Wolfgang Kreitmeier

Optimal quantization for the one–dimensional uniform distribution with Rényi-α-entropy constraints

Kybernetika, Vol. 46 (2010), No. 1, 96--113

Persistent URL: http://dml.cz/dmlcz/140047

KYBERNETIKA — VOLUME 46 (2010), NUMBER 1, PAGES 96–113

OPTIMAL QUANTIZATION FOR THE ONE–DIMENSIONAL UNIFORM DISTRIBUTION WITH RÉNYI-α-ENTROPY CONSTRAINTS

Wolfgang Kreitmeier

We establish the optimal quantization problem for probabilities under constrained Rényi-α-entropy of the quantizers. We determine the optimal quantizers and the optimal quantization error of one-dimensional uniform distributions including the known special cases $\alpha = 0$ (restricted codebook size) and $\alpha = 1$ (restricted Shannon entropy).

Keywords: optimal quantization, uniform distribution, Rényi-α-entropy

Classification: 60Exx, 62H30, 94A17, 94A29

1. INTRODUCTION AND BASIC NOTATION

The quantization of probability distributions is mainly motivated from electrical engineering in context of signal processing and data compression. A good survey about the historical development of the theory has been provided by Gray and Neuhoff [9]. The reader is also referred to the book of Gersho and Gray [6] for more applied aspects. Optimal quantization is the task of finding a best approximation of a given probability by another probability with reduced complexity. Complexity constraints used so far are restricted memory size resp. restricted Shannon entropy of the approximation. The approximating probability is always induced by a quantizer, which decomposes the space into codecells. The deviation of an optimal approximation from the original probability is described by the so-called optimal quantization error. An exact determination of the optimal quantization error and the appropriate optimal quantizers (if existing) subject to a given finite bound on the complexity of the approximation is quite hard. Only for a few classes of – mainly one-dimensional – distributions and restricted memory size exact results exist (cf. [5, 7, 8, 15, 16]). If we deal with Shannon entropy as constraint, then only the optimal quantization problem of the one-dimensional uniform distribution on $[0, 1]$ was solved completely so far (cf. [11]).

Although exact analytical results in optimal Shannon-entropy-constrained quantization are rare, this quantization method normally outperforms quantization with fixed memory size. Unfortunately, the optimal quantization error as a function of the Shannon-entropy bound turns out to be non-convex (cf. [11, Corollary 2]), but convexity would be an import property for applications (cf. [11] and the references therein). In this paper we will use Rényi-α-entropy as complexity constraint, which comprises the known special cases $\alpha = 0$ (restricted codebook size) and $\alpha = 1$ (restricted Shannon entropy). The family of Rényi-α-entropies can be characterized axiomatically (cf. [4] chapter 1.2.1) and generalizes the Shannon entropy in a canonical way. Insofar this fact motivates from a mathematical viewpoint to use Rényi-α-entropy as a complexity constraint. Moreover, Grendar [10] introduced a generalized support for discrete probabilities, which is in direct relationship to the family of Rényi-α-entropies. Insofar, quantization with entropy parameter $\alpha > 0$ is a natural generalization of classical memory-size constrained quantization ($\alpha = 0$). In addition, Rényi-α-entropy-constrained quantization has already been motivated – at least indirectly – by Harremoës and Topsøe [14, VII. B./F.]. We will completely solve the problem of optimal quantization for the one-dimensional uniform distribution on $[0, 1]$ with Rényi-α-entropy as complexity constraint. Our main result (cf. Theorem 3.1) contains the special cases $\alpha = 0$ (restricted codebook size) and $\alpha = 1$ (restricted Shannon entropy). To this end we follow in our proofs the lines of György and Linder [11]. For certain values of α this quantization technique even outperforms Shannon-entropy-constrained quantization and leads to a strict convex quantization error function (cf. Section 4 and Figure).

The paper is organized as follows. The remaining part of Section 1 introduces the exact definitions and notations for optimal quantization with Rényi-α-entropy. Section 2 contains fundamental properties of optimal scalar quantization, which are needed in the sequel but also interesting in itself. In Section 3 the optimal quantization problem for the one-dimensional uniform distribution on $[0, 1]$ with Rényi-α-entropy will be solved completely (cf. Theorem 3.1). Section 4 contains some results about the analytical behavior of the optimal quantization error against the bound on the complexity. The last section comprises an appendix with some technical results.

Let $\mathbb{N} := \{1, 2, \ldots\}$. Let $\alpha \in [0, \infty]$ and $p = (p_1, p_2, \ldots) \in [0, 1]^{\mathbb{N}}$ be a probability vector, i. e. $\sum_{i=1}^{\infty} p_i = 1$. The Rényi-$\alpha$-entropy $\tilde{H}^{\alpha}(p) \in [0, \infty]$ is defined as (see e. g. [1, Definition 5.2.35] resp. [4, Chapter 1.2.1])

$$\tilde{H}^{\alpha}(p) = \begin{cases} -\sum_{i=1}^{\infty} p_i \log(p_i), \text{ if } \alpha = 1 \\ -\log\left(\sup\{p_i : i \in \mathbb{N}\}\right), \text{ if } \alpha = \infty \\ \frac{1}{1-\alpha} \log\left(\sum_{i=1}^{\infty} p_i^{\alpha}\right), \text{ if } \alpha \in [0, \infty[\setminus\{1\}. \end{cases}$$

We use the convention $0 \cdot \log(0) := 0$ and $0^x := 0$ for all real x. The logarithm log is based on e.

Remark 1.1. With these conventions we obtain for $\alpha = 0$ that

$$\tilde{H}^{\alpha}(p) = \log\left(\text{card}\{p_i : i \in \mathbb{N}, p_i > 0\}\right),$$

if card denotes cardinality. Using de l'Hospital it is easy to see, that the case $\alpha = 1$

will be reached from $\alpha \neq 1$ by taking the limit $\alpha \to 1$. (cf. [1, Remark 5.2.34]). Moreover it is $\lim_{\alpha\to\infty} \tilde{H}^\alpha(\cdot) = \tilde{H}^\infty(\cdot)$.

Now let $d \in \mathbb{N}$ and μ be a Borel probability measure on $\mathbb{R}^d$. Denote with $\mathcal{F}_d$ the set of all Borel-measurable mappings $f : \mathbb{R}^d \to \mathbb{R}^d$ with $\operatorname{card}(f(\mathbb{R}^d)) \leq \operatorname{card}(\mathbb{N})$. A mapping $f \in \mathcal{F}_d$ is called *quantizer* and the image $f(\mathbb{R}^d)$ is called a codebook consisting of codepoints. Every quantizer f induces a partition $\{f^{-1}(z) : z \in f(\mathbb{R}^d)\}$ of $\mathbb{R}^d$. Every element of this partition is called a *codecell*. The image measure $\mu \circ f^{-1}$ has a countable support and defines an approximation of μ, the so-called quantization of μ by f. For any enumeration $\{z_1, z_2, \ldots\}$ of $f(\mathbb{R}^d)$ we define

$$H^\alpha_\mu(f) = \tilde{H}^\alpha(\mu \circ f^{-1}(z_1), \mu \circ f^{-1}(z_2), \ldots)$$

as the Rényi-α-entropy of f w.r.t μ. Now we intend to quantify the distance between μ and its approximation under f. To this end let $\|\cdot\|$ be any norm on $\mathbb{R}^d$ and $\rho : [0,\infty[\to [0,\infty[$ a strict monotone increasing (and therefore Borel-measurable) mapping. For $f \in \mathcal{F}_d$ we define as distance between μ and $\mu \circ f^{-1}$ the quantization error

$$D_\mu(f) = \int \rho(\| x - f(x) \|) \, \mathrm{d}\mu(x).$$

For any $R \geq 0$ we denote

$$D^\alpha_\mu(R) = \inf\{D_\mu(f) : f \in \mathcal{F}_d, H^\alpha_\mu(f) \leq R\} \tag{1}$$

as the optimal quantization error of μ under Rényi-α-entropy bound R. We denote with

$$C^\alpha_\mu(R) = \{f \in \mathcal{F}_d : D_\mu(f) = D^\alpha_\mu(R)\}$$

the set of all optimal quantizers of μ under Rényi-α-entropy bound R. Later we will see (cf. Theorem 3.1) that $C^\alpha_\mu(R)$ can be empty.

The problem of optimal quantization comprises the calculation of the optimal quantization error $D^\alpha_\mu(R)$ and the determination of the set of all optimal quantizers $C^\alpha_\mu(R)$. In general, this problem is rather hard to solve, even for the special cases $\alpha \in \{0, 1\}$. In this paper we will solve this problem completely for the one-dimensional uniform distribution on $[0, 1]$ and $\rho(x) = x^r$ with $r > 1$. Motivated from the problem of classification, another approach in quantization of a probability exists (cf. [17, 18]) and is based on the concept of f−divergence of probabilities. Although this approach coincides with our quantization approach in some special cases (choose $\alpha = 0$, $r = 2$ and $f(t) = (t-1)^2$) we do not use it in this paper.

2. FUNDAMENTAL PROPERTIES OF OPTIMAL SCALAR QUANTIZATION

We shortly write $\mathcal{F} = \mathcal{F}_1$. For any measurable $A \subset \mathbb{R}$ with $\mu(A) > 0$ we denote with $\mu(\cdot|A)$ the conditional probability of μ w.r.t. A. We first make the following observation (cf. [12, Theorem 2]).

Proposition 2.1. Let $\alpha \in [0,\infty]$ and $r \geq 1$. Let μ be a non-atomic Borel probability measure on $\mathbb{R}$ and $R \geq 0$. Then for the calculation of the optimal quantization error it suffices to consider only those quantizers

(i) whose Rényi-α-entropy attains the bound R if $\alpha > 0$,

(ii) whose codecells with positive $\mu-$mass are intervals,

(iii) where every codepoint a of a codecell $f^{-1}(a)$ with positive μ-mass induces an optimal quantizer f_a for $\mu(\cdot|f^{-1}(a))$, i. e. $f_a \in C^{\alpha}_{\mu(\cdot|f^{-1}(a))}(0)$ with $f_a(x) = a$ for every $x \in \mathbb{R}$.

P r o o f. If $R = 0$ there is nothing to prove. For $\alpha = 0$ the assertion follows from [8, Theorem 4.1]. Let $R > 0$ and $\alpha > 0$. Let $f \in \mathcal{F}$ with $H^{\alpha}_{\mu}(f) < R$. György and Linder have shown (cf. [12, Theorem 2]) that a quantizer $g \in \mathcal{F}$ can be constructed with the same positive codecell probabilities as f. All codecells of g with positive μ-mass are intervals and $D_\mu(g) \le D_\mu(f)$. Hence, $H^{\alpha}_{\mu}(f) = H^{\alpha}_{\mu}(g)$. As a consequence we can assume w.l.o.g. that all codecells of f with positive μ-mass are intervals. Now choose $a \in f(\mathbb{R})$ with $\mu(f^{-1}(a)) > 0$. Let

$$t_a = \sup\left\{t \in \mathbb{R} : \mu(]-\infty, t]) \le \frac{1}{2}\mu(f^{-1}(a))\right\}.$$

Because μ is non-atomic we have $\mu(]-\infty, t_a]) = \frac{1}{2}\mu(f^{-1}(a))$. Let a_1 be such that

$$\int_{f^{-1}(a)\cap]-\infty,t_a]} |x-a_1|^r \, \mathrm{d}\mu(x) = \inf\left\{\int_{f^{-1}(a)\cap]-\infty,t_a]} |x-z|^r \, \mathrm{d}\mu(x) : z \in \mathbb{R}\right\},$$

resp. a_2 such that

$$\int_{f^{-1}(a)\cap]t_a,\infty[} |x-a_2|^r \, \mathrm{d}\mu(x) = \inf\left\{\int_{f^{-1}(a)\cap]t_a,\infty[} |x-z|^r \, \mathrm{d}\mu(x) : z \in \mathbb{R}\right\}.$$

Such points a_1, a_2 exist (see e. g. [8, Lemma 2.2]) We define the quantizer

$$f_{a,t_a}(z) = \begin{cases} f(z), \text{ if } z \in \mathbb{R}\backslash f^{-1}(a) \\ a_1, \text{ if } z \in f^{-1}(a) \cap]-\infty, t_a] \\ a_2, \text{ if } z \in f^{-1}(a) \cap]t_a, \infty[. \end{cases}$$

Clearly, $D_\mu(f_{a,t_a}) \le D_\mu(f)$. Moreover, $H^{\alpha}_{\mu}(f_{a,t_a}) \ge H^{\alpha}_{\mu}(f)$, where equality can only appear in case of $\alpha = \infty$. By dividing the codecells with positive μ-mass in above manner (if necessary) we can assume w.l.o.g. that $H^{\alpha}_{\mu}(f) < R$ and $H^{\alpha}_{\mu}(f_{b,t_b}) \ge R$, if $b \in f(\mathbb{R})$ with

$$\mu(f^{-1}(b)) = \max\{\mu(f^{-1}(c)) : c \in f(\mathbb{R})\}.$$

Because μ is non-atomic and $\alpha > 0$, we can find a $t \le t_b$ with

$$\mu(]-\infty, t] \cap f^{-1}(b)) > 0$$

and $H^{\alpha}_{\mu}(f_{b,t}) = R$. Again it is $D_\mu(f_{b,t}) \le D_\mu(f)$. Moreover, all codecells of $f_{b,t}$ with positive μ-mass remain intervals and thus the assertions (i) and (ii) are proved. Assertion (iii) is obvious. $\square$

Remark 2.2. The proof of Proposition 2.1 shows also, that only quantizers, whose entropy attains the bound R can be optimal. If μ equals the uniform distribution on a compact interval, then it is obvious, that only those quantizers can be optimal, whose codecells with positive probability are consisting of intervals.

Lemma 2.3. Let $R \in [0, \infty[$ and μ be a Borel probability on $\mathbb{R}$. Then the mapping

$$[0, \infty] \ni \alpha \to D_\mu^\alpha(R)$$

is decreasing.

Proof. Let $f \in \mathcal{F}$ with $H_\mu^\alpha(f) \le R$. For arbitrary $0 < \gamma \le \beta < \infty$ we have (cf. [3, p. 53])

$$H_\mu^\beta(f) \le H_\mu^\gamma(f). \tag{2}$$

Together with Remark 1.1 we conclude that inequality (2) also holds for $0 \le \gamma \le \beta \le \infty$. Then the assertion follows immediately from Definition (1) of the optimal quantization error. □

Let T denote a similitude. The last result of this section determines how the optimal quantization error changes if we replace μ by $\mu \circ T^{-1}$. For $\alpha = 0$ the reader is also referred to [8, Lemma 3.2].

Lemma 2.4. Let $\alpha \in [0, \infty]$ and $T : \mathbb{R} \to \mathbb{R}$ be a similarity transformation with scaling number $c > 0$. Then for any $R \ge 0$ we have

$$D_{\mu \circ T^{-1}}^\alpha(R) = c^r D_\mu^\alpha(R)$$

and

$$C_{\mu \circ T^{-1}}^\alpha(R) = \{T^{-1} \circ f \circ T : f \in C_\mu^\alpha(R)\}.$$

Proof. For arbitrary $f \in \mathcal{F}$ we have $H_{\mu \circ T^{-1}}^\alpha(f) = H_\mu^\alpha(T^{-1} \circ f \circ T)$. Thus we get

$$\begin{aligned}
& D_{\mu \circ T^{-1}}^\alpha(R) \\
= \; & \inf\{\int |x - f(x)|^r \, \mathrm{d}\mu \circ T^{-1}(x) : f \in \mathcal{F}, H_{\mu \circ T^{-1}}^\alpha(f) \le R\} \\
= \; & c^r \inf\{\int |x - T^{-1} \circ f \circ T(x)|^r \, \mathrm{d}\mu(x) : f \in \mathcal{F}, H_\mu^\alpha(T^{-1} \circ f \circ T) \le R\} \\
= \; & c^r D_\mu^\alpha(R).
\end{aligned}$$

□

3. OPTIMAL QUANTIZATION OF THE UNIFORM DISTRIBUTION

Let $-\infty < a < b < \infty$ and denote with $U([a,b])$ the uniform distribution on $[a,b]$. To study the optimal quantization of this distribution, it suffices in view of Lemma 2.4 to consider only the special case $a = 0, b = 1$. For $\alpha = 1$ the following result is due to György and Linder [11]. In case of $\alpha = 0$ the reader is referred to Graf and Luschgy (cf. [8, Example 5.5]). We are following the approach for $\alpha = 1$. Let $Q(r) = \frac{1}{(1+r)2^r}$. For $R \geq 0$ and $\alpha \in [0,\infty]\backslash\{1\}$ we define the set

$$A(\alpha, R) = \left\{ p = (p_1, p_2, \ldots) \in [0,1]^{\mathbb{N}} : \sum_{i=1}^{\infty} p_i = 1, \tilde{H}^\alpha(p) = R \right\}. \tag{3}$$

For the rest of this paper we will assume w.l.o.g. that every codecell of a quantizer for $U([0,1])$ has positive $U([0,1])-$mass. Finally we denote for any quantizer $f \in \mathcal{F}$ and $A \subset \mathbb{R}$ with $f|_A$ the restriction of f to A.

Theorem 3.1. Let $r > 1$ and $\alpha \in [0,\infty]$.
Let $n \in \mathbb{N}$ and $R \in \,]\log(n), \log(n+1)]$. If $\alpha \leq r+1$, then an optimal quantizer always exists. Every restriction $f|_{]0,1[}$ of an optimal quantizer $f \in C^\alpha_{U([0,1])}(R) \neq \emptyset$ consists of $n+1$ interval cells, with one cell having length $p \in \,]0, 1/(n+1)]$ and n cells having length $(1-p)/n$, where p satisfies

$$R = \begin{cases} \dfrac{1}{1-\alpha} \log\left(p^\alpha + n\left(\dfrac{1-p}{n}\right)^\alpha\right), \text{ if } \alpha \in [0, r+1[\backslash\{1\} \\ -p\log(p) - (1-p)\log\left(\dfrac{1-p}{n}\right), \text{ if } \alpha = 1. \end{cases}$$

For each codecell of $f|_{]0,1[}$ the codepoint is the midpoint of the cell. Moreover with $p = p(R)$ we obtain the optimal quantization error

$$D^\alpha_{U([0,1])}(R) = Q(r)\left(p^{r+1} + n\left(\frac{1-p}{n}\right)^{r+1}\right). \tag{4}$$

If $\alpha > r+1$, then no optimal quantizer exist, i. e. $C^\alpha_{U([0,1])}(R) = \emptyset$. The optimal quantization error turns into

$$D^\alpha_{U([0,1])}(R) = \begin{cases} Q(r)e^{-(r+1)\frac{\alpha-1}{\alpha}R}, \text{ if } \alpha \in \,]r+1, \infty[\\ Q(r)e^{-(r+1)R}, \text{ if } \alpha = \infty. \end{cases} \tag{5}$$

Remark 3.2. It is easy to check, that the representation (5) of the optimal quantization error turns into (4), if $\alpha = r+1$. Figure illustrates $D^\alpha_{U([0,1])}(\cdot)$ for different values of α and $r = 2$.

P r o o f o f T h e o r e m 3.1. For $\alpha = 1$ the assertion was proved by György and Linder [11]. The case $\alpha = 0$ is treated by Graf and Luschgy [8, Example 5.5]. Thus

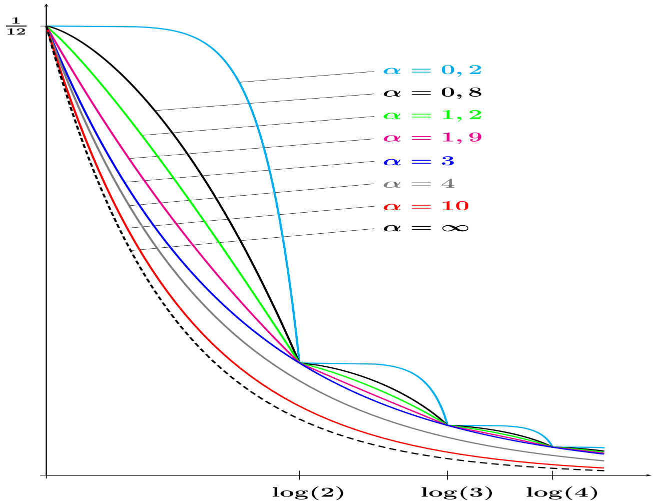

Fig. $D^{\alpha}_{U([0,1])}(\cdot)$ with $r = 2$.

we will assume that $\alpha \notin \{0,1\}$. Let $n \in \mathbb{N}$ and $R \in \,]\log(n), \log(n+1)]$. Using Proposition 2.1 (i) we obtain

$$D^{\alpha}_{U([0,1])}(R) = \inf\left\{D_{U([0,1])}(f) : f \in \mathcal{F}, H^{\alpha}_{U([0,1])}(f) = R\right\}. \tag{6}$$

From Proposition 2.1 (ii) we know, that it suffices to consider on the right hand side of (6) only quantizers $f \in \mathcal{F}$, where f consists of intervals. Due to Proposition 2.1 (iii) and with $r > 1$ we deduce from [8, Theorem 2.4] that the midpoint is the codepoint of each codecell of $f|_{]0,1[}$. According to Remark 2.2, every optimal quantizer (if existing) must necessarily have all these properties. Thus we get

$$\begin{aligned} D^{\alpha}_{U([0,1])}(R) &= \inf\left\{2\sum_{i=1}^{\infty}\int_0^{p_i/2} x^r \,\mathrm{d}x : (p_1, p_2, \ldots) \in A(\alpha, R)\right\} \\ &= Q(r) \cdot \inf\left\{\sum_{i=1}^{\infty} p_i^{r+1} : (p_1, p_2, \ldots) \in A(\alpha, R)\right\}. \end{aligned} \tag{7}$$

1. $\alpha \in \,]0, r+1[\,\backslash\{1\}$.
In the special case $R = \log(n+1)$ the assertion follows from Lemma A.5. Now let $R \in \,]\log(n), \log(n+1)[$. Let $k \in \mathbb{N}$ and $A_k(\alpha, R) \subset A(\alpha, R)$ be the set of all $(p_1, p_2, \ldots) \in A(\alpha, R)$ consisting of exactly k strictly positive components. Let us

assume w.l.o.g. that $p_{k+j}=0$ for every $j\geq 1$ and $(p_1,p_2,\ldots)\in A_k(\alpha,R)$. First we analyze if an $(p_1,p_2,\ldots)\in A_k(\alpha,R)$ exists, where

$$D^{\alpha}_{U([0,1])}(R)=Q(r)\sum_{i=1}^{k}p_i^{r+1}. \tag{8}$$

For any $(p_1,p_2,\ldots)\in A_k(\alpha,R)$ we have

$$\log(n)<R=\frac{1}{1-\alpha}\log\left(\sum_{i=1}^{k}p_i^{\alpha}\right)\leq\log(k).$$

Here the last inequality follows from the maximality of the Rényi-α-entropies (cf. [1, Remark 2, p.149]). Thus we have $k\geq n+1$. Finding an $(p_1,p_2,\ldots)\in A_k(\alpha,R)$, which satisfies (8) is equivalent to determine a global minimum of the mapping

$$]0,\infty[^k\ni(p_1,\ldots,p_k)\xrightarrow{D}Q(r)\sum_{i=1}^{k}p_i^{r+1} \tag{9}$$

subject to the constraints

$$g(p_1,\ldots,p_k)=\sum_{i=1}^{k}p_i-1=0;\quad h(p_1,\ldots,p_k)=\frac{1}{1-\alpha}\log\left(\sum_{i=1}^{k}p_i^{\alpha}\right)-R=0. \tag{10}$$

Now let $u=(u_1,\ldots,u_k)\in\,]0,\infty[^k$ and assume that u is a solution of this minimization problem. We first note, that the gradients of h and g in u are linearly dependent if and only if all u_i are equal. But in this case we would get $R=\log(k)\geq\log(n+1)$ which contradicts $R\in\,]\log(n),\log(n+1)[$. Thus it remains to deal with the case that the gradients of h and g are linearly independent. In this situation the Kuhn-Tucker conditions (cf. [2, Theorem 4.3.7]) have to be satisfied, yielding for every $j=1,\ldots,k$

$$(u_j/2)^r+\lambda_1\frac{\alpha}{(1-\alpha)\sum_{i=1}^{k}u_i^{\alpha}}u_j^{\alpha-1}+\lambda_2=0 \tag{11}$$

with $\lambda_1\geq 0$ and $\lambda_2\in\mathbb{R}$. Thus (11) is equivalent to

$$u_j^r=Au_j^{\alpha-1}+B$$

with $A=\frac{-\lambda_1\alpha 2^r}{(1-\alpha)\sum_{i=1}^{k}u_i^{\alpha}}$ resp. $B=-2^r\lambda_2$. If $\alpha\in\,]0,1[$, then $A\leq 0$. The mapping $]0,\infty[\,\ni x\to Ax^{\alpha-1}$ is concave. Together with $r>1$ this implies that (11) has at most two distinct solutions. If $\alpha\in\,]1,r+1[$, then $A\geq 0$. In case of $\alpha\in\,]1,2[$ the mapping $]0,\infty[\,\ni x\to Ax^{\alpha-1}$ is strictly concave. Thus (11) has at most two distinct solutions. If $\alpha=2$, then $r>1$ yields that (11) has at most two solutions. If $\alpha\in\,]2,r+1[$ then the strict convexity of $]0,\infty[\,\ni x\to Ax^{\alpha-1}$ implies that (11) has at most two distinct solutions. Therefore we can assume that $p,q\in\,]0,1[$ and $k_1\in\{1,\ldots,k\}$ exist, with $k_1q+(k-k_1)p=1$ and $p_i=q$ if $i=1,\ldots,k_1$ resp. $p_i=p$ if $i=k_1+1,\ldots,k$. Note that $k_1=k$ would yield with (10) the contradiction

$R = \log(k) \geq \log(n+1)$. Hence we have $k_1 \in \{1, \ldots, k-1\}$. Moreover we can assume w.l.o.g. that $q \geq p$, which yields $q \geq \frac{1}{k} \geq p$. Again, $q = p = 1/k$ would yield the contradiction $R \geq \log(n+1)$. Hence, $q > p$. If $q = 0$, then $p = 1/k_1 = 1/k$ would yield the contradiction $R \geq \log(n+1)$. We have $p = (1-k_1 q)/(k-k_1)$. Using (10) we define

$$\begin{aligned} h^*(k, k_1, q, R) &= h(p_1(k, k_1, q, R), \ldots, p_k(k, k_1, q, R)) \\ &= \frac{1}{1-\alpha} \log\left(k_1 q^\alpha + (k-k_1)^{1-\alpha}(1-k_1 q)^\alpha\right) - R. \end{aligned}$$

In view of (9) and with $q = q(k, k_1, R)$ we define

$$\begin{aligned} D^*(k, k_1, R) &= D(p_1(k, k_1, R), \ldots, p_k(k, k_1, R)) \\ &= Q(r)(k_1 q^{r+1} + (k-k_1)^{-r}(1-k_1 q)^{r+1}) \end{aligned}$$

and denote $\tilde{D}(k_1, q, k) = Q(r)(k_1 q^{r+1} + (k-k_1)^{-r}(1-k_1 q)^{r+1})$. Although $D^*(k, \cdot, R)$ has been defined only for integers, the defining formulas are allowing us to enlarge the definition to any real $k_1 \in [1, k-1]$. We will show, that $D^*(k, \cdot, R)$ is strictly decreasing on $[1, k-1]$. To this end we will show that $\frac{\partial D^*}{\partial k_1}(k, \cdot, R) < 0$ on $]1, k-1[$. First we observe that

$$\frac{\partial h^*}{\partial k_1} = \frac{q^\alpha - p^\alpha - (q-p)\alpha p^{\alpha-1}}{(1-\alpha)e^{(1-\alpha)R}}. \tag{12}$$

Additionally we calculate

$$\frac{\partial h^*}{\partial q} = \frac{\alpha k_1}{e^{(1-\alpha)R}} \left(\frac{q^{\alpha-1} - p^{\alpha-1}}{1-\alpha}\right). \tag{13}$$

Next we compute

$$\begin{aligned} \frac{\partial \tilde{D}}{\partial k_1} &= Q(r)\left(q^{r+1} + \frac{\partial}{\partial k_1}\left((k-k_1)p^{r+1}\right)\right) \\ &= Q(r)\left(q^{r+1} - p^{r+1} + (k-k_1)(r+1)p^r \frac{\partial p}{\partial k_1}\right) \\ &= Q(r)\left(q^{r+1} - p^{r+1} + (r+1)p^r(p-q)\right), \end{aligned} \tag{14}$$

resp.

$$\begin{aligned} \frac{\partial \tilde{D}}{\partial q} &= k_1 \left(\frac{q}{2}\right)^r + (k-k_1)\left(\frac{p}{2}\right)^r \frac{\partial p}{\partial q} \\ &= k_1 \left(\left(\frac{q}{2}\right)^r - \left(\frac{p}{2}\right)^r\right). \end{aligned} \tag{15}$$

From $q > p$ we get $\frac{\partial h^*}{\partial p} \neq 0$. Thus we can apply implicit differentiation and deduce

$$\frac{\partial D^*}{\partial k_1} = \frac{\partial \tilde{D}}{\partial k_1} + \frac{\partial \tilde{D}}{\partial q}\frac{\partial q}{\partial k_1} = \frac{\partial \tilde{D}}{\partial k_1} - \frac{\partial \tilde{D}}{\partial q}\frac{\partial h^*}{\partial k_1}\left(\frac{\partial h^*}{\partial q}\right)^{-1}.$$

With (12), (13), (14) and (15) we compute

$$\begin{aligned}\frac{\partial D^*}{\partial k_1} &= \frac{q^{r+1}-p^{r+1}}{2^r(r+1)}+\left(\frac{p}{2}\right)^r(p-q)\\ &\quad -\left(\left(\frac{q}{2}\right)^r-\left(\frac{p}{2}\right)^r\right)\frac{q^\alpha-p^\alpha-(q-p)\alpha p^{\alpha-1}}{\alpha\left(q^{\alpha-1}-p^{\alpha-1}\right)}.\end{aligned}$$

Now it is easy to see that $\frac{\partial D^*}{\partial k_1}<0$ if and only if

$$\frac{1}{r+1}\frac{q^{r+1}-p^{r+1}-(q-p)p^r(r+1)}{q^r-p^r}<\frac{1}{\alpha}\frac{q^\alpha-p^\alpha-(q-p)p^{\alpha-1}\alpha}{q^{\alpha-1}-p^{\alpha-1}}. \tag{16}$$

In view of Lemma A.3 we know that inequality (16) is true. Searching for a global minimum of D^* we have to restrict ourself to $k_1=k-1$. Next we will show, that $D^*(\cdot,1,R)$ is strictly increasing on $\{n+1,n+2,\ldots\}$. To this end we again enlarge the definition of $D^*(\cdot,k-1,R)$ to any real $k\in[n+1,\infty[$ and show that $\frac{\partial D^*}{\partial k}(\cdot,k-1,R)>0$ on $]n+1,\infty[$. We compute

$$\frac{\partial\tilde{D}}{\partial k}=-rQ(r)p^{r+1}.$$

Moreover we have

$$\frac{\partial h^*}{\partial k}=e^{-(1-\alpha)R}p^\alpha$$

Using implicit differentiation and (13) resp. (15) we obtain

$$\begin{aligned}\frac{\partial D^*}{\partial k} &= \frac{\partial\tilde{D}}{\partial k}+\frac{\partial\tilde{D}}{\partial q}\frac{\partial q}{\partial k}=\frac{\partial\tilde{D}}{\partial k}-\frac{\partial\tilde{D}}{\partial q}\frac{\partial h^*}{\partial k}\left(\frac{\partial h^*}{\partial q}\right)^{-1}\\ &= Q(r)\left(-rp^{r+1}+\frac{(r+1)(\alpha-1)p^\alpha\left(p^r-q^r\right)}{\alpha(p^{\alpha-1}-q^{\alpha-1})}\right).\end{aligned}$$

Hence, $\frac{\partial D^*}{\partial k}>0$ if and only if

$$\frac{\frac{r}{r+1}p^{r+1}}{p^r-q^r}>\frac{\frac{\alpha-1}{\alpha}p^\alpha}{p^{\alpha-1}-q^{\alpha-1}}. \tag{17}$$

Using Lemma A.4 and $\alpha<r+1$ we know, that (17) holds. Thus we have proved that $k=n+1$ and the infimum in (7) can only be attained by $(p,q,\ldots,q,0,\ldots)\in A_{n+1}(\alpha,R)$ or by $(p_1,p_2,\ldots)\in A(\alpha\ R)\backslash\cup_{j=1}^{\infty}A_j(\alpha,R)$. Using the results from above we obtain for any $(p_1,p_2,\ldots)\in A(\alpha\ R)\backslash\cup_{j=1}^{\infty}A_j(\alpha,R)$ that

$$\begin{aligned}\sum_{i=1}^{\infty}p_i^{r+1} &= \lim_{\substack{k\to\infty\\ k\geq n+1}}\left(\sum_{i=1}^{k}\left(\frac{p_i}{\sum_{l=1}^{k}p_l}\right)^{r+1}\right)\\ &> nq^{r+1}+p^{r+1}.\end{aligned}$$

Hence $(p, q, \ldots, q, 0, \ldots)$ is the only element of $A(\alpha, R)$ with

$$D^{\alpha}_{U([0,1])}(R) = Q(r) \sum_{i=1}^{\infty} p_i^{r+1}$$

and thus the assertion is proved for $\alpha \in [0, r+1[$.

2. $\alpha > r+1$.

In this case the assertion follows immediately from Lemma A.5.

3. $\alpha = r+1$.

Using Lemma 2.3 the assertion follows from step 1 and 2. □

Remark 3.3. If we drop the second constraint $h(p_1, \ldots, p_k) = 0$ in (10), then it follows from [14, Theorem III.1.i)] that it suffices to consider for the global minimum of the mapping in (9) only those probability vectors, who are consisting of $k-1$ equal components $p \in]0,1[$ and one component $q = 1-(k-1)p$. If we take also the second constraint into consideration, then the proof of Theorem 3.1 shows that this property remains.

Remark 3.4. Harremoës and Topsøe (cf. [14, VII. F.]) proposed the very general term *information diagram* for diagrams built with values from Shannon theory, prediction and universal coding, rate distortion analysis, error probability analysis etc. Hence our Figure represents also such an information diagram. Moreover they pointed out (cf. [14, p. 2947-2948]) that the research of György and Linder [11] can be interpreted as an information diagram related to the one's they have studied. They also motivated (cf. [14, VII. B.]) to study diagrams where Rényi-α-entropy is plotted against Shannon-entropy for discrete probabilities or to investigate even more generalized diagrams (cf. [14, VII. H.]). Insofar the results of this paper are heavily motivated by the earlier results in [11] and [14]. Finally the term complexity class has been used in [14]. In our setting the complexity class k reflects the set of all quantizers f for $U([0,1])$ with $H^{\alpha}_{U([0,1])}(f) \in]\log(k), \log(k+1)]$. Hence also this term fits naturally into the framework of this paper.

4. ANALYTICAL PROPERTIES OF THE OPTIMAL QUANTIZATION ERROR FUNCTION

In view of Theorem 3.1 we know that the mapping

$$[0, \infty[\ni R \to D^{\alpha}_{U([0,1])}(R)$$

is strictly convex and differentiable for $\alpha \geq r+1$. Moreover we deduce from Theorem 3.1 that $D^{\alpha}_{U([0,1])}(\cdot)$ is differentiable on $\cup_{n=1}^{\infty}]\log(n), \log(n+1)[$ for $\alpha < r+1$. Next we determine the right- and left-hand limit of the derivative at $\log(n)$.

Proposition 4.1. Let $n \in \mathbb{N}$ and $r' = \frac{\alpha-1}{\alpha}(r+1) > 0$ for $\alpha \in]1, \infty[$. Then

$$Q(r)^{-1} \lim_{\substack{R \to \log(n) \\ R > \log(n)}} \frac{\mathrm{d}D^{\alpha}_{\mu}(R)}{\mathrm{d}R} = \begin{cases} 0, \text{ if } \alpha \in [0,1] \\ -r' n^{-r}, \text{ if } \alpha \in]1, r+1] \\ -r' n^{-r'}, \text{ if } \alpha \in]r+1, \infty[\\ -(r+1) n^{-(r+1)}, \text{ if } \alpha = \infty \ . \end{cases}$$

and

$$Q(r)^{-1} \lim_{\substack{R\to\log(n+1)\\ R<\log(n+1)}} \frac{\mathrm{d}D_\mu^\alpha(R)}{\mathrm{d}R} = \begin{cases} -\frac{r(r+1)}{\alpha}(n+1)^{-r}, \text{ if } \alpha \in]0, r+1] \\ -r'(n+1)^{-r'}, \text{ if } \alpha \in]r+1, \infty[\\ -(r+1)(n+1)^{-(r+1)}, \text{ if } \alpha = \infty\ . \end{cases}$$

P r o o f. The case $\alpha = 0$ is obvious. Let $n \in \mathbb{N}$ and $R \in]\log(n), \log(n+1)[$. First assume that $\alpha \in]0, r+1[$. Thus we have with $p = p(R) \in [0, 1/n[$ the parametric representation

$$\begin{aligned} R &= \frac{1}{1-\alpha}\log\left(p^\alpha + n\left(\frac{1-p}{n}\right)^\alpha\right) \\ Q(r)^{-1}D_\mu^\alpha(R) &= D(p) = p^{r+1} + n\left(\frac{1-p}{n}\right)^{r+1}. \end{aligned}$$

We calculate

$$\begin{aligned} Q(r)^{-1}\frac{\mathrm{d}D_\mu^\alpha(R)}{\mathrm{d}R} &= \frac{\partial D}{\partial p}\left(\frac{\partial R}{\partial p}\right)^{-1} \\ &= \frac{r+1}{\frac{\alpha}{1-\alpha}} \frac{\left(p^r - \left(\frac{1-p}{n}\right)^r\right)\left(p^\alpha + n^{1-\alpha}(1-p)^\alpha\right)}{p^{\alpha-1} - n^{1-\alpha}(1-p)^{\alpha-1}} \end{aligned}$$

and

$$Q(r)^{-1} \lim_{R\to\log(n)} \frac{\mathrm{d}D_\mu^\alpha(R)}{\mathrm{d}R} = \frac{\frac{\alpha-1}{\alpha}(r+1)n^{-r}n^{1-\alpha}}{\lim_{p\to 0}\left(p^{\alpha-1} - n^{1-\alpha}(1-p)^{\alpha-1}\right)}$$

resp. with de l'Hospital

$$\begin{aligned} Q(r)^{-1} \lim_{R\to\log(n+1)} \frac{\mathrm{d}D_\mu^\alpha(R)}{\mathrm{d}R} &= -r'(n+1)^{1-\alpha} \lim_{p\to\frac{1}{n+1}} \frac{p^r - (\frac{1-p}{n})^r}{p^{\alpha-1} - (\frac{1-p}{n})^{\alpha-1}} \\ &= -\frac{r(r+1)}{\alpha}(n+1)^{-r} \end{aligned}$$

which yields the assertion for $\alpha \in [0, r+1]$. If $\alpha \in]r+1, \infty]$ the assertion follows directly from the analytical representation $R \to D_\mu^\alpha(R) = Q(r)e^{-(r+1)\frac{\alpha-1}{\alpha}R}; \alpha \in]r+1, \infty[$ resp. $R \to D_\mu^\infty(R) = Q(r)e^{-(r+1)R}$ of the quantization error function. □

From Proposition 4.1 we get that the quantization error function $D_{U([0,1])}^\alpha(\cdot)$ is non-differentiable in $\log(n)$ for every $n \geq 2$ and $\alpha \in [0, r+1]$. Moreover Proposition 4.1 yields that $D_{U([0,1])}^\alpha(\cdot)$ is neither concave nor convex in general for $\alpha \in [0, r+1[$.

APPENDIX

Lemma A.1. Let $a > 1$. Then

$$\log(a) > \frac{2(a-1)}{a+1}. \tag{18}$$

Proof. We compute

$$\log(a) - 2\frac{a-1}{a+1} = \int_1^a \left(\frac{1}{x} - \frac{4}{(x+1)^2}\right) \mathrm{d}x = \int_1^a \frac{(x-1)^2}{x(x+1)^2} \mathrm{d}x > 0,$$

which proves inequality (18). □

Lemma A.2. Let $a > 1$ and

$$g(x) = \begin{cases} \frac{\frac{a^x-1}{x}-a+1}{a^{x-1}-1}, \text{ if } x \in \,]0,\infty[\setminus\{1\} \\ \frac{a\log(a)-a+1}{\log(a)}, \text{ if } x = 1. \end{cases}$$

Then g is strictly monotone decreasing.

Proof. First let us note that g is continuous on $]0,\infty[$. This can be seen immediately by using de l'Hospital. On $]0,\infty[\setminus\{1\}$ the mapping g is differentiable and we compute

$$g'(x) = \frac{a^{x-1}\log(a)\left(\frac{a^x}{x} - \frac{a}{x} - \frac{a^x-1}{x} + a - 1\right) - \frac{a^x-1}{x^2}\left(a^{x-1}-1\right)}{\left(a^{x-1}-1\right)^2}.$$

Thus for every $x \in \,]0,\infty[\setminus\{1\}$ we have $g'(x) < 0$, if

$$h(x) = a^x - a - 1 + a^{1-x} - \log(a)\left((a-1)x^2 + (1-a)x\right) > 0.$$

1. $x \in \,]1,\infty[$.
A straightforward calculation gives

$$\begin{aligned} h'(x) &= \log(a)\left(a^x - a^{1-x} - 2(a-1)x - 1 + a\right) \\ h''(x) &= \log(a)\left(\log(a)\left(a^x + a^{1-x}\right) - 2(a-1)\right) \\ h'''(x) &= (\log(a))^3\left(a^x - a^{1-x}\right) \end{aligned}$$

Hence, $h'''(x) > 0$ due to $x > 1/2$. Moreover we obtain $h''(1) > 0$, if

$$\log(a)\,(a+1) - 2(a-1) > 0,$$

which is true in view of Lemma A.1. We conclude that

$$h''(x) = h''(1) + \int_1^x h'''(z)\,\mathrm{d}z > 0.$$

Thus we deduce

$$h'(x) = h'(1) + \int_1^x h''(z)dz = \int_1^x h''(z)\,\mathrm{d}z > 0$$

and, therefore

$$h(x) = h(1) + \int_1^x h'(z)\,\mathrm{d}z = \int_1^x h'(z)\,\mathrm{d}z > 0.$$

2. $x \in]0,1[$.
Obviously h is symmetric to $1/2$. Thus it suffices to consider $x \in]1/2,1[$. It is $h'(1/2) = 0 = h'(1)$. Because $x \to a^x - a^{-x}$ is strictly convex on $]1/2,\infty[$ we obtain $h'(x) < 0$ for every $x \in]1/2,1[$. Thus we get

$$h(x) = h(1) + \int_1^x h'(z)\,\mathrm{d}z = -\int_x^1 h'(z)\,\mathrm{d}z > 0.$$

Combining step 1 and 2 we obtain $g'(x) < 0$ for every $x \in]0,\infty[\setminus\{1\}$. The continuity of g on $]0,\infty[$ yields the assertion. □

Lemma A.3. Let $0 < p < q < 1$ and

$$f(x) = \begin{cases} \frac{q^x - p^x - (q-p)xp^{x-1}}{x(q^{x-1}-p^{x-1})}, \text{ if } x \in]0,\infty[\setminus\{1\} \\ \frac{q\log(q) - p\log(p) - (q-p)(1+\log(p))}{\log(q)-\log(p)}, \text{ if } x = 1. \end{cases}$$

Then f is strictly monotone decreasing.

P r o o f. With $a = q/p > 1$ and g from Lemma A.2 we obtain $f(x) = pg(x)$ for every $x > 0$. Thus the assertion follows from Lemma A.2. □

Lemma A.4. Let $0 < p < q < 1$ and

$$f(x) = \begin{cases} \frac{\frac{x-1}{x}p^x}{p^{x-1}-q^{x-1}}, \text{ if } x \in]0,\infty[\setminus\{1\} \\ \frac{p}{\log(p)-\log(q)}, \text{ if } x = 1. \end{cases}$$

Then f is strictly monotone increasing.

P r o o f. First let us note that f is continuous on $]0,\infty[$. This can be seen immediately by using de l'Hospital. On $]0,\infty[\setminus\{1\}$ the mapping f is differentiable and we compute

$$f'(x) = \frac{(\frac{p^x}{x^2} + \frac{x-1}{x}\log(p)p^x)(p^{x-1} - q^{x-1}) - \frac{x-1}{x}p^x(\log(p)p^{x-1} - \log(q)q^{x-1})}{(p^{x-1} - q^{x-1})^2}.$$

A simple calculation yields that $f'(x) > 0$ if and only if

$$\frac{1}{x}(1 - (p/q)^{x-1}) < (x-1)\log(q/p) \tag{19}$$

It is well known that for every $s > 0$ and $t > 0, t \neq 1$ the inequality

$$\frac{1}{s}(1 - t^{-s}) < \log(t) < \frac{t^s - 1}{s} \tag{20}$$

holds (see e. g. [13, p. 117]). Using (20) we obtain for $x > 1$ that

$$\frac{1}{x}(1 - (p/q)^{x-1}) < 1 - (p/q)^{x-1} = 1 - (q/p)^{-(x-1)} < (x-1)\log(q/p),$$

resp. for $x \in]0, 1[$ we deduce

$$(1 - x)\log(q/p) < (q/p)^{1-x} - 1 < \frac{1}{x}\left((q/p)^{1-x} - 1\right).$$

Hence, (19) is valid for $x \in]0, \infty[\backslash\{1\}$. As a consequence we have $f'(x) > 0$ for every $x \in]0, \infty[\backslash\{1\}$. Finally the continuity of f on $]0, \infty[$ yields the assertion. □

Recall the definition of $A(\alpha, R)$ from Eq. (3).

Lemma A.5. Let $n \in \mathbb{N}$. If $\alpha \in [0, r+1[\backslash\{1\}$, then

$$\min\left\{\sum_{i=1}^{\infty} p_i^{1+r} : (p_1, p_2, \ldots) \in A(\alpha, \log(n))\right\} = n^{-r}. \tag{21}$$

Every $(p_1, p_2, \ldots) \in A(\alpha, \log(n))$ which attains above minimum contains exactly n positive components $p_i = \frac{1}{n}$. If $\alpha \in]r+1, \infty[$ and $R > 0$, then

$$\inf\left\{\sum_{i=1}^{\infty} p_i^{1+r} : (p_1, p_2, \ldots) \in A(\alpha, R)\right\} = e^{-\frac{\alpha-1}{\alpha}(r+1)R}. \tag{22}$$

Moreover,

$$\inf\left\{\sum_{i=1}^{\infty} p_i^{1+r} : (p_1, p_2, \ldots) \in A(\infty, R)\right\} = e^{-(r+1)R}. \tag{23}$$

If $\alpha \in]r+1, \infty]$, then no probability vector $(p_1, p_2, \ldots,) \in A(\alpha, R)$ exist which attains the infimum in (22) resp. (23).

P r o o f.

1. $\alpha \in [0, r+1[\backslash\{1\}$.
Let $(p_i)_{i\in\mathbb{N}} \in A(\alpha, \log(n))$. Due to $\alpha - 1 < r$ we have by inequality of means

$$\frac{1}{n} = \left(\sum_{i=1}^{\infty} p_i p_i^{\alpha-1}\right)^{\frac{1}{\alpha-1}} \leq \left(\sum_{i=1}^{\infty} p_i p_i^{r}\right)^{\frac{1}{r}}, \tag{24}$$

which yields

$$(1/n)^r \leq \sum_{i=1}^{\infty} p_i^{r+1}. \tag{25}$$

On the other hand let $p_i = 1/n$ for every $i = 1, \ldots, n$ and $p_i = 0$ for $i > n$. Hence, $\sum_{i=1}^{\infty} p_i = 1$ and

$$\sum_{i=1}^{\infty} p_i^{\alpha} = n\,(1/n)^{\alpha} = n^{1-\alpha}.$$

Thus we get

$$\sum_{i=1}^{\infty} p_i^{1+r} = n\,(1/n)^{r+1} = (1/n)^r. \tag{26}$$

The combination of (25) and (26) proves (21). Now let $(p_1, p_2, \ldots) \in A(\alpha, \log(n))$. Assume that an $i \in \mathbb{N}$ exists, with $0 < p_i \neq \frac{1}{n}$. But then, inequality (24) will become strict. This ensures that $\sum_{i=1}^{\infty} p_i^{1+r} > n^{-r}$ and thus proves the assertion in this case.

2. $\alpha \in \,]r+1, \infty[$.
Using Jensen's inequality we obtain

$$e^{\frac{1-\alpha}{\alpha}(r+1)R} = \left(e^{(1-\alpha)R}\right)^{\frac{r+1}{\alpha}} = \left(\sum_{i=1}^{\infty} p_i^{\alpha}\right)^{\frac{r+1}{\alpha}} \leq \sum_{i=1}^{\infty} p_i^{r+1}. \tag{27}$$

On the other hand let $N > e^R$ and $p = p(N) \in \,]0,1[$ a solution of

$$p^{\alpha} + (N-1)\left(\frac{1-p}{N-1}\right)^{\alpha} = e^{(1-\alpha)R}.$$

We define the probability vector $(p_{i,N})_{i\in\mathbb{N}}$ with $p_{1,N} = p; p_{2,N}, \ldots, p_{N,N} = \frac{1-p}{N-1}$ and $p_{N+k,N} = 0$ for every $k \in \mathbb{N}$. Thus we have

$$\sum_{i=1}^{\infty} p_{i,N} = 1; \qquad \sum_{i=1}^{\infty} p_{i,N}^{\alpha} = e^{(1-\alpha)R}.$$

Moreover

$$\sum_{i=1}^{\infty} p_{i,N}^{r+1} = p^{r+1} + (N-1)\left(\frac{1-p}{N-1}\right)^{r+1} \overset{N\to\infty}{\longrightarrow} e^{\frac{1-\alpha}{\alpha}(r+1)R}. \tag{28}$$

Combing (27) with (28) we obtain (22). Now let $(p_1, p_2, \ldots) \in A(\alpha, R)$. Assume that $i, j \in \mathbb{N}$, $i \neq j$ exist with $0 < p_i < p_j$. Due to $\alpha > r+1$, inequality (27) is strict in this case. If we assume on the other hand that $(p_1, p_2, \ldots)$ contains n positive components with equal value $\frac{1}{n}$, then $n = e^R$ and we obtain

$$\sum_{i=1}^{\infty} p_i^{r+1} = e^{-rR} > e^{\frac{1-\alpha}{\alpha}(r+1)R}.$$

In both cases $(p_1, p_2, \ldots)$ does not attain the right hand side of (22).

3. $\alpha = \infty$.

Let $(p_i)_{i \in \mathbb{N}}$ be a probability vector with $\sup\{p_i : i \in \mathbb{N}\} = e^{-R}$. Hence

$$\begin{aligned} e^{-(r+1)R} &= \left(\sup\{p_i : i \in \mathbb{N}\}\right)^{r+1} \\ &\leq \left(\left(\sum_{i=1}^{\infty} p_i^{r+1}\right)^{1/(r+1)}\right)^{r+1} = \sum_{i=1}^{\infty} p_i^{r+1}. \end{aligned} \tag{29}$$

On the other hand let $N > e^R$ and define $p_{1,N} = e^{-R}$ resp.

$$p_{2,N} = \ldots = p_{N,N} = \frac{1 - e^{-R}}{N - 1}$$

and $p_{k,N} = 0$ for every $k \geq N + 1$. Obviously, $\sup\{p_{i,N} : i \in \mathbb{N}\} = e^{-R}$ and $\sum_{i=1}^{\infty} p_{i,N} = 1$. Moreover

$$\sum_{i=1}^{\infty} p_{i,N}^{r+1} = e^{-(r+1)R} + (N-1)\left(\frac{1 - e^{-R}}{N - 1}\right)^{r+1} \overset{N \to \infty}{\longrightarrow} e^{-(r+1)R}. \tag{30}$$

Inequality (29) and (30) are yielding (23). Similar to step 2 we obtain that the right hand side of (23) will not be attained by any $(p_1, p_2, \ldots) \in A(\infty, R)$. □

ACKNOWLEDGEMENT

This research was granted by the German Research Foundation (DFG). Moreover the author wants to thank the referees for communicating the references [14] and [17] to him and their suggestions leading to an improved version of this paper.

(Received February 26, 2009)

REFERENCES

[1] J. Aczél and Z. Daróczy: On Measures of Information and Their Characterizations. (Mathematics in Science and Engineering Vol. 115.) Academic Press, London 1975.

[2] M. S. Bazaraa, H. D. Sherali, and C. M. Shetty: Nonlinear Programming, Theory and Algorithms. Wiley, New York 1993.

[3] C. Beck and F. Schlögl: Thermodynamics of Chaotic Systems. Cambridge University Press, Cambridge 1993.

[4] M. Behara: Additive and Nonadditive Measures of Entropy. Wiley, New Delhi 1990.

[5] J. C. Fort and G. Pagès: Asymptotics of optimal quantizers for some scalar distributions. J. Comput. Appl. Math. *146* (2002), 253–275.

[6] A. Gersho and R. M. Gray: Vector Quantization and Signal Compression. Kluwer, Boston 1992.

[7] S. Graf and H. Luschgy: The quantization of the Cantor distribution. Math. Nachr. *183* (1997), 113–133.

[8] S. Graf and H. Luschgy: Foundations of Quantization for Probability Distributions. (Lecture Notes in Computer Science 1730.) Springer, Berlin 2000.

[9] R. M. Gray and D. Neuhoff: Quantization. IEEE Trans. Inform. Theory *44* (1998), 2325–2383.

[10] M. Grendar: Entropy and effective support size. Entropy *8* (2006), 169–174.

[11] A. György and T. Linder: Optimal entropy-constrained scalar quantization of a uniform source. IEEE Trans. Inform. Theory *46* (2000), 2704–2711.

[12] A. György and T. Linder: On the structure of optimal entropy-constrained scalar quantizers. IEEE Trans. Inform. Theory *48* (2002), 416–427.

[13] G. H. Hardy, J. E. Littlewood and G. Polya: Inequalities. Second edition. Cambridge University Press, Cambridge 1959.

[14] P. Harremoës and F. Topsøe: Inequalities between entropy and index of coincidence derived from information diagrams. IEEE Trans. Inform. Theory *47* (2001), 2944–2960.

[15] M. Kesseböhmer and S. Zhu: Stability of quantization dimension and quantization for homogeneous Cantor measure. Math. Nachr. *280* (2007), 866–881.

[16] W. Kreitmeier: Optimal quantization for dyadic homogeneous Cantor distributions. Math. Nachr. *281* (2008), 1307–1327.

[17] F. Liese, D. Morales, and I. Vajda: Asymptotically sufficient partitions and quantizations. IEEE Trans. Inform. Theory *52* (2006), 5599–5606.

[18] K. Pötzelberger and H. Strasser: Clustering and quantization by MSP-partitions. Statist. Decisions *19* (2001), 331–371.

Wolfgang Kreitmeier, Department of Informatics and Mathematics, University of Passau, Innstraße 33, 94032 Passau. Germany.

e-mail: wolfgang.kreitmeier@uni-passau.de