Wyner-Ziv coding: High-rate performance analysis and code design,” *IEEE Trans. Inf. Theory*, vol. 52, no. 10, pp. 4358–4379, Oct. 2006.
- [10] E. Martinian, G. Wornell, and R. Zamir, “Source coding with encoder side information,” *IEEE Trans. Inf. Theory*, vol. 54, no. 10, pp. 4638–4665, Oct. 2008.
- [11] T. Linder, R. Zamir, and K. Zeger, “On source coding with side-information-dependent distortion measures,” *IEEE Trans. Inf. Theory*, vol. 46, no. 7, pp. 2697–2704, Nov. 2000.
- [12] H. Yamamoto, “Wyner-Ziv theory for a general function of the correlated sources,” *IEEE Trans. Inf. Theory*, vol. IT-28, no. 5, pp. 803–807, Sep. 1982.
- [13] H. Feng, M. Effros, and S. A. Savari, “Functional source coding for networks with receiver side information,” in *Proc. 42nd Annu. Allerton Conf. Commun. Control Comput.*, Sep. 2004, pp. 1419–1427.
- [14] A. Orlicsky and J. R. Roche, “Coding for computing,” *IEEE Trans. Inf. Theory*, vol. 47, no. 3, pp. 903–917, Mar. 2001.

- [15] T. S. Han and K. Kobayashi, "A dichotomy of functions  $F(X, Y)$  of correlated sources  $(X, Y)$  from the viewpoint of the achievable rate region," *IEEE Trans. Inf. Theory*, vol. IT-33, no. 1, pp. 69–76, Jan. 1987.
- [16] V. Doshi, D. Shah, M. Médard, and S. Jaggi, "Distributed functional compression through graph coloring," in *Proc. IEEE Data Compression Conf. (DCC 2007)*, Mar. 2007, pp. 93–102.
- [17] H. Yamamoto and K. Itoh, "Source coding theory for multiterminal communication systems with a remote source," *Trans. IECE Japan*, vol. E63, no. 10, pp. 700–706, Oct. 1980.
- [18] J. Körner and K. Marton, "How to encode the modulo-two sum of binary sources," *IEEE Trans. Inf. Theory*, vol. IT-25, no. 2, pp. 219–221, Mar. 1979.
- [19] H. V. Poor, "Fine quantization in signal detection and estimation," *IEEE Trans. Inf. Theory*, vol. 34, no. 5, pp. 960–972, Sep. 1988.
- [20] G. R. Benitz and J. A. Bucklew, "Asymptotically optimal quantizers for detection of i.i.d. data," *IEEE Trans. Inf. Theory*, vol. 35, no. 2, pp. 316–325, Mar. 1989.
- [21] R. Gupta and A. O. Hero, III, "High-rate vector quantization for detection," *IEEE Trans. Inf. Theory*, vol. 49, no. 8, pp. 1951–1969, Aug. 2003.
- [22] J. Li, N. Chaddha, and R. M. Gray, "Asymptotic performance of vector quantizers with a perceptual distortion measure," *IEEE Trans. Inf. Theory*, vol. 45, no. 4, pp. 1082–1091, May 1999.
- [23] T. Linder, R. Zamir, and K. Zeger, "High-resolution source coding for non-difference distortion measures: Multidimensional companding," *IEEE Trans. Inf. Theory*, vol. 45, no. 2, pp. 548–561, Mar. 1999.
- [24] J. A. Bucklew, "Multidimensional digitization of data followed by a mapping," *IEEE Trans. Inf. Theory*, vol. IT-30, no. 1, pp. 107–110, Jan. 1984.
- [25] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*. Boston: Kluwer Academic Publishers, 1992.
- [26] A. Gyorgy and T. Linder, "On the structure of optimal entropy-constrained scalar quantizers," *IEEE Trans. Inf. Theory*, vol. 48, no. 2, pp. 416–427, Feb. 2002.
- [27] T. Linder, "On asymptotically optimal companding quantization," *Probl. Control Inf. Theory*, vol. 20, no. 6, pp. 383–393, 1991.
- [28] R. M. Gray and D. L. Neuhoff, "Quantization," *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2325–2383, Oct. 1998.
- [29] R. M. Gray and A. H. Gray, Jr., "Asymptotically optimal quantizers," *IEEE Trans. Inf. Theory*, vol. 23, no. 1, pp. 143–144, Jan. 1977.
- [30] J. J. Y. Huang and P. M. Schultheiss, "Block quantization of correlated Gaussian random variables," *IEEE Trans. Commun. Syst.*, vol. CS-11, no. 3, pp. 289–296, Sep. 1963.
- [31] B. Farber and K. Zeger, "Quantization of multiple sources using nonnegative integer bit allocation," *IEEE Trans. Inf. Theory*, vol. 52, no. 11, pp. 4945–4964, Nov. 2006.
- [32] I. Csiszar, "Generalized entropy and quantization problems," in *Trans. 6th Prague Conf. Information Theory, Statistical Decision Functions, Random Processes*, 1973, pp. 29–35.
- [33] T. Linder and K. Zeger, "Asymptotic entropy-constrained performance of tessellating and universal randomized lattice quantization," *IEEE Trans. Inf. Theory*, vol. 40, no. 2, pp. 575–579, Mar. 1994.
- [34] M. Studený and J. Vejnarová, "The multiinformation function as a tool for measuring stochastic dependence," in *Learning in Graphical Models*, M. I. Jordan, Ed. Dordrecht: Kluwer Academic Publishers, 1998, pp. 261–297.
- [35] V. Misra, "Functional quantization," M. Eng. thesis, Massachusetts Institute of Technology, Cambridge, MA, Jun. 2008.
- [36] A. György, T. Linder, and K. Zeger, "On the rate-distortion function of random vectors and stationary sources with mixed distributions," *IEEE Trans. Inf. Theory*, vol. 45, no. 6, pp. 2110–2115, Sep. 1999.

PLACE  
PHOTO  
HERE

**Vinith Misra** received the S.B. and M.Eng. degrees in electrical engineering from the Massachusetts Institute of Technology in 2008, where his thesis was awarded the David Adler Memorial M.Eng. thesis prize. He is currently pursuing a Ph.D. in Stanford University's department of electrical engineering.

He is a Stanford Graduate Fellow and a recipient of the National Defense Science and Engineering Graduate Fellowship. His research interests include information theory, signal processing, and mixed-signal circuit design, with applications to both communications and medical devices.

PLACE  
PHOTO  
HERE

**Vivek K Goyal** (S'92-M'98-SM'03) received the B.S. degree in mathematics and the B.S.E. degree in electrical engineering from the University of Iowa, Iowa City, where he received the John Briggs Memorial Award for the top undergraduate across all colleges. He received the M.S. and Ph.D. degrees in electrical engineering from the University of California, Berkeley, where he received the Eliahu Jury Award for outstanding achievement in systems, communications, control, or signal processing.

He was a Member of Technical Staff in the Mathematics of Communications Research Department of Bell Laboratories, Lucent Technologies, 1998–2001; and a Senior Research Engineer for Digital Fountain, Inc., 2001–2003. He is currently Esther and Harold E. Edgerton Associate Professor of electrical engineering at the Massachusetts Institute of Technology. His research interests include source coding theory, sampling, quantization, and computational imaging.

Dr. Goyal is a member of Phi Beta Kappa, Tau Beta Pi, Sigma Xi, Eta Kappa Nu and SIAM. He was awarded the 2002 IEEE Signal Processing Society Magazine Award and an NSF CAREER Award. He served on the IEEE Signal Processing Society's Image and Multiple Dimensional Signal Processing Technical Committee. He is a Technical Program Committee Co-chair of IEEE ICIP

2016 and a permanent Conference Co-chair of the SPIE Wavelets conference series.

PLACE  
PHOTO  
HERE

**Lav R. Varshney** received the B. S. degree with honors in electrical and computer engineering (magna cum laude) from Cornell University, Ithaca, New York in 2004. He received the S. M., E. E., and Ph. D. degrees in electrical engineering and computer science from the Massachusetts Institute of Technology (MIT), Cambridge in 2006, 2008, and 2010, respectively.

He is a research staff member at the IBM Thomas J. Watson Research Center, Hawthorne, NY. He was a National Science Foundation graduate research fellow and held various research and teaching positions at MIT. His research interests include information theory, coding, and neuroscience.

Dr. Varshney is a member of Tau Beta Pi, Eta Kappa Nu, and Sigma Xi. He received the E. A. Guillemin Thesis Award for Outstanding Electrical Engineering S.M. Thesis, the Capocelli Prize at the 2006 Data Compression Conference, the Best Student Paper Award at the 2003 IEEE Radar Conference, and was a winner of the IEEE 2004 Student History Paper Contest.