ERROR BOUNDS FOR HIGH-RESOLUTION QUANTIZATION WITH RÉNYI-α-ENTROPY CONSTRAINTS

W. KREITMEIER *

Department of Informatics and Mathematics,
University of Passau, 94032 Passau, Germany
e-mail: wolfgang.kreitmeier@uni-passau.de

Abstract

We consider the problem of optimal quantization with norm exponent $r > 0$ for Borel probability measures on $\mathbb{R}^d$ under constrained Rényi-α-entropy of the quantizers. If the bound on the entropy becomes large, then sharp asymptotics for the optimal quantization error are well-known in the special cases $\alpha = 0$ (memory-constrained quantization) and $\alpha = 1$ (Shannon-entropy-constrained quantization). In this paper we determine sharp asymptotics for the optimal quantization error under large entropy bound with entropy parameter $\alpha \in [1 + r/d, \infty]$. For $\alpha \in [0, 1 + r/d[$ we specify the asymptotical order of the optimal quantization error under large entropy bound. The optimal quantization error is decreasing exponentially fast with the entropy bound and the exact rate is determined for all $\alpha \in [0, \infty]$.

1 Introduction and basic notation

The quantization of probability distributions is mainly motivated from electrical engineering in the context of signal processing and data compression. A good survey about the historical development of the theory has been provided by Gray and Neuhoff [17]. The reader is also referred to the book of Gersho and Gray [12] for more applied aspects. Optimal quantization is the task of finding a

*This research was supported through a grant from the German Research Foundation (DFG).
Key words and phrases: Vector quantization, high-resolution quantization, Rényi-α-entropy, approximation of probability measures.
2000 Mathematics Subject Classification: primary 62E17, 94A17, secondary 60Exx, 68P30, 94A29.

best approximation of a given probability measure by another probability measure 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. Every point of a codecell will be mapped by the quantizer to a codepoint which is unique for each codecell. The set of all codepoints is called codebook. The mathematical aspects of quantization in finite dimension with restricted memory size have been investigated by Graf, Luschgy et.al. [8, 13, 14], Gruber [18], Dereich et.al. [10] and Fort, Pagès et.al. [9, 11]. A thorough mathematical treatment of (Shannon-)entropy-constrained quantization also emerged in the last few years and has been carried out by Gray, György, Linder and Li (see e.g. [16, 19, 20, 21] and the references therein). Sullivan [28] developed an algorithm for designing entropy-constrained scalar quantizers for the exponential and Laplace distribution. The question if optimal entropy-constrained quantizers induce a finite or infinite number of codecells has been investigated by György, Linder, Chou and Betts [22]. Recently, quantization has also been studied with combined entropy and memory size constraints (cf. [15]). Apart from studying high-resolution asymptotics also the asymptotic behavior of the optimal quantization error under small bounds on the complexity constraint has been investigated by several authors (cf. [24, 25, 26]).

In this paper we study the problem of quantization with arbitrary norm and norm exponent $r > 0$ by constraining the Rényi-α-entropy of the approximating probability. The known approach of restricted memory size is represented by the special case $\alpha = 0$. Shannon-entropy-constrained quantization is covered by $\alpha = 1$. Quantization with entropy parameter $\alpha = \infty$ can be interpreted as mass-constrained quantization, i.e. the maximum appearing codecell probability has to be larger or equal than a given bound. Because quantization is often linked with a subsequent lossless coding process of the codebooks, constraints on the output entropy of the quantizers are naturally arising out of the restricted channel capacity. If one restricts the average length of the code this is equivalent to bounded Shannon-entropy of the quantizers (cf. [16]). Campbell (cf. [6]) introduced a generalized measure for code length and has shown that this measure is related to Rényi's entropy. This strongly encourages to work with Rényi-α-entropy as a generalized measure of complexity for the quantizers.

For a large class of distributions, which are absolutely continuous with respect to the $d-$dimensional Lebesgue measure and entropy parameter $\alpha \in [1 + r/d, \infty]$, we derive as a main result (cf. Theorem 4.3) the exact asymptotic behavior of the optimal quantization error if the bound on the Rényi-α-entropy tends to infinity. Together with the known results for $\alpha \in \{0, 1\}$ we also provide upper and lower asymptotical bounds on the optimal quantization error for entropy parameter $\alpha \in [0, 1 + r/d[$, enabling us to determine the asymptotical order of the optimal quantization error in this region of α's if the entropy bound tends to infinity (cf. Theorem 5.2). The optimal quantization error is decreasing exponentially fast with the entropy bound and the exact rate will be determined in Corollary 5.3 for all $\alpha \in [0, \infty]$.

The paper is organized as follows. The rest of this first section contains basic notation and the setup of the optimal Rényi-α-entropy-constrained quantization problem. In section two we will prove that in the case of mass-constrained quantization ($\alpha = \infty$) the optimal quantization error can be computed in terms of a minimal moment on a ball resp. is dominated by a moment on a ball in case of $\alpha \in \,]1, \infty[$ (cf. Proposition 2.1). Using this error upper bounds and by a lower bound for moments on balls, which will be proved in section three, we are able to determine in section four the exact asymptotical behavior of the optimal quantization error for $\alpha \in [1 + r/d, \infty]$ and large entropy bound (cf. Theorem 4.3)). The last section contains our result about the asymptotical error bounds and the asymptotical order for entropy parameter $\alpha \in [0, 1 + r/d[$ (cf. Theorem 5.2). In both results (Theorem 4.3, 5.2) we are restricted to probability measures which are absolutely continuous with respect to the $d-$dimensional Lebesgue measure, having finite $r-$th moment and whose Lebesgue density has finite essential supremum.

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

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

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*

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

if card *denotes cardinality. Using de l'Hospital it is easy to see, that*

$$\lim_{\substack{\alpha \to 1 \\ \alpha \neq 1}} \hat{H}^{\alpha}(\cdot) = \hat{H}^1(\cdot)$$

(cf. [1, Remark 5.2.34]). Moreover we have $\lim_{\alpha\to\infty} \hat{H}^{\alpha}(\cdot) = \hat{H}^{\infty}(\cdot)$. *The Rényi-$\alpha$-entropy is also monotonically decreasing in α, i.e. for every $0 \le \gamma \le \beta \le \infty$ we have (cf. [2], p. 53)*

$$\hat{H}^{\beta}(\cdot) \le \hat{H}^{\gamma}(\cdot).$$

Now let $d \in \mathbb{N}$ and denote by $\mathbb{R}$ the set of all real numbers. Let μ be a Borel probability measure on $\mathbb{R}^d$. For any mapping $f : \mathbb{R}^d \to \mathbb{R}^d$ denote by $f(\mathbb{R}^d) = \{f(x) : x \in \mathbb{R}^d\}$ the image of $\mathbb{R}^d$ under f. A mapping $f \in \mathcal{F}_d$ is called a **quantizer** and the image $f(\mathbb{R}^d)$ is called codebook consisting of codepoints. We assume throughout the whole paper that the codepoints are distinct. 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 **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 $\{z_1, z_2, ..\}$ of $f(\mathbb{R}^d)$ we define

$$H_\mu^\alpha(f) = \hat{H}^\alpha(\mu \circ f^{-1}(z_1), \mu \circ f^{-1}(z_2), ...)$$

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 strictly 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

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

For any $R \geq 0$ we denote by

$$D_{\mu,\rho}^\alpha(R) = \inf\{D_{\mu,\rho}(f) : f \in \mathcal{F}_d, H_\mu^\alpha(f) \leq R\} \tag{1}$$

the optimal quantization error for μ of order r under Rényi-α-entropy bound R. The exact determination of the optimal quantization error is rather hard in general. But for large entropy bound R and $\alpha \in \{0,1\}$, the asymptotical behavior of the optimal quantization error is well-known for a large class of probability distributions (cf. [5], [13, Theorem 6.2], [16], [29]). For $r > 0$ and $\rho : x \to x^r$ we will provide asymptotical error bounds for $\alpha \in [0, 1 + r/d[$ and sharp asymptotics for $\alpha \in [1 + r/d, \infty]$.

2 Error properties in optimal Rényi-α-entropy-constrained quantization

For any $x \in \mathbb{R}^d$ and $l > 0$ we denote by

$$B(x,l) = \{z \in \mathbb{R}^d : \|z - x\| \leq l\}$$

the closed ball with midpoint x and radius l. Moreover we say that μ vanishes on spheres, if $\mu(\{z \in \mathbb{R}^d : \|z - x\| = l\}) = 0$ for every $x \in \mathbb{R}^d$ and $l > 0$.

From now on we specialize for the rest of this paper the mapping ρ to $\rho_r(x) = x^r$ with norm exponent $r > 0$. We write

$$D_{\mu,r}^\alpha(\cdot) = D_{\mu,\rho_r}^\alpha(\cdot) \text{ resp. } D_{\mu,r}(\cdot) = D_{\mu,\rho_r}(\cdot).$$

To prove the following result we use arguments presented by Graf and Luschgy (cf. [13, Lemma 2.8, Lemma 6.1]).

Proposition 2.1. *Let $R > 0$. Assume that μ has a finite $r-$th moment and vanishes on spheres. If $\alpha \in]1,\infty[$, then*

$$\begin{aligned} & D_{\mu,r}^\alpha(R) \\ \leq \quad & \inf\left\{\int_{B(a,s)} \|x - a\|^r d\mu(x) : a \in \mathbb{R}^d, s > 0, \mu(B(a,s)) = e^{-\frac{\alpha-1}{\alpha}R}\right\}. \end{aligned} \tag{2}$$

Moreover,

$$D^{\infty}_{\mu,r}(R) = \inf\left\{\int_{B(a,s)} \|x-a\|^r d\mu(x) : a \in \mathbb{R}^d, s > 0, \mu(B(a,s)) = e^{-R}\right\}.$$

Proof. The idea of the proof consists of designing a quantizer which is composed of a ball as dominating codecell and remaining cells which have an insignificant contribution to the quantization error. The construction is more apparent for $\alpha = \infty$, where clearly only the largest cell matters, as the remaining part of the space can be quantized in an arbitrary fine manner (with essentially zero distortion) without increasing the entropy.

1. $\alpha \in]1,\infty[$.

Let $\varepsilon \in]0,1[$ and $p = e^{-\frac{\alpha-1}{\alpha}R}$. Let $a \in \mathbb{R}^d$. Because μ vanishes on spheres, the mapping

$$]0,\infty[\ \ni s \to \mu(B(a,s)) \in [0,1]$$

is continuous. Therefore the intermediate value theorem yields the existence of an $s_a > 0$ with $\mu(B(a,s_a)) = p$. Let $\varepsilon > 0$. With $\delta = \varepsilon^{1/r}$ we have $t^r \leq \varepsilon$ for every $t \in [0,\delta]$. Let $(x_n)_{n\in\mathbb{N}}$ be dense in $\mathbb{R}^d$. Then $(B(x_n,\delta))_{n\in\mathbb{N}}$ is an open cover of $\mathbb{R}^d$. Hence a Borel-measurable partition $(A_n)_{n\in\mathbb{N}}$ of $\mathbb{R}^d\backslash B(a,s_a)$ exists, such that $A_n \subset B(x_n,\delta)$ for every $n \in \mathbb{N}$. Now we define the mapping $f : \mathbb{R}^d \to \mathbb{R}^d$ by

$$f(x) = \begin{cases} a, \text{ if } x \in B(a,s_a) \\ x_n, \text{ if } x \in A_n\ . \end{cases}$$

Due to $\alpha > 1$ we obtain

$$\begin{aligned} H^{\alpha}_{\mu}(f) &= \frac{1}{1-\alpha}\log\left(\mu(B(a,s_a))^{\alpha} + \sum_{n=1}^{\infty}\mu(A_n)^{\alpha}\right) \\ &\leq \frac{1}{1-\alpha}\log(\mu(B(a,s_a))^{\alpha}) = \frac{1}{1-\alpha}\log(e^{(1-\alpha)R}) = R. \end{aligned}$$

As a consequence we get

$$\begin{aligned} D^{\alpha}_{\mu,r}(R) \leq D_{\mu,r}(f) &= \int_{B(a,s_a)} \|x-a\|^r d\mu(x) + \sum_{n=1}^{\infty}\int_{A_n} \|x-x_n\|^r d\mu(x) \\ &\leq \int_{B(a,s_a)} \|x-a\|^r d\mu(x) + \sum_{n=1}^{\infty}\mu(A_n)\varepsilon \\ &\leq \int_{B(a,s_a)} \|x-a\|^r d\mu(x) + \varepsilon. \end{aligned}$$

Because $\varepsilon > 0$ was chosen arbitrarily the assertion is proved for $\alpha \in]1,\infty[$.

2. $\alpha = \infty$.

Let

$$D(R) = \inf\{\int_A \|x-a\|^r d\mu(x) : a \in \mathbb{R}^d, A \text{ measurable }, \mu(A) \geq e^{-R}\}. \tag{3}$$

2.1. $D^{\infty}_{\mu,r}(R) \geq D(R)$.

Let $f \in \mathcal{F}_d$ with $H^{\infty}_{\mu}(f) \leq R$. Then an $a \in f(\mathbb{R}^d)$ exists with $\mu(f^{-1}(a)) \geq e^{-R}$. Let $A = f^{-1}(a)$. We obtain

$$\begin{aligned}
\int \|x - f(x)\|^r d\mu(x) &= \sum_{b \in f(\mathbb{R}^d)} \int_{f^{-1}(b)} \|x - b\|^r d\mu(x) \\
&\geq \int_{f^{-1}(a)} \|x - a\|^r d\mu(x) \\
&= \int_A \|x - a\|^r d\mu(x) \geq D(R),
\end{aligned}$$

which yields $D^{\infty}_{\mu,r}(R) \geq D(R)$.

2.2. $D^{\infty}_{\mu,r}(R) \leq D(R)$.

This follows by the same argumentation as in step 1 if we replace $B(a, s_a)$ by a measurable set A with $\mu(A) \geq e^{-R}$.

From step 2.1 and 2.2 we deduce $D^{\infty}_{\mu,r}(R) = D(R)$. Obviously we can assume that $\mu(A) \in]0, 1[$ for the set A in (3). But then the assertion is an immediate consequence of [13, Lemma 2.8] if we remark that the proof of Lemma 2.8 in [13] works also for $r \in]0, 1[$. □

Remark 2.2. *The proof of Proposition 2.1 shows, that an optimal quantizer for $\alpha = \infty$ does not exist, i.e. for any quantizer $f \in \mathcal{F}_d$ with $H^{\infty}_{\mu}(f) \leq R$ we obtain $D^{\infty}_{\mu,r}(R) < D_{\mu,r}(f)$.*

3 A lower bound for moments on balls

For any set $A \subset \mathbb{R}^d$ we denote by 1_A the characteristic function of A. We denote by λ^d the Lebesgue measure on $\mathbb{R}^d$. For a measurable mapping $h : \mathbb{R}^d \to \mathbb{R}$ and a measurable nonempty set $A \subset \mathbb{R}^d$ we denote by

$$\operatorname{ess\,sup}_A h = \inf\{b \in \mathbb{R}^d : \lambda^d(\{x \in A : h(x) > b\}) = 0\} \tag{4}$$

the essential supremum of h on A. We write $\operatorname{ess\,sup} h = \operatorname{ess\,sup}_{\mathbb{R}^d} h$.

Lemma 3.1. *Let $s > 0$ and $r > 0$. Then*

$$\int_{B(0,s)} \|x\|^r d\lambda^d(x) = s^{d+r} \lambda^d(B(0,1)) \frac{d}{d+r}.$$

Proof. We compute

$$\begin{aligned}
\int_{B(0,s)} \|x\|^r d\lambda^d(x) &= \int 1_{B(0,s)}(x)\|x\|^r d\lambda^d(x) \\
&= \int_0^\infty \lambda^d(\{x : 1_{B(0,s)}(x)\|x\|^r > t\})dt \\
&= \int_0^{s^r} \left(\lambda^d(B(0,s)) - \lambda^d(B(0,t^{1/r}))\right) dt \\
&= \lambda^d(B(0,1)) \int_0^{s^r} \left(s^d - t^{d/r}\right) dt \\
&= s^{d+r}\lambda^d(B(0,1))\frac{d}{d+r},
\end{aligned}$$

which yields the assertion. □

Proposition 3.2. *Let μ be absolutely continuous w.r.t. λ^d with density h. Assume that* $\operatorname{ess\,sup} h < \infty$. *Let $a \in \mathbb{R}^d$ and $s > 0$. Then*

$$\int_{B(a,s)} \|x-a\|^r d\mu(x) \geq \frac{d}{d+r} \frac{\mu(B(a,s))^{1+r/d}}{(\lambda^d(B(0,1))\operatorname{ess\,sup} h)^{r/d}}.$$

Proof. If $\mu(B(a,s)) = 0$, then the assertion is obvious. Let us assume that $\mu(B(a,s)) > 0$. Clearly, $\operatorname{ess\,sup} h > 0$. Let

$$l = \left(\frac{\mu(B(a,s))}{\lambda^d(B(0,1))\operatorname{ess\,sup} h}\right)^{1/d}.$$

Obviously,

$$0 < l \leq \left(\frac{\lambda^d(B(a,s))\operatorname{ess\,sup} h}{\lambda^d(B(0,1))\operatorname{ess\,sup} h}\right)^{1/d} \leq s.$$

We deduce

$$\begin{aligned}
&\int_{B(a,s)} \|x-a\|^r d\mu(x) \\
=\ &\int_{B(a,l)} \|x-a\|^r d\mu(x) + \int_{B(a,s)\setminus B(a,l)} \|x-a\|^r d\mu(x) \\
\geq\ &\int_{B(a,l)} \|x-a\|^r d\mu(x) + l^r \mu(B(a,s)\setminus B(a,l)). \qquad (5)
\end{aligned}$$

From the definition of l we obtain

$$\lambda^d(B(a,l))\operatorname{ess\,sup} h = \mu(B(a,s)).$$

This implies

$$\begin{aligned}
\mu(B(a,s)\setminus B(a,l)) &= \int_{B(a,s)} h d\lambda^d - \int_{B(a,l)} h d\lambda^d \\
&= \int_{B(a,l)} (\operatorname{ess\,sup} h - h) d\lambda^d. \qquad (6)
\end{aligned}$$

Combining (5) and (6) we get

$$\begin{aligned}
&\int_{B(a,s)} \|x-a\|^r d\mu(x) \\
\geq\ & \int_{B(a,l)} \left(\|x-a\|^r h(x) + l^r(\operatorname{ess\,sup} h - h)\right) d\lambda^d \\
=\ & \int_{B(a,l)} \left(\|x-a\|^r \operatorname{ess\,sup} h + (l^r - \|x-a\|^r)(\operatorname{ess\,sup} h - h)\right) d\lambda^d \\
\geq\ & \operatorname{ess\,sup} h \int_{B(0,l)} \|x\|^r d\lambda^d(x).
\end{aligned}$$

Lemma 3.1 and the definition of l yields

$$\begin{aligned}
&\int_{B(a,s)} \|x-a\|^r d\mu(x) \\
\geq\ & l^{d+r}\lambda^d(B(0,1))\frac{d}{d+r}\operatorname{ess\,sup} h \\
=\ & \left(\frac{\mu(B(a,s))}{\lambda^d(B(0,1))\operatorname{ess\,sup} h}\right)^{1+r/d} \lambda^d(B(0,1))\frac{d}{d+r}\operatorname{ess\,sup} h \\
=\ & \frac{d}{d+r}\frac{\mu(B(a,s))^{1+r/d}}{(\lambda^d(B(0,1))\operatorname{ess\,sup} h)^{r/d}},
\end{aligned}$$

which proves the assertion. □

4 Sharp asymptotics for the optimal quantization error and $\alpha \geq 1 + r/d$

In this section we will first provide a lower bound for the optimal quantization error and $\alpha \geq 1 + r/d$. To this end we use the lower bound for moments on balls of the previous section. On the other hand we know that the optimal quantization error is dominated by a ball moment (cf. Proposition 2.1). This enables us to prove an upper bound for the error. Together with the lower bound we derive for entropy parameter $\alpha \geq 1 + r/d$ sharp asymptotics for the optimal quantization error if the bound on the entropy tends to infinity.

Proposition 4.1. *Assume that μ is absolutely continuous w.r.t. λ^d with density h. Moreover we assume that $\operatorname{ess\,sup} h < \infty$ and that the $r-$th moment of μ is finite. Let $R > 0$. Then*

$$e^{R(1+r/d)} D^{\infty}_{\mu,r}(R) \geq \frac{d}{d+r}\left(\lambda^d(B(0,1))\right)^{-r/d} (\operatorname{ess\,sup} h)^{-r/d}.$$

If $\alpha \in [1 + r/d, \infty[$, then

$$e^{\frac{\alpha-1}{\alpha}R(1+r/d)} D^{\alpha}_{\mu,r}(R) \geq \frac{d}{d+r}\left(\lambda^d(B(0,1))\right)^{-r/d} (\operatorname{ess\,sup} h)^{-r/d}. \tag{7}$$

Proof. 1. $\alpha = \infty$.
Let $\varepsilon > 0$. Applying Proposition 2.1 we can find an $a \in \mathbb{R}^d$ and $s > 0$ such that

$$\varepsilon + D_{\mu,r}^{\infty}(R) \geq \int_{B(a,s)} \|x-a\|^r d\mu(x)$$

and $\mu(B(a,s)) = e^{-R}$. From Proposition 3.2 we obtain

$$\begin{aligned} & \varepsilon + D_{\mu,r}^{\infty}(R) \\ \geq \quad & \operatorname{ess\,sup} h \frac{d}{d+r} \lambda^d(B(0,1)) \left(\frac{e^{-R}}{\lambda^d(B(0,1)) \operatorname{ess\,sup} h} \right)^{1+r/d} \\ = \quad & \frac{d}{d+r} \left(\lambda^d(B(0,1)) \right)^{-r/d} (\operatorname{ess\,sup} h)^{-r/d} e^{-R(1+r/d)}. \end{aligned}$$

Letting ε tend to zero yields the assertion.
2. $\alpha \in [1 + r/d, \infty[$.
Let $f \in \mathcal{F}_d$ with $H_\mu^\alpha(f) \leq R$. For any $a \in f(\mathbb{R}^d)$ with $\mu(f^{-1}(a)) > 0$ we can find an $s_a > 0$ such that $\mu(f^{-1}(a)) = \mu(B(a, s_a))$. Using [13, Lemma 2.8] and Proposition 3.2 we deduce

$$\begin{aligned} & D_{\mu,r}(f) \\ = \quad & \sum_{a \in f(\mathbb{R}^d)} \int_{f^{-1}(a)} \|x-a\|^r d\mu(x) \\ \geq \quad & \sum_{a \in f(\mathbb{R}^d)} \int_{B(a,s_a)} \|x-a\|^r d\mu(x) \\ \geq \quad & \left(\lambda^d(B(0,1)) \operatorname{ess\,sup} h \right)^{-r/d} \frac{d}{d+r} \sum_{a \in f(\mathbb{R}^d)} \left(\mu(f^{-1}(a)) \right)^{1+r/d}. \end{aligned} \tag{8}$$

Because $\alpha \geq 1 + r/d$ we obtain from Jensen's inequality for series (see e.g. [3, p. 18]) that

$$\left(\sum_{a \in f(\mathbb{R}^d)} \left(\mu(f^{-1}(a)) \right)^{1+r/d} \right)^{1/(1+r/d)} \geq \left(\sum_{a \in f(\mathbb{R}^d)} \left(\mu(f^{-1}(a)) \right)^{\alpha} \right)^{\frac{1}{\alpha}}.$$

Together with $H_\mu^\alpha(f) \leq R$ we obtain

$$\sum_{a \in f(\mathbb{R}^d)} \left(\mu(f^{-1}(a)) \right)^{1+r/d} \geq \left(\sum_{a \in f(\mathbb{R}^d)} \left(\mu(f^{-1}(a)) \right)^{\alpha} \right)^{\frac{1+r/d}{\alpha}} \geq e^{-\frac{\alpha-1}{\alpha}(1+r/d)R}.$$

Hence we deduce with (8) that

$$D_{\mu,r}(f) \geq \left(\lambda^d(B(0,1)) \operatorname{ess\,sup} h \right)^{-r/d} \frac{d}{d+r} e^{-\frac{\alpha-1}{\alpha}(1+r/d)R}.$$

Taking the infimum over all $f \in \mathcal{F}_d$ with $H_\mu^\alpha(f) \leq R$ yields the assertion. □

Remark 4.2. *In view of the results for the one-dimensional uniform distribution (cf. [23]) it is reasonable to conjecture and remains an open question if the lower bound (7) can be sharpened to*

$$D^{\alpha}_{\mu,r}(R) \geq \inf\left\{\int_{B(a,s)} \|x-a\|^r d\mu(x) : a \in \mathbb{R}^d, s > 0, \mu(B(a,s)) = e^{-\frac{\alpha-1}{\alpha}R}\right\},$$

i.e. that inequality (2) turns into an equation.

Theorem 4.3. *Assume that μ is absolutely continuous w.r.t. λ^d with density h. Moreover we assume that* $\operatorname{ess\,sup} h < \infty$ *and that μ has finite $r-th$ moment. Then*

$$\lim_{R\to\infty} e^{R(1+r/d)} D^{\infty}_{\mu,r}(R) = \frac{d}{d+r}\left(\lambda^d(B(0,1))\right)^{-r/d} (\operatorname{ess\,sup} h)^{-r/d}. \tag{9}$$

If $\alpha \in [1+r/d, \infty[$, then

$$\lim_{R\to\infty} e^{\frac{\alpha-1}{\alpha}R(1+r/d)} D^{\alpha}_{\mu,r}(R) = \frac{d}{d+r}\left(\lambda^d(B(0,1))\right)^{-r/d} (\operatorname{ess\,sup} h)^{-r/d}. \tag{10}$$

Proof.

1a. $\liminf_{R\to\infty} e^{R(1+r/d)} D^{\infty}_{\mu,r}(R) \geq \frac{d}{d+r}\left(\lambda^d(B(0,1))\right)^{-r/d} (\operatorname{ess\,sup} h)^{-r/d}$.

This follows immediately from Proposition 4.1.

1b. $\limsup_{R\to\infty} e^{R(1+r/d)} D^{\infty}_{\mu,r}(R) \leq \frac{d}{d+r}\left(\lambda^d(B(0,1))\right)^{-r/d} (\operatorname{ess\,sup} h)^{-r/d}$.

Let $0 < b < \operatorname{ess\,sup} h$. From definition (4) we obtain $\lambda^d(\{h > b\}) > 0$. Lebesgue's density theorem (cf. [7, Corollary 6.2.2]) implies the existence of an $a_0 \in \mathbb{R}^d$ with

$$\lim_{\substack{s\to 0\\ s>0}} \frac{\lambda^d(\{h>b\}\cap B(a_0,s))}{\lambda^d(B(a_0,s))} = 1.$$

Let $\varepsilon \in]0,1[$. Then there exists an $s(\varepsilon) > 0$ such that

$$\lambda^d(\{h>b\}\cap B(a_0,s)) > (1-\varepsilon)\lambda^d(B(a_0,s)) \tag{11}$$

for every $s \in]0, s(\varepsilon)]$. From (11) we deduce for every $s \in]0, s(\varepsilon)]$ that

$$\begin{aligned} \mu(B(a_0,s)) &\geq \int_{B(a_0,s)\cap\{h>b\}} h d\lambda^d \\ &\geq b\lambda^d(B(a_0,s)\cap\{h>b\}) \geq b(1-\varepsilon)\lambda^d(B(a_0,s)) > 0. \end{aligned} \tag{12}$$

Let $C \in]0, \mu(B(a_0, s(\varepsilon)))[$ and

$$s_C = \sup\{s > 0 : \mu(B(a_0,s)) \leq C\}.$$

Because μ is absolutely continuous w.r.t. λ^d we have $\mu(B(a_0,s_C)) = C$ and $s_C \leq s(\varepsilon)$. From Proposition 2.1 we obtain

$$\begin{aligned} D^{\infty}_{\mu,r}(-\log(C)) &= \inf\{\int_{B(a,s)} \|x-a\|^r d\mu(x) : a \in \mathbb{R}^d, \mu(B(a,s)) = C\} \\ &\leq \int_{B(a_0,s_C)} \|x-a_0\|^r d\mu(x). \end{aligned} \tag{13}$$

Moreover

$$\int_{B(a_0,s_C)} \|x-a_0\|^r d\mu(x) \le \operatorname{ess\,sup} h \int_{B(a_0,s_C)} \|x-a_0\|^r d\lambda^d(x).$$

Using Lemma 3.1 we obtain

$$\begin{aligned}\int_{B(a_0,s_C)} \|x-a_0\|^r d\mu(x) &\le \operatorname{ess\,sup} h \int_{B(0,s_C)} \|x\|^r d\lambda^d(x) \\ &= \operatorname{ess\,sup} h \frac{d}{d+r} \lambda^d(B(0,1)) s_C^{r+d}. \end{aligned} \tag{14}$$

Relation (12) yields

$$\lambda^d(B(a_0,s_C)) \le \frac{1}{(1-\varepsilon)b} \mu(B(a_0,s_C)).$$

Thus we get

$$s_C^d \le \frac{1}{(1-\varepsilon)b} \frac{\mu(B(a_0,s_C))}{\lambda^d(B(0,1))}$$

which implies

$$s_C \le \left(\frac{1}{(1-\varepsilon)b} \frac{1}{\lambda^d(B(0,1))} \right)^{1/d} C^{1/d}. \tag{15}$$

Combining (15) and (14) we obtain

$$\begin{aligned}&\int_{B(a_0,s_C)} \|x-a_0\|^r d\mu(x) \\ \le\ & \left(\operatorname{ess\,sup} h \frac{d}{d+r} \lambda^d(B(0,1)) \right) s_C^{r+d} \\ \le\ & \left(\operatorname{ess\,sup} h \frac{d}{d+r} \lambda^d(B(0,1)) \right) \left(\frac{C}{(1-\varepsilon)b\lambda^d(B(0,1))} \right)^{\frac{r+d}{d}}. \end{aligned}$$

Now we substitute $C = e^{-R}$ with large enough $R > 0$. From (13) we get

$$\begin{aligned}&D^\infty_{\mu,r}(R) \\ \le\ & \left(\operatorname{ess\,sup} h \frac{d}{d+r} \lambda^d(B(0,1)) \right) \left(\frac{e^{-R}}{(1-\varepsilon)b\lambda^d(B(0,1))} \right)^{\frac{r+d}{d}}. \end{aligned}$$

Because $b \in]0, \operatorname{ess\,sup} h[$ and $\varepsilon \in]0,1[$ were arbitrary we finally obtain

$$\limsup_{R\to\infty} e^{R(1+r/d)} D^\infty_{\mu,r}(R) \le \frac{d}{d+r} \left(\lambda^d(B(0,1)) \right)^{-r/d} (\operatorname{ess\,sup} h)^{-r/d}.$$

The combination of 1a and 1b proves the first assertion.

2a. $\liminf_{R\to\infty} e^{\frac{\alpha-1}{\alpha} R(1+r/d)} D^\infty_{\mu,r}(R) \ge \frac{d}{d+r} \left(\lambda^d(B(0,1)) \right)^{-r/d} (\operatorname{ess\,sup} h)^{-r/d}$.

This follows immediately from Proposition 4.1.

2b. $\limsup_{R\to\infty} e^{\frac{\alpha-1}{\alpha}R(1+r/d)} D^{\infty}_{\mu,r}(R) \le \frac{d}{d+r}\left(\lambda^d(B(0,1))\right)^{-r/d} (\operatorname{ess\,sup} h)^{-r/d}$.

This follows by exactly the same arguments as in step 1b if we make the substitution $C = e^{-\frac{\alpha-1}{\alpha}R}$. The combination of 2a and 2b yields the remaining part of the assertion. □

Remark 4.4. *To get sharp asymptotics for $\alpha = 0$ (cf. [13, Theorem 6.2]) it suffices that μ has only a non-vanishing part which is absolutely continuous with respect to λ^d. It is an open question if Theorem 4.3 remains also true under this weaker condition. Moreover, for $\alpha = 0$ and probability distributions μ that are continuous and singular to λ^d, the high-rate asymptotics of the optimal quantization errors is generally different from the absolutely continuous case (cf. [13, p.155]). In order to generalize these results to $\alpha > 0$ it has to be clarified, if the so-called quantization dimension ([13, p.155], [30])) exists for such distributions and $\alpha > 0$, i.e. if*

$$-rR/\log(D^{\alpha}_{\mu,r}(R))$$

converges to a positive and finite value for $R \to \infty$.

Remark 4.5. *If the density h of μ has infinite essential supremum it remains an open question if the right hand side of (9) resp. (10) equals zero.*

5 Upper and lower asymptotical bounds for the optimal quantization error and $\alpha \le 1 + r/d$

Using our results for $\alpha \ge 1 + r/d$ and the well-known sharp error asymptotics for $\alpha = 0$ (see e.g. [13, Theorem 6.2]) we will determine in this last section the asymptotical order of the optimal quantization error for $\alpha \le 1 + r/d$.

Definition 5.1. *Let $\mu = 1_{[0,1]^d}\lambda^d$. We call*

$$Q_r([0,1]^d) = \inf\{n^{r/d} D^0_{\mu,r}(\log(n)) : n \in \mathbb{N}\}$$

the $r-$th quantization coefficient of $[0,1]^d$ (cf. [13, p. 81]). According to [13, Theorem 6.2] the infimum will be achieved as $n \to \infty$.

Theorem 5.2. *Let μ be a Borel probability measure on $\mathbb{R}^d$ with finite $(r+\delta)-$th moment for some $\delta > 0$. Assume that μ is absolutely continuous w.r.t. λ^d with density h. Moreover we assume that $\operatorname{ess\,sup} h < \infty$. Then for any $\alpha \in [0, 1+r/d]$*

$$\begin{aligned}
\limsup_{R\to\infty} e^{R(r/d)} D^{\alpha}_{\mu,r}(R) &\le \lim_{R\to\infty} e^{R(r/d)} D^{0}_{\mu,r}(R) \\
&= Q_r([0,1]^d)\left(\int |h|^{d/(d+r)} d\lambda^d\right)^{1+r/d}
\end{aligned}$$

resp.

$$\begin{aligned}\liminf_{R\to\infty} e^{R(r/d)} D^{\alpha}_{\mu,r}(R) &\geq \lim_{R\to\infty} e^{R(r/d)} D^{1+r/d}_{\mu,r}(R)\\ &= \frac{d}{d+r}\left(\lambda^d(B(0,1))\right)^{-r/d}(\operatorname{ess\,sup} h)^{-r/d}.\end{aligned}$$

Proof. Let $R \geq 0$ and $f \in \mathcal{F}_d$ with $H^{\alpha}_{\mu}(f) \leq R$. Remark 1.1 yields

$$H^{1+r/d}_{\mu}(f) \leq H^{\alpha}_{\mu}(f) \leq H^{0}_{\mu}(f).$$

In view of Definition (1) we thus obtain

$$D^{1+r/d}_{\mu,r}(R) \leq D^{\alpha}_{\mu,r}(R) \leq D^{0}_{\mu,r}(R). \tag{16}$$

Now let $n \in \mathbb{N}$ with $R \in [\log(n), \log(n+1)[$. Clearly,

$$\begin{aligned}&\left(\frac{n}{n+1}\right)^{r/d} e^{\frac{r}{d}\log(n+1)} D^{0}_{\mu,r}(\log(n+1))\\ \leq\ & e^{\frac{r}{d}R} D^{0}_{\mu,r}(R)\\ \leq\ & \left(\frac{n+1}{n}\right)^{r/d} e^{\frac{r}{d}\log(n)} D^{0}_{\mu,r}(\log(n))\end{aligned}$$

From [13, Theorem 6.2] we deduce that the first and last quantity of the inequality from above tends to $Q_r([0,1]^d)\left(\int |h|^{d/(d+r)} d\lambda^d\right)^{1+r/d}$ as $R \to \infty$. Thus we obtain

$$\lim_{R\to\infty} e^{R(r/d)} D^{0}_{\mu,r}(R) = Q_r([0,1]^d)\left(\int |h|^{d/(d+r)} d\lambda^d\right)^{1+r/d}. \tag{17}$$

The combination of (17) and (16) proves the first part of the assertion. The second part follows from (16) and Theorem 4.3. □

As an immediate consequence of Theorem 5.2 and Theorem 4.3 we obtain the following result about the rate of the optimal quantization error.

Corollary 5.3. *Let μ be a Borel probability measure on $\mathbb{R}^d$ with finite $(r+\delta)-$th moment for some $\delta > 0$. Assume that μ is absolutely continuous w.r.t. λ^d with density h. Moreover we assume that* $\operatorname{ess\,sup} h < \infty$*. Let*

$$q = -\max\left\{\frac{\alpha-1}{\alpha}(1+r/d), r/d\right\}.$$

Then,

$$\lim_{R\to\infty} \frac{\log(D^{\alpha}_{\mu,r}(R))}{R} = q.$$

Remark 5.4. *If $\alpha < 1 + r/d$ the exponential rate $q = -r/d$ seems to be achievable by using a uniform quantizer, i.e. a quantizer whose codecells are cubes, which have all equal side length and whose edges are parallel to the coordinate axes. If $\alpha > 1 + r/d$ the situation changes. Here the main contribution to the quantization error comes from a ball as codecell, which is centered around the most likely region of the probability distribution. It needs also further research to investigate if both strategies are always achieving the optimal rate q in case of $\alpha = 1 + r/d$. Because the error rate can be expressed by the quantization dimension (cf. Remark 4.4), a positive answer to the last open question in Remark 4.4 would generalize Corollary 5.3 to distributions which are continuous and singular to λ^d.*

Remark 5.5. *For $\alpha = 1$ and a large class of probability distributions which are absolutely continuous to λ^d we have*

$$\lim_{R\to\infty} e^{R(r/d)} D^{\alpha}_{\mu,r}(R) = C(r,d) e^{-\frac{r}{d}\int h \log h d\lambda^d}$$

(cf. [16], [29]), where $C = C(r,d)$ is a positive constant that depends only on r and d but not on h.

Remark 5.6. *Generally we have $Q_r([0,1]^d) \geq \frac{d}{d+r}\left(\lambda^d(B(0,1))\right)^{-r/d}$, i.e. the quantization coefficient has a ball lower bound (cf. [13, Proposition 8.3]). It is well known, that this inequality is strict for low dimensions, e.g. we have*

$$Q_2([0,1]^2) = \frac{5}{18\sqrt{3}} > \frac{1}{2\pi} = \frac{2}{2+2}(\lambda^2(B(0,1)))^{-2/2}$$

(cf. Example 8.12 and Theorem 8.15 in [13]).

Remark 5.7. *Because the volume $\lambda^d(B(0,1))$ of the $d-$dimensional unit ball appears very often in this paper it is worth to remark that*

$$\lambda^d(B(0,1)) = \frac{(2\Gamma(1+1/p))^d}{\Gamma(1+d/p)}$$

for the l_p-norms and $p \in [1,\infty[$ (cf. [27, p. 11])

Remark 5.8. *It will need further research to determine sharp asymptotics for the optimal quantization error under large entropy bound if $\alpha \in]0, 1+r/d[\setminus\{1\}$.*

Acknowledgement. The author is indebted to the referee for pointing out some errors and simplifications in an earlier version of this paper as well as for the communication of reference [6].

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] C. Beck and F. Schlögl, *Thermodynamics of chaotic systems*, Cambridge University Press (Cambridge, 1993).

[3] E. F. Beckenbach and R. Bellman, *Inequalities. 4th print.*, Springer (Berlin, 1983).

[4] M. Behara, *Additive and nonadditive measures of entropy*, John Wiley (New Delhi, 1990).

[5] J. A. Bucklew and G. L. Wise, Multidimensional asymptotic quantization theory with r-th power distortion measures, *IEEE Trans. Inf. Theory*, **28** (1982), 239-247.

[6] L. L. Campbell, A coding theorem and Rényis entropy, *Information and Control*, **8** (1965), 423-429.

[7] D. L. Cohn, *Measure Theory*, Birkhäuser (Boston, 1980).

[8] S. Delattre, S. Graf, H. Luschgy and G. Pagès, Quantization of probability distributions under norm-based distortion measures, *Stat. Decis.*, **22** (2004), 261-282.

[9] S. Delattre, J. C. Fort and G. Pagès, Local distortion and μ-mass of the cells of one dimensional asymptotically optimal quantizers, *Commun. Stat., Theory Methods*, **33** (2004), 1087-1117.

[10] S. Dereich and C. Vormoor, The high resolution vector quantization problem with Orlicz norm distortion, preprint (2005).

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

[12] A. Gersho and R.M. Gray, *Vector Quantization and Signal Compression*, Kluwer (Boston, 1992).

[13] S. Graf and H. Luschgy, *Foundations of Quantization for Probability Distributions*, Lecture Notes 1730 Springer (2000).

[14] S. Graf and H. Luschgy, Quantization for probability measures with respect to the geometric mean error, *Math. Proc. Camb. Philos. Soc.*, **136** (2004), 687-717.

[15] R. M. Gray, T. Linder and J. T. Gill, Lagrangian vector quantization with combined entropy and codebook size constraints, *IEEE Trans. Inf. Theory*, **54** (2008), 2209-2242.

[16] R. M. Gray, T. Linder and J. Li, A Lagrangian formulation of Zador's entropy-constrained quantization theorem, *IEEE Trans. Inf. Theory*, **48** (2002), 695-707.

[17] R. M. Gray and D. Neuhoff, Quantization, *IEEE Trans. Inf. Theory*, **44** (1998), 2325-2383.

[18] P. M. Gruber, Optimum quantization and its applications, *Adv. Math.*, **186**(2) (2004), 456-497.

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

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

[21] A. György and T. Linder, Codecell convexity in optimal entropy-constrained vector quantization, *IEEE Trans. Inf. Theory*, **49** (2003), 1821-1828.

[22] A. György, T. Linder, P. A. Chou, and B. J. Betts, Do Optimal Entropy-Constrained Quantizers have a Finite or Infinite Number of Codewords ?, *IEEE Trans. Inf. Theory*, **49** (2003), 30313037.

[23] W. Kreitmeier, Optimal Quantization for the one-dimensional uniform distribution with Rényi-α-entropy constraints, preprint (2009).

[24] A. Magnani, A. Ghosh, and R. M. Gray, Optimal One-Bit Quantization, *Proc. Data Compression Conference*, (2005), 270278.

[25] D. Marco and D. L. Neuhoff, Low-Resolution Scalar Quantization for Gaussian Sources and Squared Error, *IEEE Trans. Inf. Theory*, **52** (2006), 1689-1697.

[26] D. Marco and D. L. Neuhoff, Low-Resolution Scalar Quantization for Gaussian Sources and Absolute Error, *IEEE Trans. Inf. Theory*, **53** (2007), 1177-1189.

[27] G. Pisier, *The Volume of Convex Bodies and Banach Space Geometry*, Cambridge University Press (Cambridge, 1989).

[28] G. Sullivan, Efficient Scalar Quantization of Exponential and Laplacian Random Variables, *IEEE Trans. Inf. Theory*, **42** (1996), 13651374.

[29] P. L. Zador, Development and evaluation of procedures for quantizing multivariate distributions, Ph.D. dissertation, Stanford Univ. (1963).

[30] P. L. Zador, Asymptotic quantization error of continuous signals and the quantization dimension, *IEEE Trans. Inf. Theory*, **28** (1982), 139-149.