

Asymptotically Scale-invariant Multi-resolution Quantization

Cheuk Ting Li

Department of Information Engineering

The Chinese University of Hong Kong

Email: ctli@ie.cuhk.edu.hk

Abstract

A multi-resolution quantizer is a sequence of quantizers where the output of a coarser quantizer can be deduced from the output of a finer quantizer. In this paper, we propose an asymptotically scale-invariant multi-resolution quantizer, which performs uniformly across any choice of average quantization step, when the length of the range of input numbers is large. Scale invariance is especially useful in worst case or adversarial settings, ensuring that the performance of the quantizer would not be affected greatly by small changes of storage or error requirements. We also show that the proposed quantizer achieves a tradeoff between rate and error that is arbitrarily close to the optimum.

I. INTRODUCTION

A multi-resolution quantizer is a sequence of quantizers, where the output of a coarser quantizer can be deduced from the output of a finer quantizer (without knowledge of the original data). It has been studied, for example, by Koshelev [1], Equitz and Cover [2], Rimoldi [3], Brunk and Farvardin [4], Jafarkhani, Brunk, and Farvardin [5], Effros [6], Wu and Dumitrescu [7], [8] and Effros and Dugatkin [9]. There are two main uses of multi-resolution quantizers: to allow a coarser quantization to be obtained from a finer quantization by discarding some information, and to allow a finer quantization to be obtained from a coarser quantization by adding some additional information from the encoder (i.e., successive refinement). We first focus on the first usage.

Consider the setting where a piece of data is relayed across a sequence of nodes, where each communication link has a different capacity. Each node only has information about the capacity of its incoming and outgoing link, and therefore can only compress the incoming data according to the capacity of the outgoing link and send it to the next node, if the outgoing link has smaller capacity than the incoming link (otherwise the node can relay the incoming data exactly). If the data is a number $x \in [0, 1)$, a simple scheme, which we call the simple uniform quantizer, is that node i would apply the uniform quantization $y_{i+1} := (\lfloor k_i y_i \rfloor + 1/2)/k_i$ to its incoming data y_i and send it to the node $i + 1$, where k_i is the number of values that can be sent through the link between node i and $i + 1$. This scheme is undesirable since if $k_1 = 4$, $k_2 = 3$, and $x = y_1 = 2/7$, then $y_2 = 3/8$, $y_3 = 1/2$, giving an absolute error $|y_3 - x| = 3/14$ that is larger than as if $k_1 = k_2 = 3$ (i.e., the data is only compressed once according to the worse link), giving an absolute error $5/42$.

To mitigate this problem, we can use a multi-resolution quantizer, where the output of a coarser quantizer can be obtained from the output of a finer quantizer, and thus the final output of the relay would convey the same information as if the input is only compressed once according to the link with the lowest capacity (by simply quantizing the final output again according to the link with the lowest capacity). One simple scheme, which we call the binary multi-resolution quantizer (BMRQ), would be to quantize only using step sizes that are powers of 2, i.e., $y_{i+1} := 2^{-\lfloor \log_2 k_i \rfloor} (\lfloor 2^{\lfloor \log_2 k_i \rfloor} y_i \rfloor + 1/2)$. For the aforementioned example $k_1 = 4$, $k_2 = 3$, $x = 2/7$, we have $y_2 = 3/8$, $y_3 = 1/4$, giving an absolute error $|y_3 - x| = 1/28$.

Nevertheless, the BMRQ does not perform well in worst case or adversarial settings. Consider the setting where an adversary can modify k_i to increase the quantization error. If $k_1 = 32$ (the quantization step is $1/32$), then the adversary can reduce k_1 by 1 to 31, increasing the quantization step two-fold to $1/16$ (and hence the average absolute error is also increased two-fold). The BMRQ performs well only when the k_i 's are powers of 2. The adversary can modify k_i slightly off powers of 2 to cause a significant degradation of the quantized data. We call this problem scale dependence, meaning that the multi-resolution quantizer does not perform uniformly well for all choices of quantization step.

In this paper, we introduce an asymptotically scale-invariant multi-resolution quantizer, called the biased binary multi-resolution quantizer (BBMRQ), that performs uniformly across any choice of average quantization step (BBMRQ is a non-uniform quantizer), when the length of the range of input numbers tends to infinity. Therefore its performance degrades gracefully when the adversary modifies the communication constraints. We show that the BBMRQ outperforms the BMRQ except when the average quantization step is close to a power of 2 (see Figure 1).

Asymptotically scale-invariant multi-resolution quantizers are also useful in successive refinement settings. Consider the scenario where an encoder observes a number x and produces a sequence of bits. Due to storage or communication constraint,

we only keep the first n bits and discard the rest, where n is chosen according to the storage constraint and the bit sequence. The BMRQ corresponds to the scheme where the bit sequence is the binary representation of x , and n is chosen only according to the storage constraint. For the BBMRQ, n also depends on the bit sequence (i.e., it is a variable length code), which allows the performance of the quantizer to vary smoothly when the storage constraint changes. The bit sequence can be produced according to the quantization tree (see Figure 2).

This paper is organized as follows. In Section II, we give the criteria of multi-resolution quantizers and define the cell size cdf for measuring the performance of a quantizer. In Section IV, we define the BBMRQ and present the main result regarding the performance of BBMRQ. In Section V, we show that the BBMRQ achieves a tradeoff between rate and error that is arbitrarily close to the optimum.

A. Previous Work

The seminal work by Equitz and Cover [2] concerns the problem of successive refinement of information, where the lossy reconstruction is iteratively refined by supplying more information. It is also studied by Rimoldi [3]. Also see [1] for a related setting. Another line of research is multiple description coding [10], [11], [12], where several descriptions are produced from the same source, and the distortion of the reconstruction depends on which subset of descriptions is available to the decoder. Also see [13] for a related setting. Note that the aforementioned papers concern the asymptotic rate-distortion problem, whereas this paper focuses on the one-shot scalar quantization setting.

Vaishampayan [14] studied multiple description scalar quantizers. Brunk and Farvardin [4] and Jafarkhani, Brunk, and Farvardin [5] studied multi-resolution scalar quantizers (or successively refinable quantizers), and provide algorithms for designing quantizers with small error given the distribution of the input. Effros [6] and Effros and Dugatkin [9] studied multi-resolution vector quantizers. Algorithms for multi-resolution quantization were studied by Wu and Dumitrescu [7], [8]. Note that the aforementioned papers concern the setting where the distribution of the input is known, and the quantizers are designed accordingly (as in the classical Lloyd-Max algorithm [15], [16]). In this paper, we do not design the quantizer according to the input distribution, but rather assume the input is (loosely speaking) uniform over a long interval.

II. MULTI-RESOLUTION QUANTIZER

In this paper, quantizer can refer to any measurable function $Q : \mathbb{R} \rightarrow \mathbb{R}$, where the range $Q(\mathbb{R})$ is a finite or countable set. We call Q a centered quantizer if each reconstruction level is the center of its corresponding quantization cell, which is formally defined below.

Definition 1. We call a function $Q : \mathbb{R} \rightarrow \mathbb{R}$ a *centered quantizer* if it is non-decreasing, the range $Q(\mathbb{R})$ is a finite or countable, and

$$Q(x) = \frac{1}{2} \left(\inf\{y : Q(y) = Q(x)\} + \sup\{y : Q(y) = Q(x)\} \right)$$

for all $x \in \mathbb{R}$.

We give the criteria of multi-resolution quantizers below.

Definition 2. We call $\{Q_s\}_{s>0}$ a *multi-resolution quantizer (MRQ)* if the functions $Q_s : \mathbb{R} \rightarrow \mathbb{R}$ are measurable and satisfy

$$Q_{s_2}(Q_{s_1}(x)) = Q_{s_2}(x)$$

for any $s_2 \geq s_1 > 0$ and $x \in \mathbb{R}$. The parameter s (that can be any positive real number) usually corresponds to the (maximum or average) quantization step. We call $\{Q_s\}_{s>0}$ a *centered multi-resolution quantizer* if the functions Q_s are centered quantizers.

This definition ensures that $Q_{s_2}(x)$ can be computed using $Q_{s_1}(x)$ for $s_2 \geq s_1 > 0$, simply by quantizing $Q_{s_1}(x)$ using Q_{s_2} . As a result, if $s_n \geq s_i$ for all $i = 1, \dots, n$, then $Q_{s_n}(Q_{s_{n-1}}(\dots Q_{s_1}(x) \dots)) = Q_{s_n}(x)$, i.e., we can quantize the final output of the relay again by the coarsest quantizer to obtain a result the same as if the input is only compressed once by the coarsest quantizer.

We remark that this definition is different from the previous definitions (e.g. [4], [5], [7]), which also concern how the quantized number is represented (e.g. by a bit sequence). Here we only concern the mapping from the input number to its reconstruction level, and assume that a suitable compression algorithm is applied to the reconstruction levels if the multi-resolution quantizer is to be used in practice.

Note that the simple uniform quantizer $Q_s(x) := s(\lfloor x/s \rfloor + 1/2)$ is not a multi-resolution quantizer, since $Q_3(x)$ cannot be deduced from $Q_2(x)$. Nevertheless, if we restrict the step size to powers of 2, i.e., $Q_s^{\text{bin}}(x) := 2^{\lfloor \log_2 s \rfloor} (\lfloor 2^{-\lfloor \log_2 s \rfloor} x \rfloor + 1/2)$, then this is a MRQ, which we call the *binary multi-resolution quantizer (BMRQ)*.

A downside of the BMRQ is that it is scale-dependent. The average absolute error of the binary multi-resolution quantizer is $2^{\lfloor \log_2 s \rfloor - 2}$ (when the input is uniformly distributed over a long interval), which must be a power of 2. Therefore, the quantizer

is suitable if the maximum allowed average absolute error is a power of 2, but not suitable if the maximum allowed average absolute error is slightly smaller than a power of 2. Scale-dependence is particularly undesirable in worst case or adversarial settings, where the quantizer must work well for any maximum allowed average absolute error (or other error metrics).

Loosely speaking, the BMRQ is the optimal uniform multi-resolution quantizer (where for each s , Q_s divides the real line into intervals of the same length)¹. Nevertheless, scale-dependence is an inherent disadvantage of uniform multi-resolution quantizers. In order to overcome this disadvantage, we consider non-uniform quantizers, where each quantization cell can have different size. The distribution of cell sizes is captured by the following definition.

Definition 3. For a (not necessarily centered) quantizer $Q : \mathbb{R} \rightarrow \mathbb{R}$, define its *cell size cumulative distribution function (cell size cdf)* on the measurable set $S \subseteq \mathbb{R}$ with positive measure as

$$F_{Q,S}(z) := \frac{\lambda(\{x \in S : \lambda(\{y \in S : Q(y) = Q(x)\}) \leq z\})}{\lambda(S)},$$

where λ denotes the Lebesgue measure. Define its *asymptotic cell size cdf* F_Q to be the cdf that is the limit (with respect to the Lévy metric) of $F_{Q,[x_0,x_1]}$ as $x_1 - x_0 \rightarrow \infty$, i.e., F_Q is a cdf and

$$\lim_{l \rightarrow \infty} \sup_{(x_0,x_1):x_1-x_0 \geq l} d_L(F_{Q,[x_0,x_1]}, F_Q) = 0, \quad (1)$$

where $d_L(F, G) := \inf\{\epsilon > 0 : F(x - \epsilon) - \epsilon \leq G(x) \leq F(x + \epsilon) + \epsilon \forall x \in \mathbb{R}\}$ is the Lévy metric. Note that F_Q may not exist for some Q .

We then give the criteria for asymptotic scale invariance.

Definition 4. For a multi-resolution quantizer $\{Q_s\}_{s>0}$, if F_{Q_s} exists for any $s > 0$, define

$$\begin{aligned} \overline{F}_{\{Q_s\}_s}(x) &:= \sup_{s>0} \{F_{Q_s}(xs)\}, \\ \underline{F}_{\{Q_s\}_s}(x) &:= \inf_{s>0} \{F_{Q_s}(xs)\}. \end{aligned}$$

We call $\{Q_s\}_{s>0}$ *asymptotically scale-invariant* if $\overline{F}_{\{Q_s\}_s} = \underline{F}_{\{Q_s\}_s}$, i.e., the functions $x \mapsto F_{Q_s}(xs)$ are the same for all $s > 0$.

One way to improve the BMRQ is to add more intermediate steps between $Q_{2^n}^{\text{bin}}$ and $Q_{2^{n+1}}^{\text{bin}}$, where only some of the adjacent pairs of quantization cells of $Q_{2^n}^{\text{bin}}$ are merged.

Definition 5. Define the *dithered binary multi-resolution quantizer (DBMRQ)* as

$$Q_s^{\text{di}}(x) := \begin{cases} 2^{\lfloor \log_2 s \rfloor + 1} (\lfloor 2^{-\lfloor \log_2 s \rfloor - 1} x \rfloor + 1/2) & \text{if } \text{frac}(\phi \lfloor 2^{-\lfloor \log_2 s \rfloor - 1} x \rfloor) < 2 - 2^{\lfloor \log_2 s \rfloor + 1} / s \\ 2^{\lfloor \log_2 s \rfloor} (\lfloor 2^{-\lfloor \log_2 s \rfloor} x \rfloor + 1/2) & \text{otherwise,} \end{cases}$$

where $\text{frac}(\gamma) := \gamma - \lfloor \gamma \rfloor$, and $\phi := (1 + \sqrt{5})/2$ is the golden ratio (or any irrational number works).

See Figure 2 for an illustration of DBMRQ. The quantizer Q_s^{di} has two cell sizes: $2^{\lfloor \log_2 s \rfloor}$ and $2^{\lfloor \log_2 s \rfloor + 1}$. The choice of which cell size to use is determined by the frac function in the definition. It can be checked that the cell size cdf of Q_s^{di} is

$$\begin{aligned} F_{Q_s^{\text{di}}}(x) &= (2 - 2^{\lfloor \log_2 s \rfloor + 1} / s) \mathbf{1}\{2^{(\lfloor \log_2 s \rfloor + 1)} \leq x\} \\ &\quad + (2^{\lfloor \log_2 s \rfloor + 1} / s - 1) \mathbf{1}\{2^{\lfloor \log_2 s \rfloor} \leq x\}. \end{aligned}$$

Note that the BMRQ and the DBMRQ are not asymptotically scale-invariant.

III. QUANTITIES OF INTEREST

The cell size cdf provides some information about the quantizer Q . We define the following useful quantity.

Definition 6. For a quantizer Q where F_Q exists, define its *Rényi entropy rate* as

$$R_\eta(Q) := \frac{1}{1 - \eta} \log_2 \int_0^\infty \gamma^{\eta-1} dF_Q(\gamma)$$

for $\eta \in \mathbb{R}_{>0} \setminus \{1\}$, and

$$R_1(Q) := \int_0^\infty \log_2(\gamma^{-1}) dF_Q(\gamma).$$

¹We can, for example, use step sizes that are powers of 3, i.e., $Q_s(x) := 3^{\lfloor \log_3 s \rfloor} (\lfloor 3^{-\lfloor \log_3 s \rfloor} x \rfloor + 1/2)$, though it provides less control over the step size, since $\{3^{\lfloor \log_3 s \rfloor}\}_s$ are spaced farther apart than $\{2^{\lfloor \log_2 s \rfloor}\}_s$.

We call $R_0(Q)$ the *log-rate* of Q .

Several quantities of interest can be obtained from the Rényi entropy rate. If $X \sim \text{Unif}[x_0, x_1]$, then:

- The number of reconstruction levels (possible values of $Q(X)$ with positive probability) is

$$|\{x : \mathbf{P}(Q(X) = x) > 0\}| = (x_1 - x_0) \int_0^\infty \gamma^{-1} dF_{Q, [x_0, x_1]}(\gamma).$$

The reason is that for a quantization cell of size γ , the probability that X is in that cell is $\gamma/(x_1 - x_0)$, and hence its contribution to $(x_1 - x_0) \int_0^\infty \gamma^{-1} dF_{Q, [x_0, x_1]}(\gamma)$ is $(x_1 - x_0)\gamma^{-1}(\gamma/(x_1 - x_0)) = 1$. Hence, the number of bits needed to encode the levels (using fixed-length code) is $\lceil \log_2((x_1 - x_0) \int_0^\infty \gamma^{-1} dF_{Q, [x_0, x_1]}(\gamma)) \rceil$. Therefore, the log-rate

$$R_0(Q) = \log_2 \int_0^\infty \gamma^{-1} dF_Q(\gamma)$$

is the logarithm of the rate of increase of the number of reconstruction levels as the interval $[x_0, x_1]$ becomes longer.

- The entropy of the output is

$$H(Q(X)) = \int_0^\infty \log_2(\gamma^{-1}(x_1 - x_0)) dF_{Q, [x_0, x_1]}(\gamma).$$

Therefore,

$$R_1(Q) = \int_0^\infty \log_2(\gamma^{-1}) dF_Q(\gamma)$$

describes how $H(Q(X))$ increases as the interval $[x_0, x_1]$ becomes longer.

- The average L^p error is lower-bounded by

$$\mathbf{E}[|X - Q(X)|^p] \geq \int_0^\infty \frac{(\gamma/2)^p}{p+1} dF_{Q, [x_0, x_1]}(\gamma). \quad (2)$$

The reason is that for a quantization cell of size γ , the expected L^p error conditioned on that X is in that cell is at least $(\gamma/2)^p/(p+1)$ (equality holds if the quantization cell is an interval, and the reconstruction level is its midpoint). The following proposition shows the relation between the asymptotic L^p error and $R_{p+1}(Q)$.

Proposition 7. Fix $p > 0$ and a quantizer Q where F_Q exists. Let $X \sim \text{Unif}[x_0, x_1]$. We have

$$\liminf_{x_1 - x_0 \rightarrow \infty} \mathbf{E}[|X - Q(X)|^p] \geq \frac{1}{p+1} 2^{-p(R_{p+1}(Q)+1)}.$$

Moreover, if Q is centered, then

$$\lim_{x_1 - x_0 \rightarrow \infty} \mathbf{E}[|X - Q(X)|^p] = \frac{1}{p+1} 2^{-p(R_{p+1}(Q)+1)}.$$

Proof: For the first part, if $x_1 - x_0 \rightarrow \infty$, then $F_{Q, [x_0, x_1]} \rightarrow F_Q$ (in Lévy metric), and hence by (2),

$$\begin{aligned} & \liminf_{x_1 - x_0 \rightarrow \infty} \mathbf{E}[|X - Q(X)|^p] \\ & \geq \liminf_{x_1 - x_0 \rightarrow \infty} \int_0^\infty \frac{(\gamma/2)^p}{p+1} dF_{Q, [x_0, x_1]}(\gamma) \\ & \geq \int_0^\infty \frac{(\gamma/2)^p}{p+1} dF_Q(\gamma) \\ & = \frac{1}{p+1} 2^{-p(R_{p+1}(Q)+1)}. \end{aligned}$$

For the second part, assume that Q is centered. We first show that there exists $b > 0$ such that each quantization cell has size upper-bounded by b . Assume the contrary that the quantization cells can be arbitrarily large. Then there exists a sequence $\{(x_{0,i}, x_{1,i})\}_i$ such that $x_{1,i} - x_{0,i} \rightarrow \infty$ and $F_{Q, [x_{0,i}, x_{1,i}]}(\gamma) = \mathbf{1}\{\gamma \geq x_{1,i} - x_{0,i}\}$ (take $x_{0,i}, x_{1,i}$ to be the end points of cells in a sequence of cells with sizes tend to ∞). This contradicts (1) since $F_{Q, [x_{0,i}, x_{1,i}]}$ does not have a limit. Hence such $b > 0$ exists. For $X \sim \text{Unif}[x_0, x_1]$, we have

$$\begin{aligned} & \mathbf{E}[|X - Q(X)|^p] \\ & \leq \mathbf{E}[\mathbf{1}\{X \in [x_0 + b, x_1 - b]\}|X - Q(X)|^p] + \mathbf{P}(X \notin [x_0 + b, x_1 - b]) \frac{(b/2)^p}{p+1} \\ & \leq \int_0^\infty \frac{(\gamma/2)^p}{p+1} dF_{Q, [x_0, x_1]}(\gamma) + \frac{2b}{x_1 - x_0} \cdot \frac{(b/2)^p}{p+1} \\ & \rightarrow \int_0^\infty \frac{(\gamma/2)^p}{p+1} dF_Q(\gamma) = \frac{1}{p+1} 2^{-p(R_{p+1}(Q)+1)} \end{aligned}$$

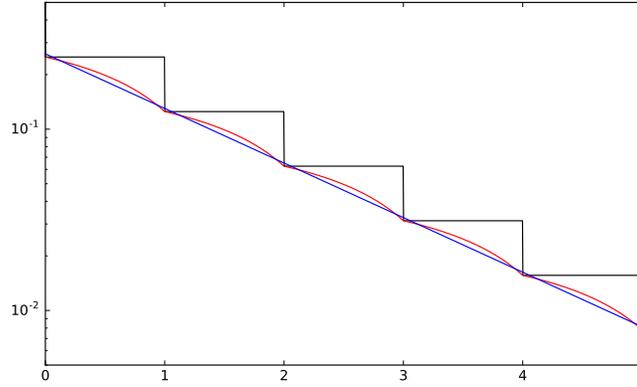


Figure 1. Plot of the asymptotic L^1 error against the log-rate (more precisely, plot of $\inf\{2^{-R_2(Q_s)^{-2}} : s > 0, R_0(Q_s) \leq x\}$ against x) for the BMRQ $\{Q_s^{\text{bin}}\}$ (black), the DBMRQ $\{Q_s^{\text{di}}\}$ (red), and the BBMRQ $\{Q_s^{\text{bias},\alpha}\}$ for α close to $1/2$ (blue). We can see that the BBMRQ provides a smoother trade-off between the error and the rate (the log plot is a straight line due to scale invariance). Also, the BBMRQ outperforms the BMRQ and DBMRQ except when the log rate is close to an integer.

as $x_1 - x_0 \rightarrow \infty$. ■

IV. BIASED BINARY MULTI-RESOLUTION QUANTIZER

We now state the main result in this paper, which is proved later in this section.

Theorem 8. *For any $\epsilon > 0$, there exists an asymptotically scale-invariant centered MRQ $\{Q_s\}_{s>0}$ with $\bar{F}_{\{Q_s\}_s}(1/2 - \epsilon) = 0$, $\bar{F}_{\{Q_s\}_s}(1) = 1$, and $d_L(\bar{F}_{\{Q_s\}_s}, F_{2^{\text{Unif}[-1,0]}}) < \epsilon$, where*

$$F_{2^{\text{Unif}[-1,0]}}(x) := \min\{\max\{\log_2(x) + 1, 0\}, 1\},$$

i.e., $F_{2^{\text{Unif}[-1,0]}}$ is the cdf of 2^Z where $Z \sim \text{Unif}[-1, 0]$.

As a result, the Rényi entropy rate $R_\eta(Q_s)$ can be arbitrarily close to

$$\frac{1}{1-\eta} \log_2 \left(\frac{1-2^{1-\eta}}{\eta-1} \log_2 e \right) - \log_2 s \quad (3)$$

for $\eta \neq 1$, and $R_1(Q_s)$ can be arbitrarily close to $1/2 - \log_2 s$. We will show in Corollary 11 that this tradeoff between log-rate and L^p error can be arbitrarily close to optimal.

To prove Theorem 8, we introduce the following construction.

Definition 9. We define the *biased quantization tree* with parameter $0 < \alpha < 1$ recursively as follows. For any $n \in \mathbb{Z}$ and sequence $\{z_i\}_{i \leq n}$ ($z_i \in \{0, 1\}$, the index i is over $\mathbb{Z} \cap (-\infty, n]$) with finitely many 1's, define

$$\begin{aligned} & (\underline{T}_n^{\text{bias},\alpha}(\{z_i\}_{i \leq n}), \bar{T}_n^{\text{bias},\alpha}(\{z_i\}_{i \leq n})) \\ & := \begin{cases} (0, \alpha^n) & \text{if } z_n = z_{n-1} = \dots = 0 \\ (\underline{T}_{n-1}, \alpha \bar{T}_{n-1} + (1-\alpha)\underline{T}_{n-1}) & \text{else if } z_n = 0 \\ (\alpha \bar{T}_{n-1} + (1-\alpha)\underline{T}_{n-1}, \bar{T}_{n-1}) & \text{otherwise.} \end{cases} \end{aligned}$$

where we write $\bar{T}_{n-1} = \bar{T}_{n-1}^{\text{bias},\alpha}(\{z_i\}_{i \leq n-1})$ (likewise for \underline{T}_{n-1}) for brevity. Note that the first case above serves as the base case of the recursive definition.

Define the *biased binary multi-resolution quantizer (BBMRQ)* with parameter $0 < \alpha < 1$ as follows. For $x \geq 0$, define

$$Q_s^{\text{bias},\alpha}(x) := \frac{1}{2} \left(\underline{T}_n^{\text{bias},\alpha}(\{z_i\}_{i \leq n}) + \bar{T}_n^{\text{bias},\alpha}(\{z_i\}_{i \leq n}) \right),$$

where $\{z_i\}_{i \leq n}$ satisfies $x \in [\underline{T}_n^{\text{bias},\alpha}(\{z_i\}_{i \leq n}), \bar{T}_n^{\text{bias},\alpha}(\{z_i\}_{i \leq n})]$ and $\bar{T}_n^{\text{bias},\alpha}(\{z_i\}_{i \leq n}) - \underline{T}_n^{\text{bias},\alpha}(\{z_i\}_{i \leq n}) \leq s$, and we select the $\{z_i\}_{i \leq n}$ with the smallest n satisfying these two constraints. For $x < 0$, define $Q_s^{\text{bias},\alpha}(x) := -Q_s^{\text{bias},\alpha}(-x)$.

Intuitively, the BBMRQ repeatedly divides an interval into two subintervals of proportion α and $1 - \alpha$, until the lengths of the intervals fall below s (see Figure 2). It is clear that the BBMRQ is a centered MRQ.

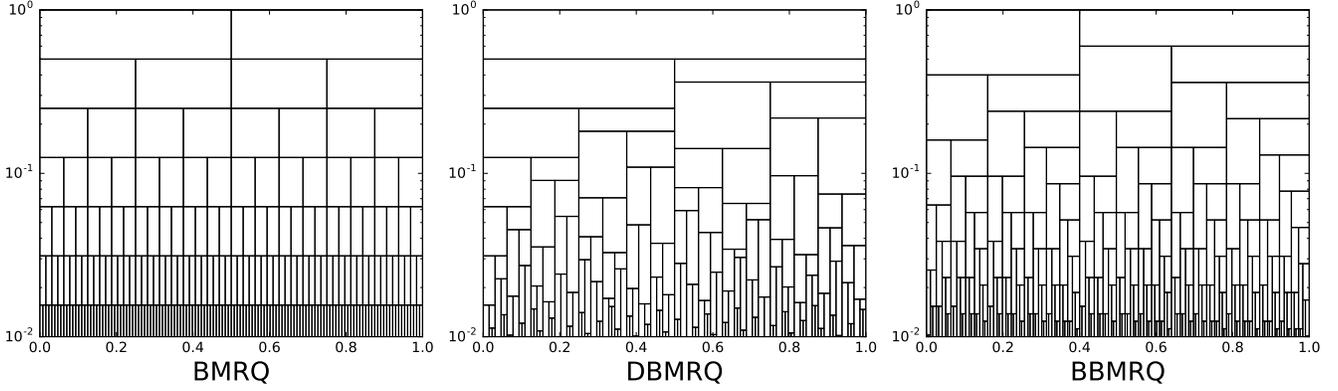


Figure 2. The quantization tree of BMRQ, DBMRQ and BMMRQ (from left to right). The x-axis is the input x , and the y-axis is the parameter s . A rectangle in the figure means that each point (x, s) in the rectangle has the same quantized value $Q_s(x)$.

We then use the BMMRQ to prove Theorem 8.

Proof: Fix any $1/2 < \alpha < 3/4$ such that $(\ln \alpha)/(\ln(1 - \alpha))$ is irrational. Write $Q_s = Q_s^{\text{bias}, \alpha}$ for brevity. Let $X \sim \text{Unif}[0, 1]$, and

$$W_t := 2^t \lambda(\{y : Q_{2^{-t}}(y) = Q_{2^{-t}}(X)\})$$

for $t \geq 0$, where λ denotes the Lebesgue measure. Note that $-\log_2 W_t$ is the residual life of a renewal process with interarrival time distribution $\alpha \delta_{-\log_2 \alpha} + (1 - \alpha) \delta_{-\log_2(1 - \alpha)}$, where δ_w is the degenerate distribution at w (as t increases, each time the quantization cell containing X splits into two, there is a probability α for X to be in the cell with proportion α , and a probability $1 - \alpha$ for X to be in the cell with proportion $1 - \alpha$). Since $(\ln \alpha)/(\ln(1 - \alpha))$ is irrational, the interarrival times have a non-lattice distribution. Denote the Markov kernel (conditional distribution of W_{s+t} given W_s) as κ_α^t . By the key renewal theorem [17], we have

$$d_L(\kappa_\alpha^t \delta_1, F_{\text{bias}, \alpha}) \rightarrow 0$$

as $t \rightarrow \infty$, where $\kappa_\alpha^t \delta_1$ denotes the distribution of W_t conditioned on $W_0 = 1$, and $F_{\text{bias}, \alpha}$ is the cdf of the stationary distribution of the process $\{W_t\}$, given by

$$\begin{aligned} F_{\text{bias}, \alpha}(\gamma) &:= (-\alpha \log_2 \alpha - (1 - \alpha) \log_2(1 - \alpha))^{-1} \\ &\quad \cdot \left(\alpha \max\{\log_2(\min\{\gamma, 1\}/\alpha), 0\} \right. \\ &\quad \left. + (1 - \alpha) \max\{\log_2(\min\{\gamma, 1\}/(1 - \alpha)), 0\} \right). \end{aligned} \quad (4)$$

For $\tau > 0$, let ϵ_τ be such that $d_L(\kappa_\alpha^t \delta_1, F_{\text{bias}, \alpha}) \leq \epsilon_\tau$ for all $t \geq \tau$, and $\epsilon_\tau \rightarrow 0$ as $\tau \rightarrow \infty$.

Fix any $s > 0$, $\tau > 1$ and $l > \tau 2^{\tau+3} s$. Fix any x_0, x_1 such that $x_1 - x_0 \geq l$. Let $X \sim \text{Unif}[x_0, x_1]$. Let $[\underline{T}_{n^{(i)}}(z^{(i)}), \bar{T}_{n^{(i)}}(z^{(i)})]$ be the quantization cells of $Q_{2^{\tau+2} s}$ that are subsets of $[x_0, x_1]$, where $n^{(i)} \in \mathbb{Z}$ and $z^{(i)} \in \{0, 1\}^{\mathbb{Z} \cap (-\infty, n^{(i)})}$ for $i = 1, 2, \dots, \tilde{N}$. Since X is not in one of $[\underline{T}_{n^{(i)}}(z^{(i)}), \bar{T}_{n^{(i)}}(z^{(i)})]$ only when X is in a quantization cell of $Q_{2^{\tau+2} s}$ that includes x_0 or x_1 , and each quantization cell of $Q_{2^{\tau+2} s}$ has length between $2^\tau s$ and $2^{\tau+2} s$ (since $1/2 < \alpha < 3/4$), we have

$$\begin{aligned} \mathbf{P}\left(X \notin \bigcup_{i=1}^{\tilde{N}} [\underline{T}_{n^{(i)}}(z^{(i)}), \bar{T}_{n^{(i)}}(z^{(i)})]\right) &\leq \frac{2(2^{\tau+2} s)}{x_1 - x_0} \\ &\leq 1/\tau. \end{aligned}$$

Let $W_t := 2^t \lambda(\{y : Q_{2^{-t}}(y) = Q_{2^{-t}}(X)\})$. Conditioned on $X \in [\underline{T}_{n^{(i)}}(z^{(i)}), \bar{T}_{n^{(i)}}(z^{(i)})]$ (let $L := \bar{T}_{n^{(i)}}(z^{(i)}) - \underline{T}_{n^{(i)}}(z^{(i)})$), we have $W_{-\log_2 L} = 1$, and $\{W_t\}_{t \geq -\log_2 L}$ is a stochastic process with Markov kernel κ_α^t , and thus $W_{-\log_2 s} \sim \kappa_\alpha^{\log_2 L - \log_2 s} \delta_1$, and $d_L(\kappa_\alpha^{\log_2 L - \log_2 s} \delta_1, F_{\text{bias}, \alpha}) \leq \epsilon_\tau$ since $\log_2 L - \log_2 s \geq \log_2 2^\tau s - \log_2 s = \tau$. Let $\bar{W} := s^{-1} \lambda(\{y \in [x_0, x_1] : Q_s(y) = Q_s(x)\})$. Since $\bar{W} = W_{-\log_2 s}$ if $X \in \bigcup_{i=1}^{\tilde{N}} [\underline{T}_{n^{(i)}}(z^{(i)}), \bar{T}_{n^{(i)}}(z^{(i)})]$, we have $d_L(F_{\bar{W}}, F_{\text{bias}, \alpha}) < \epsilon_\tau + 1/\tau$. Therefore $F_{\bar{W}} \rightarrow F_{\text{bias}, \alpha}$ as $\tau \rightarrow \infty$, and thus $\{Q_s\}$ is asymptotically scale-invariant with cell size cdf $F_{\text{bias}, \alpha}$. The result follows from the fact that $d_L(F_{\text{bias}, \alpha}, F_{2 \text{Unif}[-1, 0]}) \rightarrow 0$ as $\alpha \rightarrow 1/2$. ■

V. CONVERSE RESULTS

In this section, we show a fundamental tradeoff between the log-rate and the L^p error of asymptotically scale-invariant multi-resolution quantizers, and that BBMRQ can be arbitrarily close to optimal in this regard. We first show an inequality that must be satisfied by all asymptotically scale-invariant multi-resolution quantizers.

Theorem 10. *For any asymptotically scale-invariant multi-resolution quantizer $\{Q_s\}_{s>0}$ with finite log-rate (i.e., $R_0(Q_1) = \log_2 \int_0^\infty \gamma^{-1} d\bar{F}(\gamma) < \infty$, where we write $\bar{F} = \bar{F}_{\{Q_s\}_s}$), the distribution given by the cdf \bar{F} is a continuous distribution, and its pdf \bar{f} satisfies*

$$\int_0^\infty x^{-1} (1 + \mathbf{1}\{\bar{f}(x) \geq \zeta \bar{f}(x\zeta)\}) (\bar{f}(x) - \zeta \bar{f}(x\zeta)) dx \leq 0 \quad (5)$$

for any $\zeta > 1$. As a result,

$$\bar{f}(y) \leq \int_0^\infty x^{-1} (1 + \mathbf{1}\{x > y\}) \bar{f}(x) dx \quad (6)$$

for almost all $y > 0$.

Proof: Let $p_{s,\gamma}$ be the probability mass function corresponding to the cdf $F_{Q_s,[-\gamma,\gamma]}$ for $\gamma > 0$ (note that it is a discrete distribution so pmf exists). The number of quantization cells of Q_s in $[-\gamma, \gamma]$ of size x is $2\gamma x^{-1} p_{s,\gamma}(x)$. Let n_x (resp. \tilde{n}_x) be the number of quantization cells of Q_s (resp. $Q_{s/\gamma}$) in $[-\gamma, \gamma]$ of size x that are not quantization cells of $Q_{s/\gamma}$ (resp. Q_s). We have $n_x - \tilde{n}_x = 2\gamma x^{-1} (p_{s,\gamma}(x) - p_{s/\zeta,\gamma}(x))$. Since each cell of Q_s is split into 2 or more cells in $Q_{s/\gamma}$, we have $2 \sum_x n_x \leq \sum_x \tilde{n}_x$ (the summation is over x in the support of $p_{s,\gamma}$ or $p_{s/\zeta,\gamma}$), and hence

$$\begin{aligned} & 2 \sum_{x: p_{s,\gamma}(x) > p_{s/\zeta,\gamma}(x)} x^{-1} (p_{s,\gamma}(x) - p_{s/\zeta,\gamma}(x)) \\ &= 2\gamma^{-1} \sum_{x: n_x > \tilde{n}_x} (n_x - \tilde{n}_x) \\ &= \gamma^{-1} \left(2 \sum_x n_x - 2 \sum_x \min\{n_x, \tilde{n}_x\} \right) \\ &\leq \gamma^{-1} \left(\sum_x \tilde{n}_x - 2 \sum_x \min\{n_x, \tilde{n}_x\} \right) \\ &\leq \gamma^{-1} \sum_{x: n_x < \tilde{n}_x} (\tilde{n}_x - n_x) \\ &= \sum_{x: p_{s,\gamma}(x) < p_{s/\zeta,\gamma}(x)} x^{-1} (p_{s/\zeta,\gamma}(x) - p_{s,\gamma}(x)). \end{aligned} \quad (7)$$

Let ν be the signed measure induced by $\bar{F}(x) - \bar{F}(x\zeta)$ (i.e., $\nu(B) = \int_B d(\bar{F}(x) - \bar{F}(x\zeta))$). Fix any $\epsilon > 0$. Let $A \subseteq \mathbb{R}_{>0}$ be a measurable set. Since the measure $B \mapsto \int_B x^{-1} d(\bar{F}(x) + \bar{F}(x\zeta))$ is a regular measure over $\mathbb{R}_{>0}$ ², there exists an open set A' such that $A \subseteq A'$ and $\int_{A' \setminus A} x^{-1} d(\bar{F}(x) + \bar{F}(x\zeta)) < \epsilon$. Since A' is open, it can be expressed as a countable disjoint union of open intervals $A' = \bigcup_i (a_i, b_i)$. Let $A'_\delta := \bigcup_i (a_i + \delta, b_i - \delta)$ for $\delta > 0$ (where $(a_i + \delta, b_i - \delta) = \emptyset$ if $a_i + \delta \geq b_i - \delta$). Let $g_\delta(x) := \min\{(x - a_i)\delta^{-1}, (b_i - x)\delta^{-1}, 1\}$ if $x \in (a_i, b_i)$ ($g_\delta(x) = 0$ if x is not in any of those intervals). Then g_δ is continuous and $\mathbf{1}\{x \in A'_\delta\} \leq g_\delta(x) \leq \mathbf{1}\{x \in A'\}$. Hence,

$$\begin{aligned} & \int_A x^{-1} \nu(dx) \\ &\leq \int_{A'} x^{-1} \nu(dx) + \int_{A' \setminus A} x^{-1} d(\bar{F}(x) + \bar{F}(x\zeta)) \\ &\leq \int x^{-1} g_\delta(x) \nu(dx) + \int_{A' \setminus A'_\delta} x^{-1} d(\bar{F}(x) + \bar{F}(x\zeta)) + \epsilon \\ &\rightarrow \int x^{-1} g_\delta(x) \nu(dx) + \epsilon \end{aligned}$$

²We can reparameterize the space by $y = \ln x$. Then $\int_B x^{-1} d(\bar{F}(x) + \bar{F}(x\zeta)) = \int_{\ln(B)} e^{-y} d(\bar{F}(e^y) + \bar{F}(e^y\zeta))$. We have $\int_{B'} e^{-y} d(\bar{F}(e^y) + \bar{F}(e^y\zeta)) \leq 2e^{-\inf B'}$, and hence the measure $B' \mapsto \int_{B'} e^{-y} d(\bar{F}(e^y) + \bar{F}(e^y\zeta))$ is locally finite, and hence regular.

as $\delta \rightarrow 0$ since $A'_\delta \nearrow A'$. Also,

$$\begin{aligned} & \int_A x^{-1} \nu(dx) \\ & \geq \int_{A'} x^{-1} \nu(dx) \\ & \geq \int x^{-1} g_\delta(x) \nu(dx) - \int_{A' \setminus A'_\delta} x^{-1} d(\bar{F}(x) + \bar{F}(x\zeta)) \\ & \rightarrow \int x^{-1} g_\delta(x) \nu(dx) \end{aligned}$$

as $\delta \rightarrow 0$. Therefore, we can let $g := g_\delta$ for a $\delta > 0$ small enough that

$$\left| \int_A x^{-1} \nu(dx) - \int x^{-1} g(x) \nu(dx) \right| < 2\epsilon. \quad (8)$$

Since $F_{Q_{s,[-\gamma,\gamma]}} \rightarrow \bar{F}$ as $\gamma \rightarrow \infty$, and $x \mapsto 1 + g(x)$ is bounded and continuous, we have $\int x^{-1}(1 + g(x)) dF_{Q_{1,[-\gamma,\gamma]}}(x) \rightarrow \int x^{-1}(1 + g(x)) d\bar{F}(x)$. Also, $\int x^{-1}(1 + g(x)) dF_{Q_{1/\zeta,[-\gamma,\gamma]}}(x) \rightarrow \int x^{-1}(1 + g(x)) d\bar{F}(x\zeta)$. As a result,

$$\begin{aligned} & \int x^{-1}(1 + g(x)) d(F_{Q_{1,[-\gamma,\gamma]}}(x) - F_{Q_{1/\zeta,[-\gamma,\gamma]}}(x)) \\ & \rightarrow \int x^{-1}(1 + g(x)) d(\bar{F}(x) - \bar{F}(x\zeta)) \\ & = \int x^{-1}(1 + g(x)) \nu(dx). \end{aligned} \quad (9)$$

It can be deduced from (7) that

$$\sum_x x^{-1}(1 + \tilde{g}(x))(p_{s,\gamma}(x) - p_{s/\zeta,\gamma}(x)) \leq 0$$

for any $\tilde{g} : \mathbb{R}_{>0} \rightarrow [0, 1]$, where the summation is over x in the support of $p_{s,\gamma}$ or $p_{s/\zeta,\gamma}$ (because to maximize the above expression, we should assign $\tilde{g}(x) = 0$ to x 's where $p_{s/\zeta,\gamma}(x) - p_{s,\gamma}(x) > 0$, and $\tilde{g}(x) = 1$ otherwise). Substituting $\tilde{g} = g$,

$$\int x^{-1}(1 + g(x)) d(F_{Q_{1,[-\gamma,\gamma]}}(x) - F_{Q_{1/\zeta,[-\gamma,\gamma]}}(x)) \leq 0.$$

By (9),

$$\int x^{-1}(1 + g(x)) \nu(dx) \leq 0.$$

By (8),

$$\int x^{-1} \nu(dx) + \int_A x^{-1} \nu(dx) \leq 2\epsilon.$$

Taking $\epsilon \rightarrow 0$ and rearranging the terms,

$$\begin{aligned} & \int x^{-1}(1 + \mathbf{1}\{x \in A\}) \nu(dx) \leq 0, \quad (10) \\ & \int_0^\infty x^{-1}(1 + \mathbf{1}\{x \in A\}) d\bar{F}(x) \leq \zeta \int_0^\infty x^{-1}(1 + \mathbf{1}\{x/\zeta \in A\}) d\bar{F}(x), \\ & \int_0^\infty x^{-1}(1 + \mathbf{1}\{x \in A\} - \zeta - \zeta \mathbf{1}\{x/\zeta \in A\}) d\bar{F}(x) \leq 0. \end{aligned}$$

Substituting $A = \{x : x > y\}$,

$$\begin{aligned} 0 & \geq \int_0^\infty x^{-1}(1 + \mathbf{1}\{x > y\} - \zeta - \zeta \mathbf{1}\{x > y\zeta\}) d\bar{F}(x) \\ & = \int_0^\infty x^{-1}(\mathbf{1}\{y < x \leq y\zeta\} - (\zeta - 1)(1 + \mathbf{1}\{x > y\zeta\})) d\bar{F}(x) \\ & \geq \frac{\bar{F}(y\zeta) - \bar{F}(y)}{y\zeta} - (\zeta - 1) \int_0^\infty x^{-1}(1 + \mathbf{1}\{x > y\zeta\}) d\bar{F}(x). \end{aligned}$$

Hence,

$$\frac{\bar{F}(y\zeta) - \bar{F}(y)}{y\zeta - y} \leq \zeta \int_0^\infty x^{-1}(1 + \mathbf{1}\{x > y\zeta\}) d\bar{F}(x).$$

Since the right hand side is bounded by $\zeta 2^{R_0+1}$, \bar{f} exists and (6) holds for almost all $y > 0$. Therefore, (10) becomes

$$\int x^{-1}(1 + \mathbf{1}\{x \in A\}) (\bar{f}(x) - \zeta \bar{f}(x\zeta)) dx \leq 0.$$

We can obtain (5) by substituting $A = \{x : \bar{f}(x) \geq \zeta \bar{f}(x\zeta)\}$. ■

As a result, we can bound the log-rate and the L^p error of an asymptotically scale-invariant multi-resolution quantizer using the following corollary and Proposition 7.

Corollary 11. *For any asymptotically scale-invariant multi-resolution quantizer $\{Q_s\}_{s>0}$ with finite log-rate, for any $s > 0$, $p > 0$,*

$$R_0(Q_s) - R_{p+1}(Q_s) \geq \frac{1}{p} \log_2 \left(\frac{1 - 2^{-p}}{p} (\log_2 e)^{p+1} \right). \quad (11)$$

Recall that for the BBMRQ, by (3), $R_0(Q_s)$ can be arbitrarily close to $\log_2 \log_2 e - \log_2 s$, and $R_{p+1}(Q_s)$ can be arbitrarily close to

$$-\frac{1}{p} \log_2 \left(\frac{1 - 2^{-p}}{p} \log_2 e \right) - \log_2 s.$$

Therefore the lower bound in (11) can be approached. This shows that the tradeoff between log-rate and L^p error of BBMRQ can be arbitrarily close to optimal.

Nevertheless, it is unknown whether the lower bound in (11) can be attained exactly. It can be tracked in the proof of Corollary 11 that the equality in (11) holds if and only if there exists $s > 0$ such that $\bar{F}_{\{Q_s\}_s}(\gamma) = F_{2^{\text{Unif}[-1,0]}}(\gamma s)$ for all γ , i.e., Theorem 8 can be attained exactly. We conjecture that this is impossible.

Conjecture 12. *There does not exist an asymptotically scale-invariant centered MRQ $\{Q_s\}_{s>0}$ with $\bar{F}_{\{Q_s\}_s} = F_{2^{\text{Unif}[-1,0]}}$.*

We now prove Corollary 11.

Proof: By Theorem 10, for any pdf $\psi : \mathbb{R}_{>0} \rightarrow \mathbb{R}_{\geq 0}$,

$$\begin{aligned} \int \bar{f}(y)\psi(y)dy &\leq \int_0^\infty \int_0^\infty x^{-1}(1 + \mathbf{1}\{x > y\}) \bar{f}(x)dx \cdot \psi(y)dy \\ &= \int_0^\infty x^{-1} \left(1 + \int_0^x \psi(y)dy \right) \bar{f}(x)dx, \\ \int_0^\infty \left(x^{-1} \left(1 + \int_0^x \psi(y)dy \right) - \psi(x) \right) \bar{f}(x)dx &\geq 0. \end{aligned} \quad (12)$$

Substitute

$$\psi(y) = \mathbf{1}\{1/2 \leq y \leq 1\} \frac{(1 - 2^{-p}) \log_2 y - y^p + 1}{(1 - 2^{-p}) \left(\frac{1 - \log_2 e}{2} \right) - \frac{1 - 2^{-(p+1)}}{p+1} + 1/2}.$$

For $1/2 \leq x \leq 1$,

$$\begin{aligned} &x^{-1} \left(1 + \int_0^x \psi(y)dy \right) - \psi(x) \\ &= x^{-1} + \frac{x^{-1} \int_{1/2}^x ((1 - 2^{-p}) \log_2 y - y^p + 1) dy - ((1 - 2^{-p}) \log_2 x - x^p + 1)}{(1 - 2^{-p}) \left(\frac{1 - \log_2 e}{2} \right) - \frac{1 - 2^{-(p+1)}}{p+1} + 1/2} \\ &= x^{-1} + \frac{x^{-1} \left((1 - 2^{-p}) (-x \log_2 e + x \log_2 x + (\log_2 e)/2 + 1/2) - \frac{x^{p+1} - 2^{-(p+1)}}{p+1} + x - 1/2 \right) - ((1 - 2^{-p}) \log_2 x - x^p + 1)}{(1 - 2^{-p}) \left(\frac{1 - \log_2 e}{2} \right) - \frac{1 - 2^{-(p+1)}}{p+1} + 1/2} \\ &= x^{-1} + \frac{(1 - 2^{-p}) (-\log_2 e + \log_2 x + (\log_2 e)x^{-1}/2 + x^{-1}/2) - \frac{x^p - 2^{-(p+1)}x^{-1}}{p+1} + 1 - x^{-1}/2 - (1 - 2^{-p}) \log_2 x + x^p - 1}{(1 - 2^{-p}) \left(\frac{1 - \log_2 e}{2} \right) - \frac{1 - 2^{-(p+1)}}{p+1} + 1/2} \\ &= x^{-1} + \frac{(1 - 2^{-p}) (-\log_2 e + (\log_2 e)x^{-1}/2 + x^{-1}/2) - \frac{x^p - 2^{-(p+1)}x^{-1}}{p+1} - x^{-1}/2 + x^p}{(1 - 2^{-p}) \left(\frac{1 - \log_2 e}{2} \right) - \frac{1 - 2^{-(p+1)}}{p+1} + 1/2} \\ &= \left(1 + \frac{(1 - 2^{-p}) ((\log_2 e)/2 + 1/2) + \frac{2^{-(p+1)}}{p+1} - 1/2}{(1 - 2^{-p}) \left(\frac{1 - \log_2 e}{2} \right) - \frac{1 - 2^{-(p+1)}}{p+1} + 1/2} \right) x^{-1} + \frac{(1 - 2^{-p}) (-\log_2 e) - \frac{x^p}{p+1} + x^p}{(1 - 2^{-p}) \left(\frac{1 - \log_2 e}{2} \right) - \frac{1 - 2^{-(p+1)}}{p+1} + 1/2} \end{aligned}$$

$$\begin{aligned}
&= \left(\frac{\frac{p}{p+1}(1-2^{-p})}{(1-2^{-p})\left(\frac{1-\log_2 e}{2}\right) - \frac{1-2^{-(p+1)}}{p+1} + 1/2} \right) x^{-1} + \frac{(1-2^{-p})(-\log_2 e) + x^p \frac{p}{p+1}}{(1-2^{-p})\left(\frac{1-\log_2 e}{2}\right) - \frac{1-2^{-(p+1)}}{p+1} + 1/2} \\
&= \frac{\frac{p}{p+1}((1-2^{-p})x^{-1} + x^p) - (1-2^{-p})\log_2 e}{(1-2^{-p})\left(\frac{1-\log_2 e}{2}\right) - \frac{1-2^{-(p+1)}}{p+1} + 1/2} \\
&= \frac{\frac{p}{p+1}((1-2^{-p})x^{-1} + x^p) - (1-2^{-p})\log_2 e}{(1-2^{-p})\left(\frac{1-\log_2 e}{2}\right) + \frac{p/2-(1/2)(1-2^{-p})}{p+1}} \\
&= 2 \cdot \frac{\frac{p}{p+1}((1-2^{-p})x^{-1} + x^p) - (1-2^{-p})\log_2 e}{\frac{p}{p+1}(2-2^{-p}) - (1-2^{-p})\log_2 e}
\end{aligned}$$

For $x \geq 1$, we have

$$\begin{aligned}
p &\geq (1-2^{-p})\log_2 e, \\
\frac{p+1}{p} &\geq \frac{p+1}{p^2}(1-2^{-p})\log_2 e, \\
x^{p+1} &\geq 1 \geq \frac{p+1}{p^2} \left((1-2^{-p})\log_2 e - \frac{p}{p+1} \right), \\
\frac{p^2}{p+1} x^{p-1} &\geq \left((1-2^{-p})\log_2 e - \frac{p}{p+1} \right) x^{-2}.
\end{aligned}$$

Integrating both sides from 1 to x ,

$$\begin{aligned}
\frac{p}{p+1}(x^p - 1) &\geq \left((1-2^{-p})\log_2 e - \frac{p}{p+1} \right) (1 - x^{-1}), \\
\left((1-2^{-p})\log_2 e - \frac{p}{p+1} \right) x^{-1} + \frac{p}{p+1} x^p - (1-2^{-p})\log_2 e &\geq 0.
\end{aligned}$$

Therefore,

$$\begin{aligned}
&2 \cdot \frac{\frac{p}{p+1}((1-2^{-p})x^{-1} + x^p) - (1-2^{-p})\log_2 e}{\frac{p}{p+1}(2-2^{-p}) - (1-2^{-p})\log_2 e} - \left(x^{-1} \left(1 + \int_0^x \psi(y) dy \right) - \psi(x) \right) \\
&= 2 \cdot \frac{\frac{p}{p+1}((1-2^{-p})x^{-1} + x^p) - (1-2^{-p})\log_2 e}{\frac{p}{p+1}(2-2^{-p}) - (1-2^{-p})\log_2 e} - 2x^{-1} \\
&= 2 \cdot \frac{\left((1-2^{-p})\log_2 e - \frac{p}{p+1} \right) x^{-1} + \frac{p}{p+1} x^p - (1-2^{-p})\log_2 e}{\frac{p}{p+1}(2-2^{-p}) - (1-2^{-p})\log_2 e} \\
&\geq 0.
\end{aligned}$$

For $0 < x \leq 1/2$, by the concavity of $y \mapsto y \log_2(1/y)$,

$$\begin{aligned}
p2^{-p} &= 2^{-p} \log_2 2^p \leq (1-2^{-p})\log_2 e, \\
\frac{p+1}{p} 2^{-(p+1)} &\leq \frac{p+1}{2p^2} (1-2^{-p})\log_2 e, \\
x^{p+1} &\leq \frac{p+1}{2p^2} (1-2^{-p})\log_2 e - \frac{1}{2p} 2^{-p}, \\
x^{p+1} &\leq \frac{p+1}{2p^2} \left((1-2^{-p})\log_2 e - \frac{p}{p+1} 2^{-p} \right), \\
\frac{2p^2}{p+1} x^{p-1} &\leq \left((1-2^{-p})\log_2 e - \frac{p}{p+1} 2^{-p} \right) x^{-2},
\end{aligned}$$

Integrating both sides from x to $1/2$,

$$\frac{2p}{p+1}(2^{-p} - x^p) \leq \left((1-2^{-p})\log_2 e - \frac{p}{p+1} 2^{-p} \right) (x^{-1} - 2),$$

$$\left((1 - 2^{-p}) \log_2 e - \frac{p}{p+1} 2^{-p} \right) x^{-1} + \frac{2p}{p+1} x^p - 2(1 - 2^{-p}) \log_2 e \geq 0.$$

Therefore,

$$\begin{aligned} & 2 \cdot \frac{\frac{p}{p+1} \left((1 - 2^{-p}) x^{-1} + x^p \right) - (1 - 2^{-p}) \log_2 e}{\frac{p}{p+1} (2 - 2^{-p}) - (1 - 2^{-p}) \log_2 e} - \left(x^{-1} \left(1 + \int_0^x \psi(y) dy \right) - \psi(x) \right) \\ &= 2 \cdot \frac{\frac{p}{p+1} \left((1 - 2^{-p}) x^{-1} + x^p \right) - (1 - 2^{-p}) \log_2 e}{\frac{p}{p+1} (2 - 2^{-p}) - (1 - 2^{-p}) \log_2 e} - x^{-1} \\ &= \frac{\left((1 - 2^{-p}) \log_2 e - \frac{p}{p+1} 2^{-p} \right) x^{-1} + \frac{2p}{p+1} x^p - 2(1 - 2^{-p}) \log_2 e}{\frac{p}{p+1} (2 - 2^{-p}) - (1 - 2^{-p}) \log_2 e} \\ &\geq 0. \end{aligned}$$

Hence, for any $x > 0$,

$$x^{-1} \left(1 + \int_0^x \psi(y) dy \right) - \psi(x) \leq \frac{\frac{p}{p+1} \left((1 - 2^{-p}) x^{-1} + x^p \right) - (1 - 2^{-p}) \log_2 e}{\frac{p}{p+1} (2 - 2^{-p}) - (1 - 2^{-p}) \log_2 e}.$$

By (12),

$$\begin{aligned} & \frac{p(1 - 2^{-p})}{p+1} \int_0^\infty x^{-1} \bar{f}(x) dx + \frac{p}{p+1} \int_0^\infty x^p \bar{f}(x) dx \geq (1 - 2^{-p}) \log_2 e, \\ & (1 - 2^{-p}) \int_0^\infty x^{-1} \bar{f}(x) dx + \int_0^\infty x^p \bar{f}(x) dx \geq \frac{p+1}{p} (1 - 2^{-p}) \log_2 e. \end{aligned}$$

Therefore,

$$(1 - 2^{-p}) 2^{R_0(Q_1)} + 2^{-p} 2^{R_{p+1}(Q_1)} \geq \frac{p+1}{p} (1 - 2^{-p}) \log_2 e.$$

Note that the above also holds when Q_1 is replaced by Q_s . Since $R_\eta(Q_s) = R_\eta(Q_1) - \log_2 s$, we have, for any $s > 0$,

$$\begin{aligned} & (1 - 2^{-p}) 2^{R_0(Q_1) - \log_2 s} + 2^{-p} 2^{R_{p+1}(Q_1) - \log_2 s} \geq \frac{p+1}{p} (1 - 2^{-p}) \log_2 e, \\ & (1 - 2^{-p}) s^{-1} 2^{R_0(Q_1)} + s^p 2^{-p} 2^{R_{p+1}(Q_1)} \geq \frac{p+1}{p} (1 - 2^{-p}) \log_2 e. \end{aligned}$$

Substituting

$$s = \left(\frac{1}{p} (1 - 2^{-p}) 2^{R_0(Q_1) + p R_{p+1}(Q_1)} \right)^{\frac{1}{p+1}},$$

we have

$$\begin{aligned} & \frac{p+1}{p} (1 - 2^{-p}) \log_2 e \\ & \leq (1 - 2^{-p}) s^{-1} 2^{R_0(Q_1)} + s^p 2^{-p} 2^{R_{p+1}(Q_1)} \\ & = (1 - 2^{-p}) \left(\frac{1}{p} (1 - 2^{-p}) 2^{R_0(Q_1) + p R_{p+1}(Q_1)} \right)^{-\frac{1}{p+1}} 2^{R_0(Q_1)} + \left(\frac{1}{p} (1 - 2^{-p}) 2^{R_0(Q_1) + p R_{p+1}(Q_1)} \right)^{\frac{p}{p+1}} 2^{-p} 2^{R_{p+1}(Q_1)} \\ & = (1 - 2^{-p}) \frac{p}{p+1} p^{\frac{1}{p+1}} 2^{\frac{p}{p+1} R_0(Q_1) - \frac{p}{p+1} R_{p+1}(Q_1)} + (1 - 2^{-p}) \frac{p}{p+1} p^{-\frac{p}{p+1}} 2^{\frac{p}{p+1} R_0(Q_1) - \frac{p}{p+1} R_{p+1}(Q_1)} \\ & = \left((1 - 2^{-p}) 2^{R_0(Q_1) - R_{p+1}(Q_1)} \right)^{\frac{p}{p+1}} \left(p^{\frac{1}{p+1}} + p^{-\frac{p}{p+1}} \right). \end{aligned}$$

Hence,

$$\begin{aligned} & \left((1 - 2^{-p}) 2^{R_0(Q_1) - R_{p+1}(Q_1)} \right)^{\frac{p}{p+1}} \geq \frac{(p+1)(1 - 2^{-p}) \log_2 e}{p^{1 + \frac{1}{p+1}} + p^{\frac{1}{p+1}}} \\ & = p^{-\frac{1}{p+1}} (1 - 2^{-p}) \log_2 e, \\ & (1 - 2^{-p}) 2^{R_0(Q_1) - R_{p+1}(Q_1)} \geq p^{-\frac{1}{p}} \left((1 - 2^{-p}) \log_2 e \right)^{\frac{p+1}{p}}, \\ & 2^{R_0(Q_1) - R_{p+1}(Q_1)} \geq p^{-\frac{1}{p}} (1 - 2^{-p})^{\frac{1}{p}} (\log_2 e)^{\frac{p+1}{p}}, \end{aligned}$$

$$\begin{aligned}
R_0(Q_1) - R_{p+1}(Q_1) &\geq \log_2 \left(p^{-\frac{1}{p}} (1 - 2^{-p})^{\frac{1}{p}} (\log_2 e)^{\frac{p+1}{p}} \right) \\
&= \frac{1}{p} \log_2 \left(\frac{1 - 2^{-p}}{p} (\log_2 e)^{p+1} \right).
\end{aligned}$$

■

REFERENCES

- [1] V. Koshelev, "Hierarchical coding of discrete sources," *Problemy peredachi informatsii*, vol. 16, no. 3, pp. 31–49, 1980.
- [2] W. H. Equitz and T. M. Cover, "Successive refinement of information," *IEEE Transactions on Information Theory*, vol. 37, no. 2, pp. 269–275, 1991.
- [3] B. Rimoldi, "Successive refinement of information: Characterization of the achievable rates," *IEEE Transactions on Information Theory*, vol. 40, no. 1, pp. 253–259, 1994.
- [4] H. Brunk and N. Farvardin, "Fixed-rate successively refinable scalar quantizers," in *Proceedings of Data Compression Conference - DCC '96*, March 1996, pp. 250–259.
- [5] H. Jafarkhani, H. Brunk, and N. Farvardin, "Entropy-constrained successively refinable scalar quantization," in *Proceedings DCC '97. Data Compression Conference*, March 1997, pp. 337–346.
- [6] M. Effros, "Practical multi-resolution source coding: TSVQ revisited," in *Proceedings DCC'98 Data Compression Conference (Cat. No. 98TB100225)*, IEEE, 1998, pp. 53–62.
- [7] Xiaolin Wu and S. Dumitrescu, "On optimal multi-resolution scalar quantization," in *Proceedings DCC 2002. Data Compression Conference*, April 2002, pp. 322–331.
- [8] S. Dumitrescu and X. Wu, "Algorithms for optimal multi-resolution quantization," *Journal of Algorithms*, vol. 50, no. 1, pp. 1 – 22, 2004.
- [9] M. Effros and D. Dugatkin, "Multiresolution vector quantization," *IEEE Transactions on Information Theory*, vol. 50, no. 12, pp. 3130–3145, Dec 2004.
- [10] J. K. Wolf, A. D. Wyner, and J. Ziv, "Source coding for multiple descriptions," *The Bell System Technical Journal*, vol. 59, no. 8, pp. 1417–1426, 1980.
- [11] L. Ozarow, "On a source-coding problem with two channels and three receivers," *Bell System Technical Journal*, vol. 59, no. 10, pp. 1909–1921, 1980.
- [12] A. El Gamal and T. Cover, "Achievable rates for multiple descriptions," *IEEE Transactions on Information Theory*, vol. 28, no. 6, pp. 851–857, 1982.
- [13] R. W. Yeung, "Multilevel diversity coding with distortion," *IEEE Transactions on Information Theory*, vol. 41, no. 2, pp. 412–422, 1995.
- [14] V. A. Vaishampayan, "Design of multiple description scalar quantizers," *IEEE Transactions on Information Theory*, vol. 39, no. 3, pp. 821–834, 1993.
- [15] S. Lloyd, "Least squares quantization in PCM," *IEEE transactions on information theory*, vol. 28, no. 2, pp. 129–137, 1982.
- [16] J. Max, "Quantizing for minimum distortion," *IRE Transactions on Information Theory*, vol. 6, no. 1, pp. 7–12, 1960.
- [17] S. M. Ross, *Stochastic processes*, 2nd ed. New York: Wiley, 1995.