A Necessary and Sufficient Condition for the Asymptotic Tightness of the Shannon Lower Bound

Tobias Koch
Universidad Carlos III de Madrid, Spain
& Gregorio Marañón Health Research Institute
Email: koch@tsc.uc3m.es

***Abstract*—The Shannon lower bound is one of the few lower bounds on the rate-distortion function that holds for a large class of sources. In this paper, it is demonstrated that its gap to the rate-distortion function vanishes as the allowed distortion tends to zero for all sources that have a finite differential entropy and whose integer parts have a finite entropy. Conversely, it is demonstrated that if the integer part of the source has an infinite entropy, then its rate-distortion function is infinite for any finite distortion. Consequently, the Shannon lower bound provides an asymptotically tight bound on the rate-distortion function if, and only if, the integer part of the source has a finite entropy.**

I. Introduction

Suppose a source produces the sequence of independent and identically distributed (i.i.d.), real-valued, random variables $\{X_k,\, k \in \mathbb{Z}\}$ according to the distribution P_X, and suppose that we employ a vector quantizer that produces a sequence of reconstruction symbols $\{\hat{X}_k,\, k \in \mathbb{Z}\}$ satisfying

$$\varlimsup_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \mathsf{E}\left[\left(X_k - \hat{X}_k\right)^2\right] \le D. \tag{1}$$

(We use $\varlimsup$ to denote the *limit superior* and $\varliminf$ to denote the *limit inferior*.) Rate-distortion theory tells us that if for every blocklength n and distortion constraint D we quantize the sequence of source symbols $X_1, \ldots, X_n$ to one of $e^{nR(D)}$ possible sequences of reconstruction symbols $\hat{X}_1, \ldots, \hat{X}_n$, then the smallest rate $R(D)$ (in nats per source symbol) for which there exists a vector quantizer satisfying (1) is given by [1], [2]

$$R(D) = \inf_{P_{\hat{X}|X}} I\big(X; \hat{X}\big) \tag{2}$$

where the infimum is over all conditional distributions of $\hat{X}$ given X satisfying

$$\mathsf{E}\left[\left(X - \hat{X}\right)^2\right] \le D \tag{3}$$

and where the expectation is computed with respect to the joint distribution $P_X P_{\hat{X}|X}$. Here and throughout the paper we omit the time indices where they are immaterial. The rate $R(D)$ as a function of D is referred to as the *rate-distortion function*.

This work has been supported in part by a Marie Curie Career Integration Grant through the 7th European Union Framework Programme under Grant 333680, by the Ministerio de Economía y Competitividad of Spain under Grants TEC2013-41718-R, RYC-2014-16332, and TEC2015-69648-REDC, and by the Comunidad de Madrid under Grant S2013/ICE-2845.

Unfortunately, the rate-distortion function is unknown except in very few special cases. It therefore needs to be assessed by means of upper and lower bounds. Arguably, for sources with a finite differential entropy, the most important lower bound is the *Shannon lower bound* [1], [2]

$$R_{\mathrm{SLB}}(D) = h(X) - \frac{1}{2}\log(2\pi e D) \tag{4}$$

where $\log(\cdot)$ denotes the natural logarithm. While this lower bound is tight only for some special sources, it converges to the rate-distortion function as the allowed distortion D tends to zero, provided that the source satisfies some conditions [3]–[6]. Thus, in this case the Shannon lower bound provides a good approximation of the rate-distortion function for small distortions. A finite-blocklength refinement of the Shannon lower bound has recently been given by Kostina [7].

To the best of our knowledge, the most general conditions for the asymptotic tightness of the Shannon lower bound are due to Linder and Zamir [6]. While Linder and Zamir considered more general distortion measures, specialized to quadratic distortion (3), they showed the following.

Theorem 1 (Linder and Zamir [6, Cor. 1]): Assume that X has a probability density function (pdf) and $h(X) > -\infty$. Further assume that there exists an $\alpha > 0$ such that $\mathsf{E}[|X|^\alpha] < \infty$. Then the Shannon lower bound is asymptotically tight, i.e.,

$$\lim_{D\downarrow 0}\big\{R(D) - R_{\mathrm{SLB}}(D)\big\} = 0. \tag{5}$$

The theorem's conditions are very mild and satisfied by the most common source distributions. In fact, Theorem 1 demonstrates that the Shannon lower bound is asymptotically tight even if there exists no quantizer with a finite number of codevectors and of finite distortion, i.e., when $\mathsf{E}\big[X^2\big] = \infty$. However, the conditions are more stringent than the ones sometimes required for the analysis of the rate and distortion redundancies of high-resolution quantizers. For example, in [8] Gray et al. analyzed the asymptotic distortion of entropy-constrained vector quantization in the limit as the rate tends to infinity, thereby rigorously proving a theorem by Zador [9]. In their work, they considered source vectors X that have a density, whose differential entropy is finite, and that satisfy

$$H(\lfloor X \rfloor) < \infty \tag{6}$$

where $\lfloor a \rfloor$ denotes the integer part of a, i.e., the largest integer not larger than a. Furthermore, Koch and Vazquez-Vilar [10] demonstrated that these assumptions are also sufficient to recover the result by Gish and Pierce [11] that, among all scalar quantizers, uniform quantizers are asymptotically optimal as the allowed distortion tends to zero. In words, condition (6) demands that quantizing the source with a uniform quantizer of unit-length cells gives rise to a discrete random variable of finite entropy. This ensures that the quantizer output can be further compressed using a lossless variable-length code of finite expected length.

The quantity $H(\lfloor X \rfloor)$ is intimately related with the *Rényi information dimension* [12], defined as

$$d(X) \triangleq \lim_{m\to\infty} \frac{H(\lfloor mX \rfloor / m)}{\log m}, \quad \text{if the limit exists} \tag{7}$$

which in turn coincides with the *rate-distortion dimension* introduced by Kawabata and Dembo [13]; see also [14]. Indeed, the Rényi information dimension is finite if, and only if, condition (6) is satisfied [14, Prop. 1]. Furthermore, a sufficient condition for finite Rényi information dimension is $\mathsf{E}[\log(1+|X|)] < \infty$ [14, Prop. 1], which in turn holds for any source for which $\mathsf{E}[|X|^\alpha] < \infty$ for some $\alpha > 0$. Thus, (6) is weaker than the assumption that $\mathsf{E}[|X|^\alpha] < \infty$.

It is common to assume that the differential entropy of the source is finite, since otherwise the Shannon lower bound (4) is uninteresting. One may thus wonder how (6) and the assumption of a finite differential entropy are related. As demonstrated, for example, in the proof of Theorem 3 in [15], condition (6) implies that $h(X) < \infty$. In fact, one can show that if (6) holds and X has a pdf, then $h(X) \leq H(\lfloor X \rfloor)$ [16, Cor. 1]. Conversely, one can find sources for which the differential entropy is finite but $H(\lfloor X \rfloor)$ is infinite. For example, consider a source with pdf

$$f_X(x) = \sum_{m=2}^{\infty} p_m m \, \mathbb{1}\left\{ m \leq x < m + \frac{1}{m} \right\}, \quad x \in \mathbb{R} \tag{8}$$

where

$$p_m = \frac{1}{\mathsf{K} m \log^2 m}, \qquad m = 2, 3, \dots \tag{9a}$$

$$\mathsf{K} = \sum_{m=2}^{\infty} \frac{1}{m \log^2 m} \tag{9b}$$

and $\mathbb{1}\{\cdot\}$ denotes the indicator function. It is easy to check that for such a source

$$\begin{aligned} H(\lfloor X \rfloor) &= \sum_{m=2}^{\infty} p_m \log \frac{1}{p_m} \\ &= \sum_{m=2}^{\infty} \frac{\log \mathsf{K} + \log m + 2 \log \log m}{\mathsf{K} m \log^2 m} \\ &= \infty \end{aligned} \tag{10}$$

and

$$\begin{aligned} h(X) &= -\int_{\mathbb{R}} f_X(x) \log f_X(x) \, \mathrm{d}x \\ &= \sum_{m=2}^{\infty} \frac{\log \mathsf{K} + 2 \log \log m}{\mathsf{K} m \log^2 m} \\ &< \infty. \end{aligned} \tag{11}$$

(See remark after Theorem 1 in [12, pp. 197–198].) Thus, for sources satisfying $h(X) > -\infty$, a finite Rényi information dimension implies a finite differential entropy but not vice versa.

In this paper, we demonstrate that for sources that have a pdf and whose differential entropy is finite, the Shannon lower bound (4) is asymptotically tight if (6) is satisfied. This ensures the asymptotic tightness of the Shannon lower bound under the most general conditions imposed in the analysis of high-resolution quantizers. Conversely, we demonstrate that for sources that do not satisfy (6) the rate-distortion function is infinite for any finite distortion.

II. Problem Setup and Main Result

We consider a one-dimensional, real-valued source X with support $\mathcal{X} \subseteq \mathbb{R}$ whose distribution is absolutely continuous with respect to the Lebesgue measure, and we denote its pdf by f_X. We assume that $x \mapsto f_X(x) \log f_X(x)$ is integrable, ensuring that the differential entropy

$$h(X) \triangleq -\int_{\mathcal{X}} f_X(x) \log f_X(x) \, \mathrm{d}x \tag{12}$$

is well-defined and finite. We have the following result.

Theorem 2 (Main Result): Assume that the one-dimensional, real-valued source X has a pdf and $h(X) > -\infty$. If $H(\lfloor X \rfloor) < \infty$, then the Shannon lower bound is asymptotically tight, i.e.,

$$\lim_{D \downarrow 0} \left\{ R(D) - R_{\text{SLB}}(D) \right\} = 0. \tag{13}$$

Conversely, if $H(\lfloor X \rfloor) = \infty$, then $R(D) = \infty$ for $D > 0$.

Proof: See Section III. ∎

Theorem 2 thus demonstrates that the Shannon lower bound is asymptotically tight if, and only if, $H(\lfloor X \rfloor)$ is finite.

III. Proof of Theorem 2

The proof consists of two parts. In the first part, we show that if $H(\lfloor X \rfloor) < \infty$, then the Shannon lower bound is asymptotically tight (Section III-A). In the second part, we show that if $H(\lfloor X \rfloor) = \infty$, then $R(D) = \infty$ for every $D > 0$ (Section III-B).

A. Asymptotic Tightness

In this section, we demonstrate the asymptotic tightness of the Shannon lower $R_{\text{SLB}}(D)$ for sources that satisfy $H(\lfloor X \rfloor) < \infty$ and $h(X) > -\infty$. The first steps in our proof are identical to the ones in the proof of Theorem 1 in [6]. To keep this paper self-contained, we reproduce the main steps.

To prove the asymptotic tightness of $R_{\text{SLB}}(D)$, we derive an upper bound on $R(D)$ whose gap to $R_{\text{SLB}}(D)$ vanishes as D tends to zero. In view of (2), an upper bound on $R(D)$ follows by choosing $\hat{X} = X + Z_D$, where Z_D is a zero-mean, variance-D, Gaussian random variable that is independent of X. It follows that

$$\begin{aligned} R(D) &\leq I(X; X + Z_D) \\ &= h(X + Z_D) - h(Z_D). \end{aligned} \tag{14}$$

Furthermore, using that $h(Z_D) = \frac{1}{2}\log(2\pi e D)$, the Shannon lower bound (4) can be written as

$$R_{\text{SLB}}(D) = h(X) - h(Z_D). \tag{15}$$

Combining (14) and (15) gives

$$0 \le R(D) - R_{\text{SLB}}(D) \le h(X + Z_D) - h(X). \tag{16}$$

Thus, the asymptotic tightness of $R_{\text{SLB}}(D)$ follows by proving that

$$\overline{\lim_{D\downarrow 0}}\, h(X + Z_D) \le h(X). \tag{17}$$

To this end, we follow the steps (17)–(21) in [6] but with $Y_{\Delta(D)}$ and $Y_{\Delta(0)}$ replaced by the random variables Y_D and Y_0 with respective pdfs

$$f_{Y_D}(y) = \sum_{i\in\mathbb{Z}} \Pr\bigl(\lfloor X + Z_D\rfloor = i\bigr)\mathbb{1}\{\lfloor y\rfloor = i\} \tag{18a}$$

$$f_{Y_0}(y) = \sum_{i\in\mathbb{Z}} \Pr\bigl(\lfloor X\rfloor = i\bigr)\mathbb{1}\{\lfloor y\rfloor = i\}. \tag{18b}$$

It follows that

$$D(f_{X+Z_D}\|f_{Y_D}) = H(\lfloor X + Z_D\rfloor) - h(X + Z_D) \tag{19}$$

and

$$D(f_X\|f_{Y_0}) = H(\lfloor X\rfloor) - h(X). \tag{20}$$

The random variable Z_D converges to zero almost surely as D tends to zero and, hence, also in distribution. Since X and Z_D are independent, it follows that $X + Z_D \to X$ in distribution as D tends to zero. Furthermore, since by assumption the distribution of X is absolutely continuous with respect to the Lebesgue measure, for every $i \in \mathbb{Z}$ the interval $[i, i+1)$ is a continuity set of X, so

$$\lim_{D\downarrow 0} \Pr\bigl(\lfloor X + Z_D\rfloor = i\bigr) = \Pr\bigl(\lfloor X\rfloor = i\bigr), \quad i \in \mathbb{Z}. \tag{21}$$

Thus, the pdf of Y_D (18a) converges pointwise to the pdf of Y_0 (18b), which by Scheffe's lemma [17, Th. 16.12] implies that $Y_D \to Y_0$ in distribution as D tends to zero.

By the lower semicontinuity of relative entropy (see, e.g., the proof of Lemma 4 in [18] and references therein),

$$\underline{\lim_{D\downarrow 0}}\, D(f_{X+Z_D}\|f_{Y_D}) \ge D(f_X\|f_{Y_0}). \tag{22}$$

Combining (22) with (19) and (20) yields

$$\underline{\lim_{D\downarrow 0}}\bigl\{H(\lfloor X+Z_D\rfloor) - h(X+Z_D)\bigr\} \ge H(\lfloor X\rfloor) - h(X). \tag{23}$$

Since $h(X) > -\infty$ and $H(\lfloor X\rfloor) < \infty$, the claim (17) (and hence the asymptotic tightness of $R_{\text{SLB}}(D)$) follows by showing that $H(\lfloor X + Z_D\rfloor)$ tends to $H(\lfloor X\rfloor)$ as D tends to zero. We present this result in the following lemma.

Lemma 1: Assume that X has a pdf and $H(\lfloor X\rfloor) < \infty$. Let Z_D be a zero-mean, variance-D, Gaussian random variable that is independent of X. Then

$$\lim_{D\downarrow 0} H(\lfloor X + Z_D\rfloor) = H(\lfloor X\rfloor). \tag{24}$$

Proof: Using basic properties of entropy, we obtain

$$\begin{aligned} H(\lfloor X + Z_D\rfloor) &\le H(\lfloor X\rfloor) + H\bigl(\lfloor X + Z_D\rfloor \bigm| \lfloor X\rfloor\bigr) \\ &\le H(\lfloor X\rfloor) + H(V_D) \end{aligned} \tag{25}$$

and

$$\begin{aligned} H(\lfloor X + Z_D\rfloor) &\ge H(\lfloor X\rfloor) - H\bigl(\lfloor X\rfloor \bigm| \lfloor X + Z_D\rfloor\bigr) \\ &\ge H(\lfloor X\rfloor) - H(V_D) \end{aligned} \tag{26}$$

where $V_D \triangleq \lfloor X + Z_D\rfloor - \lfloor X\rfloor$. Lemma 1 follows therefore by showing that $H(V_D)$ vanishes as D tends to zero.

We first show that

$$\lim_{D\downarrow 0} \Pr(V_D = i) = \mathbb{1}\{i = 0\}. \tag{27}$$

Indeed, let $\bar{X} \triangleq X - \lfloor X\rfloor$, and recall that $Z_D \to 0$ in distribution as D tends to zero. Noting that $V_D = \lfloor \bar{X} + Z_D\rfloor$, the probability mass function of V_D can be written as

$$\Pr\bigl(V_D = i\bigr) = \Pr\bigl(\lfloor \bar{X} + Z_D\rfloor = i\bigr), \quad i \in \mathbb{Z}. \tag{28}$$

Furthermore, the independence of X and Z_D implies that $\bar{X} + Z_D \to \bar{X}$ in distribution as D tends to zero. Since the distribution of X is absolutely continuous with respect to the Lebesgue measure, so is the distribution of $\bar{X}$. Consequently, for every $i \in \mathbb{Z}$ the interval $[i, i+1)$ is a continuity set of $\bar{X}$ and

$$\lim_{D\downarrow 0} \Pr\bigl(\lfloor \bar{X} + Z_D\rfloor = i\bigr) = \Pr\bigl(\lfloor \bar{X}\rfloor = i\bigr) = \mathbb{1}\{i = 0\} \tag{29}$$

where the last step follows because the support of $\bar{X}$ is $[0, 1)$. This proves (27).

We continue by expressing the entropy of V_D as

$$\begin{aligned} H(V_D) = &\sum_{i=-1}^{1} \Pr(V_D = i)\log\frac{1}{\Pr(V_D = i)} \\ &+ \sum_{i\in\mathbb{Z}:\ |i|>1} \Pr(V_D = i)\log\frac{1}{\Pr(V_D = i)}. \end{aligned} \tag{30}$$

The first sum on the right-hand side (RHS) of (30) consists of finitely many terms, so (27) and the continuity of $x \mapsto x\log(1/x)$ give[1]

$$\begin{aligned} &\lim_{D\downarrow 0} \sum_{i=-1}^{1} \Pr(V_D = i)\log\frac{1}{\Pr(V_D = i)} \\ &\quad = \sum_{i=-1}^{1} \lim_{D\downarrow 0} \Pr(V_D = i)\log\frac{1}{\Pr(V_D = i)} \\ &\quad = 0. \end{aligned} \tag{31}$$

To show that the second sum on the RHS of (30) vanishes as $D \to 0$, it suffices to show that

$$\overline{\lim_{D\downarrow 0}} \sum_{i\in\mathbb{Z}:\ |i|>1} \Pr(V_D = i)\log\frac{1}{\Pr(V_D = i)} \le 0 \tag{32}$$

since the summands are nonnegative. As observed above, the distribution of $\bar{X}$ is absolutely continuous with respect to the

[1] Here and throughout the paper we define $0\log(1/0) \triangleq 0$.

Lebesgue measure, so $\bar{X}$ has a pdf which we shall denote by $f_{\bar{X}}$. Since Z_D and $\bar{X}$ are independent, the pdf of $\bar{X}+Z_D$ is given by [19, Th. 4.10, p. 29]

$$f_{\bar{X}+Z_D}(\xi)=\int_0^1 f_{\bar{X}}(\bar{x})\frac{1}{\sqrt{2\pi D}}e^{-\frac{(\xi-\bar{x})^2}{2D}}\,\mathrm{d}\bar{x},\quad \xi\in\mathbb{R}. \tag{33}$$

Combining (28) and (33), we obtain

$$\begin{aligned}\Pr\big(V_D=i\big)&=\Pr\big(\lfloor\bar{X}+Z_D\rfloor=i\big)\\&=\int_i^{i+1}\int_0^1 f_{\bar{X}}(\bar{x})\frac{1}{\sqrt{2\pi D}}e^{-\frac{(\xi-\bar{x})^2}{2D}}\,\mathrm{d}\bar{x}\,\mathrm{d}\xi\\&\geq\frac{1}{\sqrt{2\pi D}}e^{-\frac{(|i|+1)^2}{2D}},\quad i\in\mathbb{Z}\end{aligned} \tag{34}$$

where the inequality follows because for $\xi\in[i,i+1)$ and $\bar{x}\in[0,1)$ we have $|\xi-\bar{x}|\leq|i|+1$. Applying (34) to the second sum on the RHS of (30) gives

$$\begin{aligned}\sum_{i\in\mathbb{Z}\colon|i|>1}&\Pr(V_D=i)\log\frac{1}{\Pr(V_D=i)}\\&\leq\frac{1}{2}\log(2\pi D)\sum_{i\in\mathbb{Z}\colon|i|>1}\Pr(V_D=i)\\&\quad+\sum_{i\in\mathbb{Z}\colon|i|>1}\Pr(V_D=i)\frac{(|i|+1)^2}{2D}.\end{aligned} \tag{35}$$

To demonstrate that the first term on the RHS of (35) vanishes as $D\to 0$, we use that any variable z satisfying $|\lfloor\bar{x}+z\rfloor|>1$ must also satisfy $|z|>1$, irrespective of $\bar{x}\in[0,1)$. Consequently,

$$\begin{aligned}\sum_{i\in\mathbb{Z}\colon|i|>1}\Pr(V_D=i)&=\Pr\big(|\lfloor\bar{X}+Z_D\rfloor|>1\big)\\&\leq\Pr\big(|Z_D|>1\big)\\&\leq D\end{aligned} \tag{36}$$

where the last inequality follows from Chebyshev's inequality [20, Th. 4.10.7, p. 192]. Combining (36) with (35), we obtain

$$\left|\frac{1}{2}\log(2\pi D)\sum_{i\in\mathbb{Z}\colon|i|>1}\Pr(V_D=i)\right|\leq\frac{D}{2}\big|\log(2\pi D)\big| \tag{37}$$

which tends to zero as $D\to 0$.

We next consider the second term on the RHS of (35). To this end, we write

$$\Pr(V_D=i)(|i|+1)^2=\int_i^{i+1}f_{\bar{X}+Z_D}(\xi)(|i|+1)^2\,\mathrm{d}\xi. \tag{38}$$

By Fubini's theorem [20, Th. 2.6.4, p. 105], we obtain from (38) and (33) that

$$\begin{aligned}&\sum_{i\in\mathbb{Z}\colon|i|>1}\Pr(V_D=i)(|i|+1)^2\\&=\sum_{i\in\mathbb{Z}\colon|i|>1}\int_i^{i+1}\int_0^1 f_{\bar{X}}(\bar{x})\frac{1}{\sqrt{2\pi D}}e^{-\frac{(\xi-\bar{x})^2}{2D}}(|i|+1)^2\,\mathrm{d}\bar{x}\,\mathrm{d}\xi\\&=\int_0^1 f_{\bar{X}}(\bar{x})\sum_{i\in\mathbb{Z}\colon|i|>1}\int_{i-\bar{x}}^{i+1-\bar{x}}\frac{(|i|+1)^2}{\sqrt{2\pi D}}e^{-\frac{z^2}{2D}}\,\mathrm{d}z\,\mathrm{d}\bar{x}.\end{aligned} \tag{39}$$

For every $|i|=2,3,\ldots$, $z\in[i-\bar{x},i+1-\bar{x})$, and $\bar{x}\in[0,1)$ we have $|i|+1\leq 3|z|$. Hence,

$$\begin{aligned}\sum_{i\in\mathbb{Z}\colon|i|>1}&\int_{i-\bar{x}}^{i+1-\bar{x}}\frac{(|i|+1)^2}{\sqrt{2\pi D}}e^{-\frac{z^2}{2D}}\,\mathrm{d}z\\&\leq\sum_{i\in\mathbb{Z}\colon|i|>1}\int_{i-\bar{x}}^{i+1-\bar{x}}\frac{9z^2}{\sqrt{2\pi D}}e^{-\frac{z^2}{2D}}\,\mathrm{d}z\\&\leq 9\int_{\{|z|\geq 1\}}\frac{z^2}{\sqrt{2\pi D}}e^{-\frac{z^2}{2D}}\,\mathrm{d}z\end{aligned} \tag{40}$$

where the last inequality follows because, for every $\bar{x}\in[0,1)$,

$$\bigcup_{i\in\mathbb{Z}\colon|i|>1}[i-\bar{x},i+1-\bar{x})\subseteq\{z\in\mathbb{R}\colon|z|\geq 1\}.$$

The RHS of (40) does not depend on $\bar{x}$, so together with (39) this yields

$$\sum_{i\in\mathbb{Z}\colon|i|>1}\Pr(V_D=i)(|i|+1)^2\leq 9\mathsf{E}\big[Z_D^2\mathbb{1}\{|Z_D|\geq 1\}\big]. \tag{41}$$

Writing Z_D as $Z_D=\sqrt{D}Z_1$, where Z_1 is a zero-mean, unit-variance, Gaussian random variable, the expected value on the RHS of (41) can be written as

$$\mathsf{E}\big[Z_D^2\mathbb{1}\{|Z_D|\geq 1\}\big]=D\mathsf{E}\big[Z_1^2\mathbb{1}\big\{Z_1^2\geq 1/D\big\}\big]. \tag{42}$$

Combining (42) and (41), we obtain

$$\begin{aligned}\sum_{i\in\mathbb{Z}\colon|i|>1}&\Pr(V_D=i)\frac{(|i|+1)^2}{2D}\\&\leq\frac{9}{2}\mathsf{E}\big[Z_1^2\mathbb{1}\big\{Z_1^2\geq 1/D\big\}\big].\end{aligned} \tag{43}$$

Since the function $z\mapsto z^2\mathbb{1}\big\{z^2\geq 1/D\big\}$ is dominated by $z\mapsto z^2$, and since $\mathsf{E}\big[Z_1^2\big]=1$, it follows from the Dominated Convergence Theorem [20, Th. 1.6.9, p. 50] that

$$\lim_{D\downarrow 0}\mathsf{E}\big[Z_1^2\,\mathbb{1}\big\{Z_1^2\geq 1/D\big\}\big]=0. \tag{44}$$

Together with (43) this demonstrates that the second term on the RHS of (35) vanishes as D tends to zero.

Thus, (35), (37), (43), and (44) prove (32), which together with (30) and (31) demonstrates that

$$\lim_{D\downarrow 0}H(V_D)=\lim_{D\downarrow 0}\sum_{i\in\mathbb{Z}}\Pr(V_D=i)\log\frac{1}{\Pr(V_D=i)}=0. \tag{45}$$

This was the last step required to prove Lemma 1. ■

Combining Lemma 1 with (23) implies (17), which in turn demonstrates that the Shannon lower bound is asymptotically tight if $H(\lfloor X\rfloor)<\infty$ and $h(X)>-\infty$. This proves the first part of Theorem 2.

B. Infinite Rate-Distortion Function

To prove that $H(\lfloor X \rfloor) = \infty$ implies $R(D) = \infty$ for every $D > 0$, we show that $I(X;\hat{X}) = \infty$ for every pair of random variables $(X,\hat{X})$ satisfying (3) and $H(\lfloor X \rfloor) = \infty$. To this end, we follow along the lines of the proof of Theorem 6 in [16, App. A]. Indeed, by the Data Processing Inequality [21, Cor 7.16],

$$I(X;\hat{X}) \geq I\bigl(\lfloor X \rfloor; \lfloor \hat{X} \rfloor\bigr). \tag{46}$$

The mutual information on the RHS of (46) can be written as

$$I\bigl(\lfloor X \rfloor; \lfloor \hat{X} \rfloor\bigr) = H\bigl(\lfloor X \rfloor\bigr) - H\bigl(\lfloor X \rfloor \bigm| \lfloor \hat{X} \rfloor\bigr). \tag{47}$$

Since $H(\lfloor X \rfloor) = \infty$ by assumption, the claim follows by showing that the conditional entropy on the RHS of (47) is bounded for every pair of random variables $(X,\hat{X})$ satisfying (3). Indeed, we have

$$\begin{aligned} H\bigl(\lfloor X \rfloor \bigm| \lfloor \hat{X} \rfloor\bigr) &\leq H\bigl(\lfloor X - \hat{X} \rfloor\bigr) \\ &\quad + H\bigl(\lfloor X \rfloor \bigm| \lfloor \hat{X} \rfloor, \lfloor X - \hat{X} \rfloor\bigr). \end{aligned} \tag{48}$$

Since $\mathsf{E}[\log(1 + |X - \hat{X}|)] < \infty$ for $(X,\hat{X})$ satisfying (3), Proposition 1 in [14] yields that

$$H\bigl(\lfloor X - \hat{X} \rfloor\bigr) < \infty. \tag{49}$$

Furthermore, denoting $Y = X - \hat{X}$, we obtain

$$\begin{aligned} H\bigl(\lfloor X \rfloor \bigm| \lfloor \hat{X} \rfloor, \lfloor X - \hat{X} \rfloor\bigr) &= H\bigl(\lfloor \hat{X} + Y \rfloor \bigm| \lfloor \hat{X} \rfloor, \lfloor Y \rfloor\bigr) \\ &\leq \log 2 \end{aligned} \tag{50}$$

since, conditioned on $\lfloor \hat{X} \rfloor$ and $\lfloor Y \rfloor$, the random variable $\lfloor \hat{X} + Y \rfloor$ can only take on the values $\lfloor \hat{X} \rfloor + \lfloor Y \rfloor$ or $\lfloor \hat{X} \rfloor + \lfloor Y \rfloor + 1$. Combining (48)–(50) yields

$$H\bigl(\lfloor X \rfloor \bigm| \lfloor \hat{X} \rfloor\bigr) < \infty. \tag{51}$$

Summing up, (46)–(51) demonstrate that $I(X;\hat{X}) = \infty$ for every pair of random variables $(X,\hat{X})$ satisfying (3) and $H(\lfloor X \rfloor) = \infty$. Hence, the rate-distortion function $R(D)$ is infinite for every finite D. This proves the second part of Theorem 2.

IV. Conclusions

The Shannon lower bound is one of the few lower bounds on the rate-distortion function that hold for a large class of sources. We have demonstrated that this lower bound is asymptotically tight as the allowed distortion vanishes for all sources having a finite differential entropy and a finite Rényi information dimension. Conversely, we have demonstrated that if the source has an infinite Rényi information dimension, then its rate-distortion function is infinite for any finite distortion.

Assuming a finite Rényi information dimension is tantamount to assuming that quantizing the source with a uniform scalar quantizer of unit-length cells gives rise to a discrete random variable of finite entropy. The latter assumption is natural in rate-distortion theory and often encountered. To this effect, we have demonstrated that this assumption is not only natural, but it is also a necessary and sufficient condition for the asymptotic tightness of the Shannon lower bound.

Finally, the presented results can be generalized to d-dimensional, real-valued sources and distortion measures of the form $\|\mathbf{x} - \hat{\mathbf{x}}\|^r$, where $\|\cdot\|$ is an arbitrary norm on $\mathbb{R}^d$ and $r > 0$. For details, see our paper [22] on arXiv.

Acknowledgment

The author wishes to thank Helmut Bölcskei, David Stotz, and Gonzalo Vazquez-Vilar for helpful discussions. The author further wishes to thank Giuseppe Durisi and Tamás Linder for calling his attention to references [14] and [15], respectively.

References

[1] C. E. Shannon, "Coding theorems for a discrete source with a fidelity criterion," *IRE International Convention Record*, vol. 7, pp. 142–163, 1959.

[2] T. Berger, *Rate Distortion Theory: Mathematical Basis for Data Compression*, ser. Electrical Engineering Series. Prentice Hall, 1971.

[3] Y. N. Linkov, "Evaluation of epsilon entropy of random variables for small epsilon," *Problemy Peredachi Informatsii (Problems of Inform. Transm.)*, vol. 1, pp. 12–18, 1965.

[4] A. M. Gerrish and P. M. Schultheiss, "Information rates of non-Gaussian processes," *IEEE Trans. Inform. Theory*, vol. 10, pp. 265–271, Oct. 1964.

[5] J. Binia, M. Zakai, and J. Ziv, "On the ϵ-entropy and the rate-distortion function of certain non-Gaussian process," *IEEE Trans. Inform. Theory*, vol. 20, pp. 514–524, July 1974.

[6] T. Linder and R. Zamir, "On the asymptotic tightness of the Shannon lower bound," *IEEE Trans. Inform. Theory*, vol. 40, no. 6, pp. 2026–2031, Nov. 1994.

[7] V. Kostina, "Data compression with low distortion and finite blocklength," in *Proc. 53rd Allerton Conf. Comm., Contr. and Comp.*, Allerton H., Monticello, Il, Sep. 30 – Oct. 2, 2015.

[8] R. M. Gray, T. Linder, and J. Li, "A Langrangian formulation of Zador's entropy-constrained quantization theorem," *IEEE Trans. Inform. Theory*, vol. 28, no. 3, pp. 695–707, Mar. 2002.

[9] P. L. Zador, "Topics in the asymptotic quantization of continuous random variables," Bell Laboratories, Tech. Rep., 1966.

[10] T. Koch and G. Vazquez-Vilar, "Rate-distortion bounds for high-resolution vector quantization via Gibbs's inequality," July 2015. [Online]. Available: http://arxiv.org/abs/1507.08349

[11] H. Gish and J. N. Pierce, "Asymptotically efficient quantizing," *IEEE Trans. Inform. Theory*, vol. 14, no. 5, pp. 676–683, Sept. 1968.

[12] A. Rényi, "On the dimension and entropy of probability distributions," *Acta Mathematica Hungarica*, vol. 10, no. 1–2, Mar. 1959.

[13] T. Kawabata and A. Dembo, "The rate-distortion dimension of sets and measures," *IEEE Trans. Inform. Theory*, vol. 40, no. 5, pp. 1564–1572, Sept. 1994.

[14] Y. Wu and S. Verdú, "Rényi information dimension: Fundamental limits of almost lossless analog compression," *IEEE Trans. Inform. Theory*, vol. 56, no. 8, pp. 3721–3748, Aug. 2010.

[15] I. Csiszár, "Some remarks on the dimension and entropy of random variables," *Acta Mathematica Hungarica*, vol. 12, no. 3–4, pp. 399–408, Sept. 1961.

[16] D. Stotz and H. Bölcskei, "Degrees of freedom in vector interference channels," Sept. 26, 2014, subm. to *IEEE Trans. Inform. Theory*. [Online]. Available: www.nari.ee.ethz.ch/commth/pubs/p/dof_transit

[17] P. Billingsley, *Probability and Measure*, 3rd ed., ser. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons, 1995.

[18] I. Csiszár, "Arbitrarily varying channel with general alphabets and states," *IEEE Trans. Inform. Theory*, vol. 38, no. 6, pp. 1725–1742, Nov. 1992.

[19] R. Durrett, *Probability: Theory and Examples*, 3rd ed. Brooks/Cole, 2005.

[20] R. B. Ash and C. A. Doléans-Dade, *Probability and Measure Theory*, 2nd ed. Elsevier/Academic Press, 2000.

[21] R. M. Gray, *Entropy and Information Theory*, 2nd ed. Springer Verlag, 2011.

[22] T. Koch, "The Shannon lower bound is asymptotically tight for sources with a finite Rényi information dimension," Apr. 2015. [Online]. Available: http://arxiv.org/abs/1504.08245