

## DISTORTION MISMATCH IN THE QUANTIZATION OF PROBABILITY MEASURES

SIEGFRIED GRAF<sup>1</sup>, HARALD LUSCHGY<sup>2</sup> AND GILLES PAGÈS<sup>3</sup>

**Abstract.** We elucidate the asymptotics of the  $L^s$ -quantization error induced by a sequence of  $L^r$ -optimal  $n$ -quantizers of a probability distribution  $P$  on  $\mathbb{R}^d$  when  $s > r$ . In particular we show that under natural assumptions, the optimal rate is preserved as long as  $s < r + d$  (and for every  $s$  in the case of a compactly supported distribution). We derive some applications of these results to the error bounds for quantization based cubature formulae in numerical integration on  $\mathbb{R}^d$  and on the Wiener space.

**Mathematics Subject Classification.** 60G15, 60G35, 41A25.

Received June 2, 2006. Revised November 24, 2006.

### 1. INTRODUCTION

Optimal quantization is devoted to the best approximation in  $L^r_{\mathbb{R}^d}(\mathbb{P})$  ( $r > 0$ ) of a random vector  $X : (\Omega, \mathcal{A}, \mathbb{P}) \rightarrow \mathbb{R}^d$  by random vectors taking finitely many values in  $\mathbb{R}^d$  (endowed with a norm  $\|\cdot\|$ ). When  $X \in L^r(\mathbb{P})$ , this leads for every  $n \geq 1$  to the following  $n$ -level  $L^r(\mathbb{P})$ -optimal quantization problem for the random vector  $X$  defined by

$$e_{n,r}(X) := \inf \{ \|X - q(X)\|_r, q : \mathbb{R}^d \rightarrow \mathbb{R}^d, \text{Borel function, } \text{card}(q(\mathbb{R}^d)) \leq n \}. \quad (1.1)$$

One shows that the above infimum can be taken over the Borel functions

$$q : \mathbb{R}^d \rightarrow \alpha := q(\mathbb{R}^d), \quad \alpha \subset \mathbb{R}^d, \quad \text{card} \alpha \leq n$$

which are some *projection following the nearest neighbour rule* on their image *i.e.*

$$q(x) = \sum_{a \in \alpha} a \mathbf{1}_{V_a(\alpha)}(x),$$

---

*Keywords and phrases.* Optimal quantization, Zador Theorem.

<sup>1</sup> Universität Passau, Fakultät für Informatik und Mathematik, 94030 Passau, Germany; [graf@fim.uni-passau.de](mailto:graf@fim.uni-passau.de)

<sup>2</sup> Universität Trier, FB IV-Mathematik, 54286 Trier, Germany; [luschgy@uni-trier.de](mailto:luschgy@uni-trier.de)

<sup>3</sup> Laboratoire de Probabilités et Modèles aléatoires, UMR 7599, Université Paris 6, case 188, 4, pl. Jussieu, 75252 Paris cedex 5, France; [gpa@ccr.jussieu.fr](mailto:gpa@ccr.jussieu.fr)

$(V_a(\alpha))_{a \in \alpha}$  being a Borel partition of  $\mathbb{R}^d$  satisfying

$$V_a(\alpha) \subset \left\{ x \in \mathbb{R}^d : \|x - a\| = \min_{b \in \alpha} \|x - b\| \right\}.$$

The set  $\alpha = q(\mathbb{R}^d)$  is (also) called a *Voronoi  $n$ -quantizer* and one denotes

$$\widehat{X}^\alpha := q(X).$$

Then, if  $d(x, \alpha) := \min_{a \in \alpha} \|x - a\|$  denotes the distance of  $x$  to the set  $\alpha$ , one has

$$\|X - \widehat{X}^\alpha\|_r^r = \mathbb{E} d(X, \alpha)^r = \int_{\mathbb{R}^d} d(x, \alpha)^r \mathbb{P}_X(dx)$$

which shows that  $e_{n,r}(X)$  actually only depends on the distribution  $P = \mathbb{P}_X$  of  $X$  so that

$$e_{n,r}(X) = e_{n,r}(P) = \inf_{\text{card}(\alpha) \leq n} \left( \int d(x, \alpha)^r dP(x) \right)^{\frac{1}{r}}.$$

The first two basic results in optimal quantization theory are the following (see [6]):

- The above infimum is in fact a minimum: there exists for every  $n \geq 1$  (at least) one  $L^r(\mathbb{P})$ -optimal  $n$ -quantizer  $\alpha_n^*$ . If  $\text{supp}(P)$  is infinite, then  $\text{card}(\alpha_n^*) = n$ .
- Zador's Theorem: If  $X \in L^{r+\eta}(\mathbb{P})$  i.e.  $\int_{\mathbb{R}^d} \|x\|^{r+\eta} dP(x) < +\infty$  for some  $\eta > 0$ , then

$$\lim_n n^{\frac{1}{d}} e_{n,r}(P) = (Q_r(P))^{\frac{1}{r}} \in \mathbb{R}_+.$$

A more explicit expression is known for the real constant  $Q_r(P)$  (see (2.3) below). In particular,  $Q_r(P) > 0$  if and only if  $P$  has an absolutely continuous part (with respect to the Lebesgue measure  $\lambda_d$  on  $\mathbb{R}^d$ ). When  $P$  has an absolutely continuous part, a sequence  $(\alpha_n)_{n \geq 1}$  of  $n$ -quantizers is  *$L^r$ -rate optimal* for  $P$  if

$$\limsup_n n^{\frac{r}{d}} \int_{\mathbb{R}^d} d(x, \alpha_n)^r \mathbb{P}_X(dx) < +\infty$$

and is *asymptotically  $L^r$ -optimal* if  $\lim_n \frac{\int_{\mathbb{R}^d} d(x, \alpha_n)^r \mathbb{P}_X(dx)}{e_{n,r}(P)^r} = 1$ .

Our aim in this paper is to deeply investigate the (asymptotic)  $L^s$ -quantization error induced by a sequence  $(\alpha_n)_{n \geq 1}$  of  $L^r$ -optimal  $n$ -quantizers. It follows from the monotony of  $s \mapsto \|\cdot\|_s$  that  $(\alpha_n)_{n \geq 1}$  remains an  $L^s$ -rate optimal sequence as long as  $s \leq r$ . As soon as  $s > r$  no such straightforward answer is available (except for the uniform distribution over the unit interval since the sequence  $((\frac{2k-1}{2n})_{1 \leq k \leq n})_{n \geq 1}$  is  $L^r$ -optimal for every  $r > 0$ ).

Our main motivation for investigating this problem comes from the recent application of optimal quantization to numerical integration (see [11]) and to the computation of conditional expectation (e.g. for the pricing of American options, see [1]). Let us consider for the sake of simplicity the case of the error bound in the quantization based cubature formulae for numerical integration. Let  $F : \mathbb{R}^d \rightarrow \mathbb{R}$  be a  $\mathcal{C}^1$  function with a Lipschitz continuous differential  $DF$ . It follows from a simple Taylor expansion (see [11]) that for any random vector  $X$  with distribution  $P = \mathbb{P}_X$  quantized by  $\widehat{X}^\alpha$  ( $\alpha \subset \mathbb{R}^d$ )

$$\left| \mathbb{E}(F(X)) - \mathbb{E}(F(\widehat{X}^\alpha)) - \mathbb{E}(DF(\widehat{X}^\alpha) \cdot (X - \widehat{X}^\alpha)) \right| \leq [DF]_{\text{Lip}} \mathbb{E}|X - \widehat{X}^\alpha|^2$$

where  $[DF]_{\text{Lip}}$  denotes the Lipschitz coefficient of  $DF$ . If  $\alpha$  is an  $L^2$ -optimal (or *quadratic*) quantizer then it is *stationary* (see [11] or [6]) so that

$$\widehat{X}^\alpha = \mathbb{E}(X | \widehat{X}^\alpha)$$

which makes the first order term vanish since

$$\mathbb{E}(DF(\widehat{X}^\alpha) \cdot (X - \widehat{X}^\alpha)) = \mathbb{E}(DF(\widehat{X}^\alpha) \cdot \mathbb{E}(X - \widehat{X}^\alpha | \widehat{X}^\alpha)) = 0.$$

Finally, if  $(\alpha_n)_{n \geq 1}$  is a sequence of *quadratic* optimal  $n$ -quantizers

$$\left| \mathbb{E}(F(X)) - \mathbb{E}(F(\widehat{X}^{\alpha_n})) \right| \leq [DF]_{\text{Lip}}(e_{n,2}(P))^2 \sim [DF]_{\text{Lip}} Q_2(P) n^{-\frac{2}{d}}. \tag{1.2}$$

Now, if the Hessian  $D^2F$  does exist, is  $\rho$ -Hölder ( $\rho \in (0, 1]$ ) and computable, the same approach yields

$$\left| \mathbb{E}(F(X)) - \mathbb{E}(F(\widehat{X}^{\alpha_n})) - \mathbb{E}((X - \widehat{X}^{\alpha_n})^* D^2F(\widehat{X}^{\alpha_n})(X - \widehat{X}^{\alpha_n})) \right| \leq [D^2F]_\rho \|X - \widehat{X}^{\alpha_n}\|_{2+\rho}^{2+\rho}. \tag{1.3}$$

Consequently elucidating the asymptotic behaviour of  $\|X - \widehat{X}^{\alpha_n}\|_{2+\rho} = (\int_{\mathbb{R}^d} d(x, \alpha_n)^{2+\rho} dP(x))^{\frac{1}{2+\rho}}$  is necessary to evaluate to what extent the cubature formula in (1.3) does improve the former one (1.2). Similar problems occur when evaluating the error in the first order quantization based scheme designed for the pricing of multi-asset American options or for non-linear filtering (see [2, 13]). One also meets such mismatch problems in infinite dimensions when dealing with (product) functional quantization on the Wiener space in order to price path-dependent European options (see the example in Sect. 6 and [12]).

The paper is organized as follows: in Section 2, a lower bound for the  $L^s(\mathbb{P})$ -quantization rate of convergence of an asymptotically  $L^r$ -optimal sequence  $(\alpha_n)_{n \geq 1}$  of  $n$ -quantizers is established. In particular this result implies that for absolutely continuous distributions  $P$  with unbounded support, the quantization rate  $n^{-\frac{1}{d}}$  in  $L^s$  cannot be preserved as soon as  $s > r + d$ . If  $s \leq r + d$ , then the lower bound can be finite. We conjecture that, when  $(\alpha_n)_{n \geq 1}$  is  $L^r$ -rate optimal the lower bound is in fact the sharp rate. In Section 3, several natural criteria on the distribution  $P$  are derived. They ensure that  $(\alpha_n)_{n \geq 1}$  is  $L^s$ -rate optimal for a given  $s \in (r, r + d)$  or even for all  $s \in (r, r + d)$ . Our criteria are applied to many parametrized families of distributions on  $\mathbb{R}^d$ . We investigate by the same method in Section 4 the critical case  $s = r + d$  and the super-critical case  $s > r + d$ . In Section 5 we show that for compactly supported distributions on the real line the lower bound obtained in Section 2 does hold as a sharp rate. Finally, in Section 6 we apply our results to the evaluation of errors in numerical integration by quantization based cubature formulae in finite and infinite dimensions.

NOTATIONS :  $\bullet \|\cdot\|$  will denote a norm on  $\mathbb{R}^d$  and  $B(x, r)$  will denote the closed ball centred at  $x$  with radius  $r > 0$  (with respect to this norm),  $d(x, A)$  will denote the distance between  $x \in \mathbb{R}^d$  and a subset  $A \subset \mathbb{R}^d$ .

- $\lambda_d$  will denote the Lebesgue measure on  $\mathbb{R}^d$  (equipped with its Borel  $\sigma$ -field  $\mathcal{B}(\mathbb{R}^d)$ ).
- Let  $(a_n)_{n \geq 0}$  and  $(b_n)_{n \geq 0}$  be two sequences of positive real numbers. The symbol  $a_n \asymp b_n$  is for  $a_n = O(b_n)$  and  $b_n = O(a_n)$  whereas the symbol  $a_n \sim b_n$  means  $a_n = b_n + o(b_n)$  as  $n \rightarrow \infty$ .
- $[x]$  is for the integral part of the real number  $x$ .
- $f \propto g$  means that the functions  $f$  and  $g$  are proportional.

## 2. THE LOWER ESTIMATE

In this section we derive an explicit lower bound in the  $(r, s)$ -problem for non-purely singular probability distributions  $P$  and asymptotically  $L^r$ -optimal quantizers. This bound is expected to be best possible.

Let  $r \in (0, \infty)$ . Let  $P$  be a probability measure on  $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$  satisfying

$$\int_{\mathbb{R}^d} \|x\|^r dP(x) < +\infty \tag{2.1}$$

and  $\text{supp}(P)$  is not finite. Then  $e_{n,r}(P) \in (0, \infty)$  for every  $n$  and  $e_{n,r}(P) \rightarrow 0$  as  $n \rightarrow \infty$ . A sequence  $(\alpha_n)_{n \geq 1}$  of quantizers is called asymptotically  $L^r$ -optimal for  $P$  if  $\text{card } \alpha_n \leq n$  for every  $n$  and

$$\int_{\mathbb{R}^d} d(x, \alpha_n)^r dP(x) \sim e_{n,r}(P)^r \quad \text{as } n \rightarrow \infty. \quad (2.2)$$

Let  $P^a = f \cdot \lambda_d$  denote the absolutely continuous part of  $P$  with respect to  $\lambda_d$ . Assume that  $\int_{\mathbb{R}^d} \|x\|^{r+\eta} dP(x) < +\infty$  for some  $\eta > 0$ . Then by the Zador Theorem (see [6])

$$\lim_{n \rightarrow \infty} n^{r/d} e_{n,r}(P)^r = Q_r(P) \quad (2.3)$$

where

$$Q_r(P) := J_{r,d} \left( \int_{\mathbb{R}^d} f^{d/(d+r)} d\lambda_d \right)^{(d+r)/d} \in [0, \infty) \quad (2.4)$$

and

$$J_{r,d} := \inf_{n \geq 1} n^{r/d} e_{n,r}(U([0, 1]^d))^r \in (0, \infty),$$

( $U([0, 1]^d)$  denotes the uniform distribution on the hyper-cube  $[0, 1]^d$ ). This theorem was first stated by Zador in [14, 15] and then generalized by Bucklew and Wise (see [3]), the first completely rigorous proof has been proposed by Graf & Luschgy in [6]. Note that the finiteness of  $\int_{\mathbb{R}^d} f^{d/(d+r)} d\lambda_d$  is a simple consequence of the Hölder Inequality and the moment assumption  $\int_{\mathbb{R}^d} \|x\|^{r+\eta} dP(x) < +\infty$ : first note that  $\int_{\|x\| \leq 1} f^{d/(d+r)} d\lambda_d < +\infty$  since  $\lambda_d(\|x\| \leq 1) < +\infty$  and  $\frac{d}{d+r} \leq 1$ . Then setting  $p = 1 + r/d$ ,  $q = 1 + d/r$ ,  $\alpha = (r + \eta)d/(d + r)$

$$\begin{aligned} \int_{B(0,1)^c} f^{d/(d+r)} d\lambda_d &= \int_{\|x\| > 1} \|x\|^{-\alpha} \|x\|^\alpha f^{d/(d+r)}(x) d\lambda_d(x) \\ &\leq \left( \int_{\|x\| > 1} \|x\|^{-\alpha q} d\lambda_d(x) \right)^{1/q} \left( \int \|x\|^{\alpha p} f(x) d\lambda_d(x) \right)^{1/p} \\ &= \left( \int_{\|x\| > 1} \|x\|^{-(1+\eta/r)d} d\lambda_d(x) \right)^{1/q} \left( \int \|x\|^{r+\eta} dP(x) \right)^{1/p} < +\infty. \end{aligned}$$

Furthermore, for probabilities  $P$  on  $\mathbb{R}^d$  with  $P^a \neq 0$ , the empirical measures associated to an asymptotically  $L^r$ -optimal sequence  $(\alpha_n)_{n \geq 1}$  of  $n$ -quantizers satisfy (see [6] Th. 7.5 or [4] for this slight extension)

$$\frac{1}{n} \sum_{a \in \alpha_n} \delta_a \xrightarrow{w} P_r \quad (2.5)$$

where  $P_r$  denotes the  $L^r$ -point density measure of  $P$  defined by

$$P_r := f_r \cdot \lambda_d \quad \text{with} \quad f_r := \frac{f^{d/(d+r)}}{\int f^{d/(d+r)} d\lambda_d}. \quad (2.6)$$

Note that the limit  $Q_r(P)$  in the Zador Theorem reads

$$Q_r(P) = J_{r,d} \int f_r^{-r/d} dP^a.$$

The quantity that naturally comes out in the  $(r, s)$ -problem,  $r, s \in (0, \infty)$ , is

$$\begin{aligned} Q_{r,s}(P) &:= J_{s,d} \int f_r^{-s/d} dP^a \\ &= J_{s,d} \left( \int_{\mathbb{R}^d} f^{d/(d+r)} d\lambda_d \right)^{s/d} \int_{\{f>0\}} f^{1-s/(d+r)} d\lambda_d \in (0, +\infty]. \end{aligned} \quad (2.7)$$

**Theorem 1.** *Let  $r, s \in (0, \infty)$ . Assume  $P^a \neq 0$  and  $\int_{\mathbb{R}^d} \|x\|^{r+\eta} dP(x) < +\infty$  for some  $\eta > 0$ . Let  $(\alpha_n)_{n \geq 1}$  be an asymptotically  $L^r$ -optimal sequence of  $n$ -quantizers for  $P$ . Then*

$$\liminf_{n \rightarrow \infty} n^{s/d} \int d(x, \alpha_n)^s dP(x) \geq Q_{r,s}(P). \quad (2.8)$$

Prior to the proof, let us provide a few comments on this lower bound.

COMMENTS. • The main corollary that can be directly derived from Theorem 1 is that

$$\int_{\{f>0\}} f^{1-s/(d+r)} d\lambda_d = +\infty \implies \lim_{n \rightarrow \infty} n^{s/d} \int d(x, \alpha_n)^s dP(x) = +\infty$$

since then  $Q_{r,s}(P) = +\infty$ .

By contraposition, a necessary condition for an asymptotically  $L^r$ -optimal sequence of quantizers  $(\alpha_n)$  to achieve the optimal rate  $n^{-s/d}$  for the  $L^s$ -quantization error is that  $Q_{r,s}(P) < +\infty$ . But, under the moment assumption of Theorem 1 the following equivalence holds true

$$Q_{r,s}(P) < +\infty \iff \int f^{-\frac{s}{d+r}} dP^a = \int_{\{f>0\}} f^{1-\frac{s}{d+r}} d\lambda_d < +\infty \quad (2.9)$$

since  $\int_{\mathbb{R}^d} f^{d/(d+r)} d\lambda_d < +\infty$ .

In turn, for probability measures  $P$  satisfying  $\lambda_d(f > 0) = +\infty$  a necessary condition for the right hand side of (2.9) to be satisfied is that

$$s < d + r. \quad (2.10)$$

Indeed, if  $s \geq d + r$ , the following chain of inequalities holds true

$$\lambda_d(f > 0) = \int \mathbf{1}_{\{f>0\}} f^{-1} dP^a \leq \left( \int (\mathbf{1}_{\{f>0\}} f^{-1})^{\frac{s}{d+r}} dP^a \right)^{\frac{d+r}{s}} = \left( \int_{\{f>0\}} f^{1-s/(d+r)} d\lambda_d \right)^{\frac{d+r}{s}}$$

where we used that  $p \mapsto \|\cdot\|_{L^p(P^a)}$  is non-decreasing since  $P^a(\mathbb{R}^d) \leq 1$ .

On the other hand, still when  $s < d + r$ , the following criterion holds for the finiteness of  $Q_{r,s}(P)$ :

$$\left( \exists \vartheta > 0, \int_{\mathbb{R}^d} \|x\|^{ds/(d+r-s)+\vartheta} dP(x) < +\infty \right) \implies Q_{r,s}(P) < +\infty. \quad (2.11)$$

Set  $\rho = 1 - \frac{s}{d+r} \in (0, 1)$  and  $u = \frac{ds}{d+r-s} + \vartheta$ . Then (2.11) follows from the regular Hölder inequality applied with  $\tilde{p} = \frac{1}{\rho} = \frac{d+r}{d+r-s}$  and  $\tilde{q} = \frac{1}{1-\rho} = \frac{s}{d+r-s}$ ,

$$\begin{aligned} \int_{B(0,1)^c} f^\rho d\lambda_d &\leq \left( \int_{B(0,1)^c} (f(x)^\rho \|x\|^{u\rho})^{\tilde{p}} d\lambda_d(x) \right)^{1/\tilde{p}} \left( \int_{B(0,1)^c} \|x\|^{-u\rho\tilde{q}} d\lambda_d(x) \right)^{1/\tilde{q}} \\ &= \left( \int_{B(0,1)^c} f(x) \|x\|^u d\lambda_d(x) \right)^\rho \left( \int_{B(0,1)^c} \|x\|^{-u\rho/(1-\rho)} d\lambda_d(x) \right)^{1-\rho} < +\infty \end{aligned}$$

using the moment assumption in (2.11) and  $u\rho/(1-\rho) = d + \vartheta \frac{\rho}{1-\rho} > d$ .

- It is generally not true in the general setting of Theorem 1 that  $\lim_{n \rightarrow \infty} n^{s/d} \int d(x, \alpha_n)^s dP(x) = Q_{r,s}(P)$  (see Counter-Example 2 in Sect. 3). However, one may reasonably conjecture that this limiting result holds true for sequences  $(\alpha_n)$  of *exactly*  $L^r$ -optimal  $n$ -quantizers. Our result in one dimension for compactly supported distributions (see Sect. 5) supports this conjecture.

- In any case, note that (2.8) improves the obvious lower bound

$$\liminf_{n \rightarrow \infty} n^{s/d} \int d(x, \alpha_n)^s dP(x) \geq \liminf_{n \rightarrow \infty} n^{s/d} e_{n,s}(P)^s \geq Q_s(P).$$

(The right inequality needs no moment assumption on  $P$  as can be checked from the proof of the Zador Theorem, see [6].) In fact, one even has that, for every  $r, s \in (0, +\infty)$ ,

$$Q_{r,s}(P) \geq Q_s(P).$$

Furthermore, this inequality is strict when  $r \neq s$  (except if  $f$  is  $\lambda_d$ -a.e. constant on  $\{f > 0\}$ ). Let us provide a short proof of this fact. Set  $p = (d+s)/s > 1$ ,  $q = (d+s)/d >$  and  $a = ds/(d+r)(d+s)$ ,  $b = (d+r-s)d/(d+r)(d+s)$ . Then the Hölder inequality yields

$$\begin{aligned} (Q_s(P))^{\frac{d}{d+s}} &= \int f^{d/(d+s)} d\lambda_d = \int_{\{f>0\}} f^a f^b d\lambda_d \\ &\leq \left( \int f^{ap} d\lambda_d \right)^{1/p} \left( \int_{\{f>0\}} f^{bq} d\lambda_d \right)^{1/q} \quad (\text{"<" if } f^{ap} \text{ and } f^{bq} \text{ are not proportional}) \\ &= \left( \int f^{d/(d+r)} d\lambda_d \right)^{s/(d+s)} \left( \int_{\{f>0\}} f^{1-\frac{s}{d+r}} d\lambda_d \right)^{d/(d+s)} \\ &= (Q_{r,s}(P))^{\frac{d}{d+s}}. \end{aligned}$$

*Proof of Theorem 1.* First keep in mind that, the  $r + \eta$ -moment assumption on  $P$  implies the finiteness of  $\int f^{\frac{d}{d+r}} d\lambda_d$ . The existence of at least one asymptotically  $L^r$ -optimal sequence  $(\alpha_n)_{n \geq 1}$  follows from the existence of an  $r$ -moment for  $P$ . For every integer  $m \geq 1$ , set

$$f_m := \sum_{k=0}^{m2^m-1} \frac{k}{2^m} \mathbf{1}_{E_k^m} \quad \text{with} \quad E_k^m = \left\{ \frac{k}{2^m} \leq f < \frac{k+1}{2^m} \right\} \cap B(0, m).$$

The sequence  $(f_m)_{m \geq 1}$  is non-decreasing and converges to  $f \mathbf{1}_{\{0 \leq f < +\infty\}} = f$   $\lambda_d$ -a.e.

Let  $I_m := \{k \in \{0, \dots, m2^m - 1\} : \lambda_d(E_k^m) > 0\}$ .

For every  $k \in I_m$ , there exists a closed set  $A_k^m \subset E_k^m$  satisfying

$$\lambda_d(E_k^m \setminus A_k^m) \leq \frac{1}{m^3 2^m}.$$

Let  $\varepsilon_m \in (0, 1]$  be a positive real number such that the closed sets  $\tilde{A}_k^m := \{x \in \mathbb{R}^d : d(x, A_k^m) \leq \varepsilon_m\}$ ,  $k \in I_m$ , satisfy

$$\int_{\tilde{A}_k^m} f^{\frac{d}{d+r}} d\lambda_d \leq (1 + 1/m) \int_{A_k^m} f^{\frac{d}{d+r}} d\lambda_d < +\infty.$$

Set

$$\tilde{f}_m := \sum_{k=0}^{m2^m-1} \frac{k}{2^m} \mathbf{1}_{A_k^m}.$$

It is clear that

$$\{f_m \neq \tilde{f}_m\} \subset \bigcup_{0 \leq k \leq m2^m-1} (E_k^m \setminus A_k^m).$$

Hence

$$\lambda_d(\{f_m \neq \tilde{f}_m\}) \leq \sum_{k=0}^{m2^m-1} \frac{1}{m^3 2^m} = \frac{1}{m^2}$$

so that

$$\sum_{m \geq 1} \mathbf{1}_{\{f_m \neq \tilde{f}_m\}} < +\infty \quad \lambda_d\text{-a.e.}$$

*i.e.*, for  $\lambda_d$ -a.e.  $x$ ,  $f_m(x) = \tilde{f}_m(x)$  for large enough  $m$  so that  $\tilde{f}_m$  converges to  $f$   $\lambda_d$ -a.e.. Finally, as a result  $\tilde{f}_m \leq f_m \leq f$  and  $f_m$  converges to  $f$   $\lambda_d$ -a.e. Then, for every  $n \geq 1$ ,

$$\begin{aligned} n^{\frac{s}{d}} \int_{\mathbb{R}^d} (d(x, \alpha_n))^s dP(x) &\geq n^{\frac{s}{d}} \int_{\mathbb{R}^d} (d(x, \alpha_n))^s \tilde{f}_m(x) d\lambda_d(x) \\ &= n^{\frac{s}{d}} \sum_{k=0}^{m2^m-1} \frac{k}{2^m} \int_{A_k^m} (d(x, \alpha_n))^s d\lambda_d(x). \end{aligned} \quad (2.12)$$

Since all the sets  $\tilde{A}_k^m$ ,  $k = 0, \dots, m2^m - 1$  are bounded (as subsets of  $B(0, m+1)$ ), there exists for every  $m \geq 1$  and every  $k \in \{0, \dots, m2^m - 1\}$  a finite “firewall”  $\beta_k^m \subset \mathbb{R}^d$  (see [6] or Lem. 4.3 in [4] and note that  $A_k^m \subset (\tilde{A}_k^m)_{\varepsilon_m/2} := \{x \in \mathbb{R}^d : d(x, (\tilde{A}_k^m)^c) > \varepsilon_m/2\}$ ) such that

$$\forall n \geq 1, \quad \forall x \in A_k^m, \quad d(x, \alpha_n \cup \beta_k^m) = d(x, (\alpha_n \cup \beta_k^m) \cap \tilde{A}_k^m).$$

Set  $\beta^m = \bigcup_{0 \leq k \leq m2^m-1} \beta_k^m$ . Then, for every  $k \in \{0, \dots, m2^m - 1\}$ , for every  $x \in A_k^m$ ,

$$d(x, \alpha_n) \geq d(x, \alpha_n \cup \beta_k^m) = d(x, (\alpha_n \cup \beta_k^m) \cap \tilde{A}_k^m) \geq d(x, (\alpha_n \cup \beta^m) \cap \tilde{A}_k^m).$$

Set temporarily  $n_k^m := \text{card}((\alpha_n \cup \beta^m) \cap \tilde{A}_k^m)$ . First note that it is clear that

$$\frac{n_k^m}{n} \sim \frac{\text{card}(\alpha_n \cap \tilde{A}_k^m)}{n} \quad \text{as } n \rightarrow \infty.$$

It follows from the asymptotic  $L^r$ -optimality of the sequence  $(\alpha_n)$  and the empirical measure theorem (see (2.5)) that

$$\limsup_n \frac{\text{card}(\alpha_n \cap \tilde{A}_k^m)}{n} \leq \frac{\int_{\tilde{A}_k^m} f^{\frac{d}{d+r}} d\lambda_d}{\int f^{\frac{d}{d+r}} d\lambda_d} \quad (2.13)$$

so that

$$\liminf_n \frac{n}{n_k^m} \geq \frac{\int f^{\frac{d}{d+r}} d\lambda_d}{\int_{\tilde{A}_k^m} f^{\frac{d}{d+r}} d\lambda_d} \geq \frac{m}{m+1} \frac{\int f^{\frac{d}{d+r}} d\lambda_d}{\int_{A_k^m} f^{\frac{d}{d+r}} d\lambda_d}.$$

On the other hand, for every  $k \in I_m$ ,

$$\int_{A_k^m} (d(x, \alpha_n))^s d\lambda_d(x) \geq \int_{A_k^m} d(x, (\alpha_n \cup \beta^m) \cap \tilde{A}_k^m)^s d\lambda_d(x) \geq \lambda_d(A_k^m) e_{n_k^m, s}^s(U(A_k^m))$$

where  $U(A_k^m)$  denotes the uniform distribution over  $A_k^m$  (note that the inequality is trivial when  $\lambda_d(A_k^m) = 0$ ). Then one may apply Zador's theorem which yields, combined with (2.13),

$$\begin{aligned} \liminf_n n^{\frac{s}{d}} \int_{A_k^m} (d(x, \alpha_n))^s d\lambda_d(x) &\geq \lambda_d(A_k^m) \times \liminf_n \left(\frac{n}{n_k^m}\right)^{\frac{s}{d}} \times \lim_{n_k^m} ((n_k^m)^{\frac{s}{d}} e_{n_k^m, s}^s(U(A_k^m)))^s \\ &\geq \lambda_d(A_k^m) \times \left(\frac{m}{m+1} \frac{\int f^{\frac{d}{d+r}} d\lambda_d}{\int_{A_k^m} f^{\frac{d}{d+r}} d\lambda_d}\right)^{\frac{s}{d}} J_{s,d} \times (\lambda_d(A_k^m))^{\frac{s}{d}} \\ &\geq J_{s,d} \left(\frac{m}{m+1}\right)^{\frac{s}{d}} \left(\int f^{\frac{d}{d+r}} d\lambda_d\right)^{\frac{s}{d}} \left(\frac{\lambda_d(A_k^m)}{\int_{A_k^m} f^{\frac{d}{d+r}} d\lambda_d}\right)^{\frac{s}{d}} \lambda_d(A_k^m) \\ &\geq J_{s,d} \left(\frac{m}{m+1}\right)^{\frac{s}{d}} \left(\int f^{\frac{d}{d+r}} d\lambda_d\right)^{\frac{s}{d}} \left(\frac{k+1}{2^m}\right)^{-\frac{s}{d+r}} \lambda_d(A_k^m) \end{aligned}$$

with the convention  $\frac{0}{0} = 0$ .

Consequently, using (2.12) and the super-additivity of  $\liminf$  yield that, for every  $m \geq 1$ ,

$$\begin{aligned} \liminf_n n^{\frac{s}{d}} \int_{\mathbb{R}^d} (d(x, \alpha_n))^s dP(x) &\geq J_{s,d} \left(\frac{m}{m+1}\right)^{\frac{s}{d}} \left(\int_{\mathbb{R}^d} f^{\frac{d}{d+r}} d\lambda_d\right)^{\frac{s}{d}} \sum_{k=0}^{m2^m-1} \frac{k}{2^m} \left(\frac{k+1}{2^m}\right)^{-\frac{s}{d+r}} \lambda_d(A_k^m) \\ &= J_{s,d} \left(\frac{m}{m+1}\right)^{\frac{s}{d}} \left(\int_{\mathbb{R}^d} f^{\frac{d}{d+r}} d\lambda_d\right)^{\frac{s}{d}} \int_{\{f>0\}} \tilde{f}_m (\tilde{f}_m + 2^{-m})^{-\frac{s}{d+r}} d\lambda_d. \end{aligned}$$

Now, by Fatou's Lemma, one concludes by letting  $m$  go to infinity that

$$\liminf_n n^{\frac{s}{d}} \int_{\mathbb{R}^d} (d(x, \alpha_n))^s dP(x) \geq J_{s,d} \left(\int_{\mathbb{R}^d} f^{\frac{d}{d+r}} d\lambda_d\right)^{\frac{s}{d}} \int_{\{f>0\}} f^{1-\frac{s}{d+r}} d\lambda_d. \quad \square$$

### 3. THE UPPER ESTIMATE

Let  $r, s \in (0, \infty)$ . In this section we investigate whether the upper bound  $\int d(x, \alpha_n)^s dP(x) = O(n^{-s/d})$  for  $L^r$ -optimal  $n$ -quantizers  $\alpha_n$  holds true. (This is of course less precise than the lower bound given in the previous section.) The reason for the restriction to (exactly) optimal  $n$ -quantizers (when  $s > r$ ) instead of only asymptotically optimal  $n$ -quantizers will become clear soon. See *e.g.* the subsequent Example 2. First note that the  $L^p(P)$ -norms being non-decreasing as a function of  $p$ , the above upper bound trivially holds for  $s \in (0, r]$  since

$$n^{s/d} \int d(x, \alpha_n)^s dP(x) \leq \left(n^{r/d} \int d(x, \alpha_n)^r dP(x)\right)^{\frac{s}{r}}.$$

The same argument shows that when this rate holds for some  $s > 0$ , then it holds for every  $s' \in (0, s]$ .

For a sequence  $(\alpha_n)_{n \geq 1}$  of finite codebooks in  $\mathbb{R}^d$  and  $b \in (0, \infty)$  we introduce the maximal function  $\psi_b : \mathbb{R}^d \rightarrow \mathbb{R}_+ \cup \{\infty\}$  by

$$\psi_b(x) := \sup_{n \geq 1} \frac{\lambda_d(B(x, bd(x, \alpha_n)))}{P(B(x, bd(x, \alpha_n)))} \tag{3.1}$$

(with the interpretation  $\frac{0}{0} := 0$ ). Note that  $\psi_b$  is Borel-measurable and depends on the underlying norm on  $\mathbb{R}^d$ . The theorem below provides a criterion based on these maximal functions that ensures the  $L^s$ -rate optimality of  $L^r$ -optimal  $n$ -quantizers. In Corollaries 1, 3, 4 we derive more applicable criteria which only involve the distribution  $P$ .

**Theorem 2.** *Let  $r, s \in (0, \infty)$ . Assume  $P^a \neq 0$  and  $\int \|x\|^{r+\eta} dP(x) < +\infty$  for some  $\eta > 0$ . For every  $n \geq 1$ , let  $\alpha_n$  be an  $L^r$ -optimal  $n$ -quantizer for  $P$ . Assume that the maximal function associated with the sequence  $(\alpha_n)$  satisfies*

$$\psi_b^{s/(d+r)} \in L^1(P) \tag{3.2}$$

for some  $b \in (0, 1/2)$ . Then

$$\sup_n n^{s/d} \int d(x, \alpha_n)^s dP(x) < +\infty. \tag{3.3}$$

*Proof.* Let  $y \in \mathbb{R}^d$  and set  $\delta = \delta_n = d(y, \alpha_n)$ . For every  $x \in B(y, \delta/2)$  and  $a \in \alpha_n$ , we have  $\|x - a\| \geq \|y - a\| - \|x - y\| \geq \delta/2$  and hence

$$d(x, \alpha_n) \geq \delta/2 \geq \|x - y\|, \quad x \in B(y, \delta/2).$$

Let  $\beta = \beta_n = \alpha_n \cup \{y\}$ . Then

$$d(x, \beta) = \|x - y\|, \quad x \in B(y, \delta/2).$$

Consequently, for every  $b \in (0, 1/2)$ ,

$$\begin{aligned} e_{n,r}(P)^r - e_{n+1,r}(P)^r &\geq \int d(x, \alpha_n)^r dP(x) - \int d(x, \beta)^r dP(x) \\ &\geq \int_{B(y, \delta b)} (d(x, \alpha_n)^r - d(x, \beta)^r) dP(x) \\ &= \int_{B(y, \delta b)} (d(x, \alpha_n)^r - \|x - y\|^r) dP(x) \\ &\geq \int_{B(y, \delta b)} ((\delta/2)^r - (b\delta)^r) dP(x) \\ &= ((1/2)^r - b^r) \delta^r P(B(y, b\delta)). \end{aligned}$$

One derives that

$$d(y, \alpha_n)^r \leq \frac{C(b)}{P(B(y, bd(y, \alpha_n)))} (e_{n,r}(P)^r - e_{n+1,r}(P)^r) \tag{3.4}$$

for every  $y \in \mathbb{R}^d, b \in (0, 1/2), n \geq 1$ , where  $C(r, b) = ((1/2)^r - b^r)^{-1}$ . Note that  $e_{n,r}(P)^r - e_{n+1,r}(P)^r > 0$  for every  $n \in \mathbb{N}$  (see [6]).

Now we estimate the increments  $e_{n,r}(P)^r - e_{n+1,r}(P)^r$ . (This extends a corresponding estimate in [7] to distributions with possibly unbounded support.)

Set  $e_{n,r} = e_{n,r}(P)$  for convenience. Let  $\{V_a : a \in \alpha_{n+1}\}$  with  $V_a = V_a(\alpha_{n+1})$  be a Voronoi partition of  $\mathbb{R}^d$  with respect to  $\alpha_{n+1}$ . Then  $P(V_a) > 0$  for all  $a \in \alpha_{n+1}$  and  $\text{card } \alpha_{n+1} = n + 1$  (see [6]),

$$\text{card} \left\{ a \in \alpha_{n+1} : \int_{V_a} \|x - a\|^r dP(x) > \frac{4e_{n+1,r}^r}{n + 1} \right\} \leq \frac{n + 1}{4}$$

and

$$\text{card} \left\{ a \in \alpha_{n+1} : P(V_a) > \frac{4}{n+1} \right\} \leq \frac{n+1}{4}.$$

This implies that

$$\beta_{n+1} := \left\{ a \in \alpha_{n+1} : \int_{V_a} \|x - a\|^r dP(x) \leq \frac{4e_{n+1,r}^r}{n+1}, P(V_a) \leq \frac{4}{n+1} \right\}$$

satisfies  $\text{card}\beta_{n+1} \geq (n+1)/2$ . Choose a closed hyper-cube  $K = [-m, m]^d$  such that  $P_r(K) > 3/4$ . The empirical measure theorem (see (2.5) above or [4, 6] for details) implies

$$\lim_{k \rightarrow \infty} \frac{\text{card}(\alpha_k \cap K)}{k} = P_r(K)$$

since  $P_r(\partial K) = \lambda_d(\partial K) = 0$ . We deduce  $\text{card}(\alpha_{n+1} \cap K) \geq 3(n+1)/4$  and hence  $\text{card}(\beta_{n+1} \cap K) \geq (n+1)/4$  for large enough  $n$ . Since one can find a tessellation of  $K$  into  $[(n+1)/8] \vee 1$  cubes of diameter less than  $C_1 n^{-1/d}$ , there exist  $a_1, a_2 \in \beta_{n+1}$ ,  $a_1 \neq a_2$  such that

$$\|a_1 - a_2\| \leq C_1(r)n^{-1/d}$$

for every  $n \geq 3$ . Let  $\gamma = \alpha_{n+1} \setminus \{a_1\}$ . Using

$$d(x, \gamma) \leq \|x - a_2\| \leq \|x - a_1\| + \|a_1 - a_2\|,$$

one obtains

$$\begin{aligned} e_{n,r}(P)^r - e_{n+1,r}(P)^r &\leq \int d(x, \gamma)^r dP(x) - \int d(x, \alpha_{n+1})^r dP(x) \\ &= \sum_{a \in \gamma} \int_{V_a} \|x - a\|^r dP(x) + \int_{V_{a_1}} d(x, \gamma)^r dP(x) - \sum_{a \in \alpha_{n+1}} \int_{V_a} \|x - a\|^r dP(x) \\ &= \int_{V_{a_1}} (d(x, \gamma)^r - \|x - a_1\|^r) dP(x) \\ &\leq (2^r - 1) \int_{V_{a_1}} \|x - a_1\|^r dP(x) + 2^r \|a_1 - a_2\|^r P(V_{a_1}) \\ &\leq \frac{4(2^r - 1)e_{n+1,r}^r}{n+1} + \frac{4 \cdot 2^r C_1(r)^r n^{-r/d}}{n+1}. \end{aligned}$$

Consequently, using (2.3), for every  $n \in \mathbb{N}$ ,

$$e_{n,r}(P)^r - e_{n+1,r}(P)^r \leq C_2(r)n^{-(d+r)/d} \quad (3.5)$$

where  $C_2(r)$  denotes a finite constant independent of  $n$ . Combining (3.4) and (3.5), we get

$$n^{s/d} d(x, \alpha_n)^s \leq C_3(r, b)^s \left( \frac{\lambda_d(B(x, bd(x, \alpha_n)))}{P(B(x, bd(x, \alpha_n)))} \right)^{s/(d+r)} \quad (3.6)$$

$$\leq C_3(r, b)^s \psi_b(x)^{s/(d+r)} \quad (3.7)$$

for every  $x \in \mathbb{R}^d$ ,  $n \in \mathbb{N}$ ,  $b \in (0, 1/2)$  and some finite constant  $C_3(r, b)$ . The proof is completed by integrating both sides with respect to  $P$ .  $\square$

**Application to pointwise convergence rate.** In the situation of Theorem 2, assuming  $P = P^a$ , but without assuming (3.2) (so that  $s$  is not involved in that statement), one can deduce from (3.6) that

$$\limsup_{n \rightarrow \infty} n^{1/d} d(x, \alpha_n) \leq C_3(r, b) f^{-1/(d+r)} < +\infty \quad P(dx)\text{-a.s.} \quad (3.8)$$

since  $d(x, \alpha_n) \rightarrow 0$   $P(dx)$ -a.s. (see [4]) implies in turn

$$\frac{P(B(x, bd(x, \alpha_n)))}{\lambda_d(B(x, bd(x, \alpha_n)))} \rightarrow f(x) \quad P(dx)\text{-a.s.} \quad \text{as } n \rightarrow \infty$$

by the differentiation of measures. This improves considerably for absolutely continuous distributions and (exactly)  $L^r$ -optimal quantizers an a.s. result in [4].

Next we observe that in case  $s \in (0, d+r)$  a local version of condition (3.2) is always satisfied.

**Lemma 1.** *Assume  $\int \|x\|^r dP(x) < +\infty$  for some  $r \in (0, \infty)$ . Let  $(\alpha_n)$  be a sequence of finite codebooks in  $\mathbb{R}^d$  satisfying  $\int d(x, \alpha_n)^r dP(x) \rightarrow 0$ . Then the associated maximal functions  $\psi_b$  are locally in  $L^p(P)$  for every  $p \in (0, 1)$  i.e.*

$$\forall M, b \in (0, \infty), \quad \int_{B(0, M)} \psi_b^p dP < +\infty.$$

*Proof.* Let  $M, b \in (0, \infty)$  and set  $A = \text{supp}(P)$ . Then  $\max_{x \in B(0, M) \cap A} d(x, \alpha_n) \rightarrow 0$  (see [4]) and hence

$$C(M) := \sup_{n \geq 1} \max_{x \in B(0, M) \cap A} d(x, \alpha_n) < +\infty.$$

One derives that

$$B(x, bd(x, \alpha_n)) \subset B(0, bC(M) + M)$$

for every  $x \in B(0, M) \cap A, n \in \mathbb{N}$ . Define the Hardy-Littlewood maximal function  $\varphi : \mathbb{R}^d \rightarrow \mathbb{R}_+ \cup \{\infty\}$  with respect to the finite measures  $\lambda_d(\cdot \cap B(0, bC(M) + M))$  and  $P$  by

$$\varphi(x) = \varphi_{b, M}(x) := \sup_{\rho > 0} \frac{\lambda_d(B(x, \rho) \cap B(0, bC(M) + M))}{P(B(x, \rho))}.$$

Then

$$\psi_b(x) \leq \varphi(x)$$

for every  $x \in B(0, M) \cap A$ . From the Besicovitch covering theorem follows the maximal inequality

$$P(\varphi > \rho) \leq \frac{C_1 \lambda_d(B(0, bC(M) + M))}{\rho}$$

for every  $\rho > 0$  where the finite constant  $C_1$  only depends on  $d$  and the underlying norm. (See [10], Th. 2.19. The result in [10] is stated for Euclidean norms but it obviously extends to arbitrary norms since any two norm on  $\mathbb{R}^d$  are equivalent.) Consequently,

$$\begin{aligned} \int_{B(0, M) \cap A} \psi_b^p dP &\leq \int \varphi^p dP = \int_0^\infty P(\varphi^p > t) dt \leq 1 + \int_1^\infty P(\varphi^p > t) dt \\ &\leq 1 + C_2 \int_1^\infty t^{-1/p} dt < +\infty \end{aligned}$$

where  $C_2 = C_1 \lambda_d(B(0, bC(M) + M))$ . □

In case  $s > d + r$  under a mild assumption on the support of  $P$  a local version of (3.2) holds provided the minimal condition is satisfied locally.

**Lemma 2.** *Set  $A = \text{supp}(P)$ . Assume  $P = P^a = f \cdot \lambda_d$ ,  $\int \|x\|^r dP(x) < +\infty$  for some  $r \in (0, \infty)$ ,  $\lambda_d(\cdot \cap A)$  is absolutely continuous with respect to  $P$  and  $A$  is a finite union of closed convex sets. Let  $(\alpha_n)$  be a sequence of finite codebooks in  $\mathbb{R}^d$  satisfying*

$$\int d(x, \alpha_n)^r dP(x) \rightarrow 0.$$

*Then for  $p \in (1, \infty]$ , the associated maximal functions  $\psi_b$  are locally in  $L^p(P)$  provided  $1/f$  is locally in  $L^p(P)$ .*

**Remark.** The absolute continuity assumption  $\lambda_d(\cdot \cap A) \ll P$  does not follow from the absolute continuity of  $P$ : set  $P = c \mathbf{1}_U \cdot \lambda_d$  where  $U = \cup_{n \geq 0} (r_n - 2^{-(n+1)}, r_n + 2^{-(n+1)})$ ,  $\mathbb{Q} = \{r_n, n \geq 0\}$  and  $c = 1/\lambda_d(U)$ . Then  $\text{supp}(P) = \mathbb{R}$  but  $\lambda_d \not\ll P$  since  $\lambda_d(U^c) = +\infty$  and  $P(U^c) = 0$ .

*Proof.* Let  $M, b \in (0, \infty)$  such that  $B(0, M) \cap A \neq \emptyset$ . We have

$$C = C(M) := \sup_{n \geq 1} \max_{x \in B(0, M) \cap A} d(x, \alpha_n) < +\infty$$

(see [4]) and hence

$$\psi_b(x) \leq \sup_{\rho \leq bC} \frac{\lambda_d(B(x, \rho))}{P(B(x, \rho))}$$

for every  $x \in B(0, M) \cap A$ . There exists a constant  $C_1 \in (0, \infty)$  such that

$$\lambda_d(B(x, \rho) \cap B(0, M) \cap A) \geq C_1 \rho^d$$

for every  $x \in B(0, M) \cap A$  and  $\rho \leq bC$  since  $B(0, M) \cap A$  is finite union of convex sets (see [6]). Define the maximal function  $\varphi = \varphi_{f^{-1}} : \mathbb{R}^d \rightarrow \mathbb{R}_+ \cup \{\infty\}$  by

$$\varphi(x) := \sup_{\rho > 0} \frac{\int_{B(x, \rho)} f^{-1} dP(\cdot \cap B(0, M))}{P(B(x, \rho) \cap B(0, M))}$$

and note that  $d\lambda_d(\cdot \cap A)/dP = 1/f$ . One derives that

$$\begin{aligned} \psi_b(x) &\leq C_2 \sup_{\rho \leq bC} \frac{\lambda_d(B(x, \rho) \cap B(0, M) \cap A)}{P(B(x, \rho) \cap B(0, M))} \\ &\leq C_2 \varphi(x) \end{aligned}$$

for every  $x \in B(0, M) \cap A$  with  $C_2 = \lambda_d(B(0, 1))/C$ . By the  $L^p(P(\cdot \cap B(0, M)))$ -boundedness of the maximal operator  $f^{-1} \mapsto \varphi_{f^{-1}}$  (see [10], Th. 2.19), we obtain

$$\|\psi_b \mathbf{1}_{B(0, M)}\|_{L^p(P)} \leq C_2 \|\varphi \mathbf{1}_{B(0, M)}\|_{L^p(P)} \leq C_3 \|f^{-1} \mathbf{1}_{B(0, M)}\|_{L^p(P)}.$$

This yields the assertion. □

**Remarks.**

- One can replace the assumption on  $A = \text{supp}(P)$  by a local “peakless” assumption, namely

$$\forall M, c > 0, \inf_{x \in B(0, M) \cap \text{supp}(P), 0 < \rho \leq c} \frac{\lambda_d(B(x, \rho) \cap B(0, M) \cap \text{supp}(P))}{\lambda_d(B(x, \rho))} > 0$$

which can be satisfied by many subsets which are not finite unions of closed convex sets (e.g. if  $A = C^c$ ,  $C$  convex set with a non-empty interior).

• Let  $s \in (0, d+r)$  (and  $P^a \neq 0$ ). We know by Lemma 1 that  $\psi_b \in L_{loc}^{\frac{s}{d+r}}(P)$ . Then, owing to the domination property (on a fixed ball  $B(0; M)$ ,  $M > 0$ ) induced by Inequality (3.7), we may apply the (reverse) Fatou's lemma for sequences of dominated non-negative functions (applied on every ball  $B(0; M)$ ) which yields

$$\sup_{M>0} \limsup_{n \rightarrow \infty} n^{s/d} \int_{B(0,M)} d(x, \alpha_n)^s dP(x) \leq \int \limsup_{n \rightarrow \infty} n^{s/d} d(x, \alpha_n)^s dP(x).$$

Then, provided  $P = P^a$ , Inequality (3.8) implies

$$\int \limsup_{n \rightarrow \infty} n^{s/d} d(x, \alpha_n)^s dP(x) \leq C_3(r, b)^s \int f^{-s/(d+r)} dP.$$

Consequently, the “minimal” expected condition  $\int f^{-s/(d+r)} dP < +\infty$  implies

$$\sup_{M>0} \limsup_{n \rightarrow \infty} n^{s/d} \int_{B(0,M)} d(x, \alpha_n)^s dP(x) \leq C_3(r, b)^s \int f^{-s/(d+r)} dP < +\infty.$$

• Now, coming back to Assumption (3.2), i.e.  $\psi_b \in L^{\frac{s}{d+r}}(P)$ , one derives likewise a slightly more precise result than that stated in Theorem 2:

$$\limsup_{n \rightarrow \infty} n^{s/d} \int d(x, \alpha_n)^s dP(x) \leq \int \limsup_{n \rightarrow \infty} n^{s/d} d(x, \alpha_n)^s dP(x) \leq C_3(r, b)^s \int f^{-s/(d+r)} dP < +\infty.$$

This emphasizes that, if one wishes to establish the conclusion of Theorem 2 without the global integrality Assumption (3.2) on  $\psi_b$ , one needs to show directly that

$$\sup_{M>0} \limsup_{n \rightarrow \infty} n^{s/d} \int_{B(0,M)} d(x, \alpha_n)^s dP(x) = \limsup_{n \rightarrow \infty} n^{s/d} \int d(x, \alpha_n)^s dP(x).$$

Although of little practical help a clue in that direction could be to notice that, when  $s = r$ , it follows from Theorem 4.5 in [4] and (2.3) that

$$\begin{aligned} \sup_{M>0} \lim_{n \rightarrow \infty} n^{r/d} \int_{B(0,M)} d(x, \alpha_n)^r dP(x) &= \sup_{M>0} Q_r(P) P_r(B(0, M)) = Q_r(P) \\ &= \lim_{n \rightarrow \infty} n^{r/d} \int d(x, \alpha_n)^r dP(x). \end{aligned}$$

The following corollaries deal with distributions first with bounded and then mainly with unbounded supports.

**Corollary 1** (compactly supported distributions). *Assume that  $\text{supp}(P)$  is compact and  $P^a \neq 0$ .*

(a) *For every  $s \in (0, d+r)$ , the assertion (3.3) of Theorem 2 holds true.*

(b) *Let  $s > d+r$ . Assume that  $\text{supp}(P)$  is a finite union of compact convex sets,  $P = P^a = f \cdot \lambda_d$  and*

$$\int f^{-s/(d+r)} dP < +\infty.$$

*Then the assertion (3.3) of Theorem 2 holds true for  $s$  (and every  $s' \in (0, s]$ ). In particular:*

– *if  $f \geq \varepsilon > 0$   $\lambda_d$ -a.s. on  $\text{supp}(P)$ , then (3.3) holds for every  $s \in (0, \infty)$ ;*

– *if  $\int_{\{f>0\}} f^{-(1+\delta)} dP = \int_{\{f>0\}} f^{-\delta} d\lambda_d < +\infty$  for some  $\delta > 0$ , then (3.3) holds with  $s = d+r$ .*

*Proof.* (a) This is an immediate consequence of Theorem 2 since Lemma 1 implies that  $\psi_b^{\frac{s}{d+r}}$  is integrable. (b) follows from Lemma 2 and Theorem 2.  $\square$

For distributions with unbounded support we provide the following condition.

**Corollary 2.** (Distributions with unbounded supports.) *Let  $s \in (0, d+r)$ . Assume  $P^a \neq 0$ ,  $\int \|x\|^{r+\delta} dP(x) < +\infty$  for some  $\delta > 0$  and*

$$\int_{B(0, M)^c} \left( \sup_{t \leq 2b\|x\|} \frac{\lambda_d(B(x, t))}{P(B(x, t))} \right)^{s/(d+r)} dP(x) < +\infty \quad (3.9)$$

for some  $M, b \in (0, \infty)$ . Then the assertion (3.3) of Theorem 2 holds true.

*Proof.* Let  $x_0 \in \text{supp}(P)$ . Then  $d(x_0, \alpha_n) \rightarrow 0$  (see [4]). For  $\|x\| > N := \|x_0\| + \sup_{n \geq 1} d(x_0, \alpha_n)$ , we have  $d(x, \alpha_n) \leq 2\|x\|$  for every  $n \geq 1$  and thus

$$\psi_b(x) \leq \sup_{t \leq 2b\|x\|} \frac{\lambda_d(B(x, t))}{P(B(x, t))}$$

for every  $x \in B(0, N)^c, b > 0$ . Then  $\psi_b \in L^{\frac{s}{d+r}}$  by Lemma 1 and the conclusion follows from Theorem 2.  $\square$

For distributions with radial tails we obtain a condition which is very close to the minimal condition  $\int f^{-s/(d+r)} dP < +\infty$ .

**Corollary 3** (density with radial tails). (a) *Assume  $P = f \cdot \lambda_d$ ,  $\int \|x\|^{r+\delta} dP(x) < +\infty$  for some  $\delta > 0$  and  $f = h(\|\cdot\|_0)$  on  $B_{\|\cdot\|_0}(0, R)^c$  with  $h : (R, \infty) \rightarrow \mathbb{R}_+$  non-increasing for some  $R \in \mathbb{R}_+$  and  $\|\cdot\|_0$  any norm on  $\mathbb{R}^d$ . Let  $s \in (r, d+r)$ . If*

$$\int f(cx)^{-s/(d+r)} dP(x) < +\infty \quad (3.10)$$

for some  $c > 1$ , then Assertion (3.3) of Theorem 2 holds true.

(b) *Assume  $d = 1$  and  $s \in (0, 1+r)$ . Assume  $P = f \cdot \lambda$ ,  $\int |x|^{r+\delta} dP(x) < +\infty$  for some  $\delta > 0$ . If  $\text{supp}(P) \subset [R_0, \infty)$  for some  $R_0 \in \mathbb{R}$  and  $f|_{(R_0, \infty)}$  is non-increasing for some  $R'_0 \geq R_0$ . Assume furthermore (3.10) for some  $c > 1$ . Then the assertion (3.3) of Theorem 2 holds true.*

**Remark.** Note that if Assumption (3.10) holds for a real number  $c > 1$  then it holds for every  $c' \in (1, c]$ .

*Proof.* (a) We may assume without loss of generality that  $\|\cdot\| = \|\cdot\|_0$ . For  $b \in (0, 1/2)$ , let  $M = M(b) = N/(1-2b)$ . Then for  $x \in B(0, M)^c, t \leq 2b\|x\|$  and  $z \in B(0, t)$ , we have

$$\|x+z\| \geq \|x\| - \|z\| \geq \|x\|(1-2b) > N.$$

Consequently,

$$\begin{aligned} P(B(x, t)) &= \int_{B(0, t)} f(z+x) d\lambda_d(z) = \int_{B(0, t)} h(\|z+x\|) d\lambda_d(z) \\ &\geq h(\|x\| + t) \lambda_d(B(0, t)) \\ &\geq h(\|x\|(1+2b)) \lambda_d(B(0, t)) \end{aligned}$$

for every  $x \in B(0, M)^c$ ,  $t \leq 2b\|x\|$ . Hence

$$\sup_{t \leq 2b\|x\|} \frac{\lambda_d(B(x, t))}{P(B(x, t))} \leq \frac{1}{h(\|x\|(1+2b))} = \frac{1}{f((1+2b)x)}$$

for every  $x \in B(0, M)^c$  and Corollary 2 gives the assertion. Item (b) follows similarly.  $\square$

Corollary 3 provides an efficient criterion for many families of multi-dimensional distributions since many parametric families of distributions do have radial tails. Nevertheless, it is also natural to provide a criterion which does not require such a radial assumption. This is the aim of the next Corollary (which can also treat successfully the most usual multi-dimensional distributions).

**Corollary 4.** *Assume  $P = f \cdot \lambda_d$  and  $\int \|x\|^{r+\delta} dP(x) < +\infty$  for some  $\delta > 0$ . Furthermore, assume that  $\text{supp}(P)$  has no peak i.e.*

$$\kappa_f := \inf_{x \in \text{supp}(P), \rho > 0} \frac{\lambda_d(\text{supp}(P) \cap B(x, \rho))}{\lambda_d(B(x, \rho))} > 0 \quad (3.11)$$

and that  $f$  satisfies the local growth control assumption: there exist real numbers  $\varepsilon \geq 0$ ,  $\eta \in (0, \frac{1}{2})$ ,  $M, C > 0$  such that

$$\forall x, y \in \text{supp}(P), \|x\| \geq M, \|y - x\| \leq 2\eta\|x\| \implies f(y) \geq C(f(x))^{1+\varepsilon}. \quad (3.12)$$

Then, for every  $s \in (0, \frac{d+r}{1+\varepsilon})$  such that  $\int_{\mathbb{R}^d} \frac{dP(x)}{(f(x))^{\frac{s(1+\varepsilon)}{d+r}}} < +\infty$  (if any), the assertion (3.3) of Theorem 2 holds true. In particular, if (3.12) holds either for  $\varepsilon = 0$  or for every  $\varepsilon \in (0, \underline{\varepsilon}]$  ( $\underline{\varepsilon} > 0$ ), and if

$$\forall s \in (0, d+r), \quad \int_{\mathbb{R}^d} \frac{dP(x)}{(f(x))^{\frac{s}{d+r}}} = \int_{\{f>0\}} (f(x))^{1-\frac{s}{d+r}} d\lambda_d(x) < +\infty \quad (3.13)$$

then (3.3) holds for every  $s \in (0, d+r)$ .

Note that (if  $\lambda_d(\text{supp}(P)) = +\infty$ ) Assumption (3.11) is e.g. satisfied by any finite intersection of half-spaces, the typical example being  $\mathbb{R}_+^d$ . Furthermore, a careful reading of the proof below shows that this assumption can be slightly relaxed into: there exists a real  $c > 0$  such that

$$\kappa'_f := \inf_{x \in \text{supp}(P)} \left\{ \frac{\lambda_d(\text{supp}(P) \cap B(x, \rho))}{\lambda_d(B(x, \rho))}, 0 < \rho \leq c\|x\| \right\} > 0.$$

*Proof.* Let  $x_0 \in \text{supp}(P)$ . Then, for large enough  $n$ ,  $\alpha_n \cap B(x_0, 1) \neq \emptyset$ . Hence

$$d(x, \alpha_n) \leq d(x, \alpha_n \cap B(x_0, 1)) \leq \|x - x_0\| + 1 \leq \|x\| + \|x_0\| + 1$$

so that  $d(x, \alpha_n) \leq 2\|x\|$  provided  $\|x\| \geq \|x_0\| + 1$ . We will assume from now on (without loss of generality) that  $M \geq \|x_0\| + 1$  in Assumption (3.12).

Then for every  $x \in \text{supp}(P)$ ,  $\|x\| \geq M$  and every  $\rho \in [0, 2\eta\|x\|]$

$$\begin{aligned} P(B(x, \rho)) &= \int_{B(0, \rho)} f(x+y) \mathbf{1}_{\{\text{supp}(P)\}}(x+y) dy \\ &\geq C(f(x))^{1+\varepsilon} \lambda_d(B(x, \rho) \cap \text{supp}(P)) \\ &\geq C\kappa_f(f(x))^{1+\varepsilon} \lambda_d(B(x, \rho)) \end{aligned}$$

where we used (3.11), (3.12) and  $\|x\| \leq 2\eta\|x\|$ . Then

$$\sup_{\rho \leq 2\eta\|x\|} \frac{\lambda_d(B(x, \rho))}{P(B(x, \rho))} \leq \frac{1}{C\kappa_f} \frac{1}{(f(x))^{1+\varepsilon}}.$$

Finally one concludes by Corollary 2 once noticed that

$$\int_{\|x\| \geq M} \sup_{\rho \leq 2\eta\|x\|} \left( \frac{\lambda_d(B(x, \rho))}{P(B(x, \rho))} \right)^{\frac{s}{d+r}} dP(x) \leq \frac{1}{(C\kappa_f)^{\frac{s}{d+r}}} \int_{\|x\| \geq M} \frac{dP(x)}{(f(x))^{\frac{s(1+\varepsilon)}{d+r}}} < +\infty. \quad \square$$

**Remark.** Note that the moment assumption (3.13) follows from the more natural moment assumption

$$\forall s \in (0, d+r), \quad \exists M \geq 0, \exists \delta > 0 \quad \text{such that} \quad \int_{\|x\| \geq M} \|x\|^{\frac{ds}{d+r-s} + \delta} f(x) dx < +\infty$$

which is of course satisfied by all distributions having polynomial moment of any order. This follows from a standard application of Hölder inequality (see the comments that follow the statement (2.3), of the Zador Theorem in Sect. 2), once noticed that  $\int_{\|x\| \leq M} f^{\frac{d+r-s}{d+r}} d\lambda_d < +\infty$  for every  $M > 0$  since  $0 < \frac{d+r-s}{d+r} \leq 1$  and  $f$  is a probability density function.

As concerns practical applications, the Gaussian distributions are probably the most important class of distributions satisfying Theorem 2.

**Example 1 •** Let  $P = N(0, \Sigma)$  be the  $d$ -dimensional centred normal distribution with positive definite covariance matrix  $\Sigma$ . Then  $P$  satisfies condition (3.10) from Corollary 3(a) with  $h(y) = ((2\pi)^d \det \Sigma)^{-1/2} e^{-y^2/2}$ ,  $\|x\|_0 := \|\Sigma^{-1/2}x\|$  where  $\|\cdot\|$  stands for the Euclidean norm on  $\mathbb{R}^d$  and  $M = 0$ . For  $s \in (0, d+r)$ , choose  $c \in (1, \sqrt{(d+r)/s})$ .

• For  $s \in (0, d+r)$ , the assumptions of Corollary 3(a) are satisfied by the *hyper-exponential distributions* defined by

$$f(x) \propto \exp(-a\|x\|^b), \quad x \in \mathbb{R}^d, \quad a, b > 0,$$

where  $\|\cdot\|$  denotes any norm on  $\mathbb{R}^d$  (possibly different of the underlying norm). Set  $h(u) \propto \exp(-au^b)$ . Note that if  $d = 1$ , then the normalizing positive real constant is given by  $\kappa_{a,b} = \frac{ba^{1/b}}{\Gamma(1/b)}$ . In fact the whole family of distributions

$$f(x) \propto \|x\|^c \exp(-a\|x\|^b), \quad x \in \mathbb{R}^d, \quad a, b > 0, \quad c > -d,$$

satisfies the assumptions of Corollary 3(a). In particular, the family includes the (scalar) *double Gamma distributions* on the real line where

$$f(x) = \frac{a^c}{2\Gamma(c)} |x|^{c-1} e^{-a|x|}, \quad x \in \mathbb{R}, \quad a, c > 0.$$

• As concerns scalar distributions let us first mention the *Gamma distributions*

$$f(x) = \frac{a^b}{\Gamma(b)} x^{b-1} e^{-ax} \mathbf{1}_{(0, \infty)}(x), \quad x \in \mathbb{R}, \quad a, b > 0,$$

for which the assumptions of Corollary 3(b) are satisfied for every  $s \in (1, 1+r)$ . This holds as well for the *Weibull distributions*

$$f(x) = bx^{b-1} \exp(-x)^b \mathbf{1}_{(0, \infty)}(x), \quad x \in \mathbb{R}, \quad b > 0,$$

the *lognormal distributions*

$$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2\sigma^2}(\log x - a)^2\right) \mathbf{1}_{(0,\infty)}(x), \quad x \in \mathbb{R}, \quad \sigma > 0, \quad a \in \mathbb{R},$$

• The *logistic distribution*

$$f(x) = \frac{e^x}{(1 + e^x)^2}, \quad x \in \mathbb{R},$$

satisfies Corollary 3(a) and so do the symmetric  $\rho$ -stable distributions with  $\rho \in (0, 2)$  provided  $r < \rho$  and  $s \in (0, \frac{\rho}{(1+\rho)}(1+r))$ .

• For the *Pareto distributions* where

$$f(x) = bx^{-(b+1)} \mathbf{1}_{(1,\infty)}(x), \quad x \in \mathbb{R}, \quad b > 0$$

the assumptions Corollary 3(b) are satisfied provided that  $r < b$  and  $s < b(1+r)/(b+1)$ .

• In fact all the examples of distributions provided so far also fulfil the criterion proposed in Corollary 4. There are natural classes of distributions whose density is not monotone which can only be handled by the criterion in Corollary 4. We will discuss now such an example.

Let  $\mu = g.\lambda_1$  be an absolutely continuous probability distribution on  $[0, 1]$  satisfying

$$0 < C_- \leq g \leq C_+ < +\infty.$$

Let  $(Y_n)_{n \geq 1}$  denote an i.i.d. sequence of  $\mu$ -distributed random variables and let  $N$  denote a Poisson random variable with parameter  $\lambda > 1$ , independent of  $(Y_n)_{n \geq 1}$ . Then set  $X = N + Y_N$  and  $P = \mathbb{P}_X$ . One easily checks that

$$\mathbb{P}_X = f.\lambda_1 \quad \text{with} \quad f(x) = e^{-\lambda} \frac{\lambda^{\lfloor x \rfloor}}{\lfloor x \rfloor!} g(x - \lfloor x \rfloor) \mathbf{1}_{\mathbb{R}_+}(x).$$

Let  $\varepsilon > 0$  and  $\eta \in (0, (1/2) \wedge \varepsilon)$  and  $x \in \mathbb{R}_+$ ,  $y \in [x(1-\eta), x(1+\eta)]$ . As soon as  $n \geq \lfloor \lambda \rfloor + 1$ , the sequence  $\frac{\lambda^n}{n!}$  is decreasing (to 0). Hence, if  $x \geq (1+\eta)\lfloor \lambda \rfloor + 1$

$$\begin{aligned} \frac{f(y)}{(f(x))^{1+\varepsilon}} &\geq \frac{C_-}{(C_+)^{1+\varepsilon}} e^{\lambda\varepsilon} \frac{\lambda^{\lfloor y \rfloor} (\lfloor x \rfloor!)^{1+\varepsilon}}{\lfloor y \rfloor! \lambda^{(1+\varepsilon)\lfloor x \rfloor}} \\ &\geq \frac{C_-}{(C_+)^{1+\varepsilon}} e^{\lambda\varepsilon} \lambda^{-1-x(\varepsilon-\eta)} \frac{([\lfloor x \rfloor!])^{1+\varepsilon}}{(1 + \lfloor x(1+\eta) \rfloor)!}. \end{aligned}$$

One concludes using the Stirling formula that

$$\liminf_{x \rightarrow +\infty} \inf_{y \in (x(1-\eta), x(1+\eta))} \frac{f(y)}{(f(x))^{1+\varepsilon}} = +\infty.$$

On the other hand, Assumption (3.13 of Cor. 4) is fulfilled since  $X$  has finite moment of any polynomial order (see the remark following Cor. 4). Note that when the density function  $g$  is not non-decreasing  $f$  cannot be a non-decreasing function so that Corollary 3 does not apply.

The exact optimality assumption made on the sequence  $(\alpha_n)$  in Theorem 2 (and Cor. 1) is critical to get the optimal rate  $n^{-s/d}$ . This is illustrated by the following counter-example.

**(Counter-)example 2.** Let  $P = U([0, 1])$  and for  $n \geq 2$  and  $\vartheta \in (0, \infty)$  set

$$\alpha_n = \alpha_n(\vartheta) := \left\{ \frac{1}{2n^\vartheta} \right\} \cup \left\{ \frac{1}{n^\vartheta} + \left(1 - \frac{1}{n^\vartheta}\right) \frac{2(k-1)-1}{2(n-1)} : k = 2, \dots, n \right\}.$$

Let  $r \in (0, \infty)$  and assume  $\vartheta > \frac{r}{(1+r)}$ . Using a non-Voronoi partition gives the upper estimate

$$\begin{aligned} \mathbb{E}|X - \hat{X}^{\alpha_n}|^r &\leq \int \mathbf{1}_{[0, n^{-\vartheta}]}(x) \left| x - \frac{1}{2n^\vartheta} \right|^r dx \\ &\quad + \sum_{k=2}^n \int \mathbf{1}_{\left[\frac{1}{n^\vartheta} + (1 - \frac{1}{n^\vartheta}) \frac{k-2}{n-1}, \frac{1}{n^\vartheta} + (1 - \frac{1}{n^\vartheta}) \frac{k-1}{n-1}\right]}(x) \left| x - \left( \frac{1}{n^\vartheta} + \left(1 - \frac{1}{n^\vartheta}\right) \frac{2(k-1)-1}{2(n-1)} \right) \right|^r dx \\ &= \frac{2}{r+1} \left( \frac{1}{2n^\vartheta} \right)^{r+1} + \left(1 - \frac{1}{n^\vartheta}\right)^{r+1} \frac{1}{(r+1)2^r(n-1)^r}. \end{aligned}$$

Hence

$$\begin{aligned} \limsup_{n \rightarrow \infty} n^r \mathbb{E}|X - \hat{X}^{\alpha_n}|^r &\leq \limsup_{n \rightarrow \infty} \left( \frac{1}{2^r(r+1)} \frac{1}{n^{(r+1)\vartheta-r}} + \frac{1}{(r+1)2^r} \right) \\ &= \frac{1}{2^r(r+1)} = J_{r,1} = Q_r(P) \end{aligned}$$

so that in fact

$$\lim_{n \rightarrow \infty} n^r \mathbb{E}|X - \hat{X}^{\alpha_n}|^r = Q_r(P).$$

It follows that the sequence  $(\alpha_n(\vartheta))_n$  is an asymptotically  $L^r$ -optimal  $n$ -quantizer for every  $\vartheta \in (r/(r+1), \infty)$ . Now, let  $s > r$  and  $\vartheta \in (r/(r+1), s/(s+1))$ . Then

$$\mathbb{E}|X - \hat{X}^{\alpha_n}|^s \geq \int_0^{1/2n^\vartheta} \left| x - \frac{1}{2n^\vartheta} \right|^s dx = \frac{1}{2^{s+1}(s+1)} n^{-\vartheta(s+1)}$$

so that

$$n^s \mathbb{E}|X - \hat{X}^{\alpha_n}|^s \geq \frac{1}{2^{s+1}(s+1)} n^{s-\vartheta(s+1)}.$$

Consequently,

$$\lim_{n \rightarrow \infty} n^s \mathbb{E}|X - \hat{X}^{\alpha_n}|^s = \infty.$$

#### 4. THE CRITICAL AND SUPER-CRITICAL CASES

In this section, we investigate the upper bound for the rate of convergence when  $s \geq d+r$ .

**Proposition 1.** *Let  $r \in (0, \infty)$ . Assume  $P^a \neq 0$  and  $\int \|x\|^{r+\delta} dP(x) < +\infty$  for some  $\delta > 0$ . For every  $n \geq 1$ , let  $\alpha_n$  be an  $L^r$ -optimal  $n$ -quantizer for  $P$ .*

(a) **THE CRITICAL CASE.**  $s = d+r$ : *Assume there is a real  $\underline{\vartheta} > 0$  such that for every  $\vartheta \in (0, \underline{\vartheta} \wedge s)$ , there exists a real number  $M \geq 0$  such that*

$$\int_{B(0;M)^c} \left( \sup_{r \leq 2b\|x\|} \frac{\lambda_d(B(x,r))}{P(B(x,r))} \right)^{1-\frac{\vartheta}{d+r}} \|x\|^\vartheta dP(x) < +\infty. \quad (4.14)$$

Then,

$$\forall \varepsilon \in (0, 1+r/d), \quad \limsup_n n^{1+\frac{\varepsilon}{d}-\varepsilon} \int_{\mathbb{R}^d} (d(x, \alpha_n))^{d+r} dP(x) < +\infty.$$

(b) THE SUPER-CRITICAL CASE.  $s > d + r$ : Assume there is a real  $\vartheta \in (s - (d + r), s)$  and a real number  $M \geq 0$  such that

$$\int_{B(0;M)^c} \left( \sup_{r \leq 2b\|x\|} \frac{\lambda_d(B(x,r))}{P(B(x,r))} \right)^{\frac{s-\vartheta}{d+r}} \|x\|^\vartheta dP(x) < +\infty. \quad (4.15)$$

Then,

$$\sup_n n^{\frac{s-\vartheta}{d}} \int_{\mathbb{R}^d} (d(x, \alpha_n))^s dP(x) < +\infty. \quad (4.16)$$

*Proof.* (b) It follows from (3.7) that, for every  $n \geq 1$ ,  $b \in (0, 1/2)$ ,

$$\forall x \in \mathbb{R}^d, \quad (d(x, \alpha_n))^{d+r} \leq C_b \psi_b(x) n^{-(1+\frac{r}{d})}.$$

On the other hand,  $x_0 \in \text{supp}(P)$  being fixed, for large enough  $n$ ,  $\alpha_n \cap B(x_0, 1) \neq \emptyset$  so that

$$\forall x \in \mathbb{R}^d, \quad d(x, \alpha_n) \leq (1 + |x - x_0|) \mathbf{1}_{\{|x-x_0| \geq 1\}} + 2 \mathbf{1}_{\{|x-x_0| < 1\}} \leq C_0(|x| \vee 1).$$

Let  $\psi_b$  be the maximal function associated to  $(\alpha_n)_{n \geq 1}$  and  $b \in (0, 1/2)$ . Consequently

$$\begin{aligned} \forall x \in \mathbb{R}^d, \quad (d(x, \alpha_n))^s &\leq (d(x, \alpha_n))^{s-\vartheta} (C_0(|x| \vee 1))^\vartheta \\ &\leq C'_b n^{\frac{\vartheta-s}{d}} (\psi_b(x))^{\frac{s-\vartheta}{d+r}} (\|x\| \vee 1)^\vartheta. \end{aligned}$$

It follows from Lemma 1 that, for every real number  $M \geq 0$  and  $b > 0$ ,

$$\int_{B(0;M)} (\psi_b(x))^{\frac{s-\vartheta}{d+r}} \|x\|^\vartheta dP(x) < +\infty.$$

Combined with Assumption (4.15), this clearly implies  $\psi_b^{\frac{s-\vartheta}{d+r}} \in L^1(P)$ .

(a) The above computations are still valid when  $s = d + r$ . So one concludes by considering  $\vartheta$  arbitrarily close to 0 which is precisely made possible by Assumption (4.14).  $\square$

From the above proposition one easily derives some corollaries in the formerly mentioned settings. We give them in details for the super-critical case. The adaptation to the critical case is straightforward.

**Corollary 5.** Assume the global assumption on  $P$  in Proposition 1 holds and that  $P = f \cdot \lambda_d$ .

Let  $\vartheta \in (s - (d + r), s)$ .

(a) If  $f$  is radial outside a compact subset of  $\mathbb{R}^d$  (as defined in Cor. 1) and if

$$\int f(cx)^{-\frac{s-\vartheta}{d+r}} \|x\|^\vartheta dP(x) < +\infty \text{ for some real } c > 1,$$

then Assumption (4.15) is fulfilled.

(b) If  $f$  satisfies Assumption (3.11) and Assumption (3.12) for some parameter  $\varepsilon \geq 0$  and if

$$\int f(x)^{-\frac{(s-\vartheta)(1+\varepsilon)}{d+r}} \|x\|^\vartheta dP(x) < +\infty$$

then Assumption (4.15) is fulfilled.

The above assumptions are clearly fulfilled by the normal distributions (and in fact most distributions mentioned in Ex. 1). Numerical experiments seem to suggest that the critical rate  $n^{1+\frac{r}{d}}$  cannot be improved for  $P$  with unbounded support, that is,

$$\forall s > 0, \quad \lim_n n^{1+\frac{r}{d}} \int_{\mathbb{R}^d} d(x, \alpha_n)^s dP(x) = +\infty.$$

### 5. A SHARP RATE FOR LIPSCHITZ DENSITIES ON COMPACT INTERVALS

**Proposition 2.** *Let  $P = f.\lambda_1$  where  $f : [a, b] \rightarrow \mathbb{R}_+$  is a Lipschitz continuous probability density function, bounded away from 0 on  $[a, b]$ . Let  $r > 0$ . Let  $(\alpha_n)_{n \geq 1}$  be a sequence of  $L^r$ -stationary and asymptotically  $L^r$ -optimal  $[a, b]$ -valued  $n$ -quantizers. Then for every  $s > 0$ ,*

$$\lim_n \left( n^s \int_a^b \min_{\alpha_n^i \in \alpha_n} |x - \alpha_n^i|^s dP(x) \right) = Q_{r,s}(P).$$

Note that we do not assume *a priori* that the  $n$ -quantizers are  $L^r$ -optimal but only stationary and asymptotically  $L^r$ -optimal.

*Proof.* It is possible in this scalar setting to number the elements of a  $n$ -quantizer with respect to the natural order on the real line. Furthermore, one may assume for large enough  $n$  that  $\alpha_n \subset (a, b)$  and  $\text{card}\alpha_n = n$ . Namely, we set

$$\alpha_n = \{\alpha_n^1, \dots, \alpha_n^n\}, \quad a < \alpha_n^1 < \alpha_n^2 < \dots < \alpha_n^{n-1} < \alpha_n^n < b.$$

We also set  $\alpha_n^0 = a$  and  $\alpha_n^{n+1} = b$  for convenience. For every  $n \geq 1$ , set for every  $x \in [a, b]$

$$\varphi_{r,n}(x) = n \sum_{k=1}^n \Delta\alpha_n^k \mathbf{1}_{V_k(\alpha_n)}(x) \tag{5.1}$$

where  $V_k(\alpha_n)$  denotes the Voronoi cell of  $\alpha_n^k$  and  $\Delta\alpha_n^k = \alpha_n^k - \alpha_n^{k-1}$ ,  $k \in \{1, \dots, n+1\}$ . We know from [5] (see the proof of Th. 4(a), p. 1101) that  $h$  being Lipschitz continuous and bounded away from 0 on  $[a, b]$ , there is a positive real constant  $C_{a,b,f}$  such that

$$\max_{1 \leq k \leq n+1} \Delta\alpha_n^k \leq C_{a,b,f} \min_{1 \leq k \leq n+1} \Delta\alpha_n^k. \tag{5.2}$$

A proof of this last inequality can be found in the appendix. Set  $P_{r,n} = \frac{1}{n} \sum_{k=1}^n \delta_{\alpha_n^k}$  the empirical measure associated to the  $n$ -quantizer  $\alpha_n$ . By the empirical measure theorem (recalled in (2.5)), it follows that for large enough  $n$ ,

$$\forall a', b' \in [a, b], \quad \frac{\text{card}\{i : \alpha_n^i \in [a', b']\}}{n} \rightarrow P_r([a', b']) := c_{f, \frac{1}{r+1}} \int_{a'}^{b'} f^{\frac{1}{r+1}} d\lambda_1$$

uniformly with respect to  $(a', b') \in [a, b]^2$  (where  $c_{g,\vartheta} = (\int_a^b g^\vartheta d\lambda_1)^{-1}$ ). Combining this convergence with (5.2) and using that  $f$  is bounded away from 0 on  $[a, b]$  implies that there is a real constant  $C'_{a,b,f}$  such that, for every  $n \geq 1$  and every  $a', b' \in [a, b]$

$$\frac{b' - a'}{C'_{a,b,f} n} \leq \min_{1 \leq k \leq n+1, \alpha_n^k \in [a', b']} \Delta\alpha_n^k \leq \max_{1 \leq k \leq n+1, \alpha_n^k \in [a', b']} \Delta\alpha_n^k \leq \frac{C'_{a,b,f}}{n} (b' - a').$$

A straightforward application of the Arzela-Ascoli Theorem yields that the sequence  $(\varphi_{r,n})_{n \geq 1}$  is relatively compact for the topology of uniform convergence and that all its limiting functions are continuous. Let  $\varphi_r$

denote such a limiting function. On the one hand, for every  $a', b' \in [a, b]$ ,

$$\int_{a'}^{b'} \varphi_{r,n} dP_{r,n} = \sum_{k=1}^n \Delta \alpha_n^k \mathbf{1}_{\{\alpha_n^k \in [a', b']\}} \longrightarrow b' - a' \quad \text{as } n \rightarrow \infty,$$

i.e.  $\varphi_{r,n} \cdot P_{r,n}$  converges weakly to the Lebesgue measure  $\lambda_1$  over  $[a, b]$ . On the other hand  $\varphi_{r,n} \cdot P_{r,n}$  converges weakly to  $\varphi_r \cdot P_r = c_{f, \frac{1}{r+1}} \varphi_r f^{\frac{1}{r+1}} \cdot \lambda_1$  since  $\varphi_r$  is continuous. Consequently,  $\varphi_r$  satisfies

$$c_{f, \frac{1}{r+1}} \varphi_r f^{\frac{1}{r+1}} \cdot \lambda_1 = \lambda_1.$$

The function  $\varphi_r$  is uniquely determined by this equation so that

$$\varphi_{r,n} \xrightarrow{U} \varphi_r := \left( \int_a^b f^{\frac{1}{r+1}} d\lambda_1 \right) f^{-\frac{1}{r+1}} \quad \text{as } n \rightarrow \infty.$$

Now for every  $k = 2, \dots, n-1$ ,

$$\begin{aligned} n^s \left| \int_{V_k(\alpha_n^k)} |\alpha_n^k - x|^s f(x) dx - f(\alpha_n^k) \int_{V_k(\alpha_n^k)} |\alpha_n^k - x|^s dx \right| &\leq n^s [f]_{\text{Lip}} \int_{V_k(\alpha_n^k)} |\alpha_n^k - x|^{s+1} dx \\ &= [f]_{\text{Lip}} \frac{n^s}{s+2} \left( \left( \frac{\Delta \alpha_n^{k+1}}{2} \right)^{s+2} + \left( \frac{\Delta \alpha_n^k}{2} \right)^{s+2} \right) \\ &= \frac{[f]_{\text{Lip}}}{(s+2)2^{s+2}n^2} \left( (\varphi_{r,n}(\alpha_n^{k+1}))^{s+2} + (\varphi_{r,n}(\alpha_n^k))^{s+2} \right) \end{aligned} \quad (5.3)$$

and

$$\begin{aligned} n^s f(\alpha_n^k) \int_{V_k(\alpha_n^k)} |\alpha_n^k - x|^s dx &= \frac{f(\alpha_n^k)}{(s+2)2^{s+1}n} \left( (\varphi_{r,n}(\alpha_n^{k+1}))^{s+1} + (\varphi_{r,n}(\alpha_n^k))^{s+1} \right) \\ &= \frac{J_{1,s}}{2n} \left( f(\alpha_n^k) (\varphi_{r,n}(\alpha_n^k))^{s+1} + f(\alpha_n^{k+1}) (\varphi_{r,n}(\alpha_n^{k+1}))^{s+1} \right) + O\left(\frac{1}{n^2}\right) \end{aligned} \quad (5.4)$$

since

$$|f(\alpha_n^{k+1}) - f(\alpha_n^k)| (\varphi_{r,n}(\alpha_n^{k+1}))^{s+1} \leq [f]_{\text{Lip}} \Delta \alpha_n^{k+1} (\varphi_{r,n}(\alpha_n^{k+1}))^{s+1} = \frac{[f]_{\text{Lip}}}{n} (\varphi_{r,n}(\alpha_n^{k+1}))^{s+2} \leq C/n.$$

For  $k = 1$  and  $n$ , the equalities (5.3) and (5.4) are no longer true due to edge effects but both induced errors are  $O(1/n)$ . Consequently

$$\begin{aligned} \left| n^s \mathbb{E} |X - \widehat{X}^{\alpha_n}|^s - \frac{J_{1,s}}{2} \frac{1}{n} \sum_{k=1}^n \left( f(\alpha_n^k) (\varphi_{r,n}(\alpha_n^k))^{s+1} + f(\alpha_n^{k+1}) (\varphi_{r,n}(\alpha_n^{k+1}))^{s+1} \right) \right| \\ \leq C_{s,f} \frac{1}{n^2} \sum_{k=1}^n \left( (\varphi_{r,n}(\alpha_n^{k+1}))^{s+2} + (\varphi_{r,n}(\alpha_n^k))^{s+2} \right) + O\left(\frac{1}{n}\right) \\ \leq 2C_{s,f} \frac{1}{n} \int_a^b (\varphi_{r,n})^{s+2} dP_{r,n} + O\left(\frac{1}{n}\right) \\ = O\left(\frac{1}{n}\right). \end{aligned}$$

On the other hand

$$\begin{aligned} \frac{1}{n} \sum_{k=1}^n \left( f(\alpha_n^k) (\varphi_{r,n}(\alpha_n^k))^{s+1} + f(\alpha_n^{k+1}) (\varphi_{r,n}(\alpha_n^{k+1}))^{s+1} \right) &= \frac{2}{n} \int_a^b (\varphi_{r,n})^{s+1} f \, dP_{r,n} + O\left(\frac{1}{n}\right) \\ &\longrightarrow 2 \int_a^b (\varphi_r)^{s+1} f \, dP_r \quad \text{as } n \rightarrow \infty. \end{aligned}$$

Finally

$$\int_a^b (\varphi_r)^{s+1} f \, dP_r = \left( \int_a^b f^{\frac{1}{r+1}} \, d\lambda_1 \right)^{s+1} \int_a^b f^{1-\frac{s+1}{r+1}} f^{\frac{1}{r+1}} \, d\lambda_1 = \left( \int_a^b f^{\frac{1}{r+1}} \, d\lambda_1 \right)^{s+1} \int_a^b f^{1-\frac{s}{r+1}} \, d\lambda_1$$

which completes the proof.  $\square$

**Remark.** In fact, we proved a slightly more precise statement, namely that, for every  $r, s > 0$ ,

$$\mathbb{E}|X - \widehat{X}^{\alpha_n}|^s = \frac{Q_{r,s}(P)}{n^s} + O\left(\frac{1}{n^{s+1}}\right) \quad \text{as } n \rightarrow \infty.$$

## 6. APPLICATION TO CUBATURE FORMULAE FOR NUMERICAL INTEGRATION

### 6.1. Numerical integration on $\mathbb{R}^d$

Thanks to the optimal  $L^s$ -quantization rate provided by Theorem 2 for sequences of  $L^r$ -optimal quantizers, it is now possible to take advantage of the stationarity of quadratic quantizers for a wider class of functions.

Let  $f : \mathbb{R}^d \rightarrow \mathbb{R}$  be a twice differentiable function such that

$$\|D^2 f(x)\| \leq A(\|x\|^{d-\eta} + 1)$$

for some non-negative real constants  $A$  and  $B$  and  $\eta \in (0, d]$ . Let  $X$  be a r.v. with distribution  $P = \mathbb{P}_X$ . Assume that  $X \in L^{2+d}(\mathbb{P})$  and that for every sequence of optimal quadratic  $n$ -quantizers  $(\alpha_n)_{n \geq 1}$  the optimal  $L^s$ -quantization rate holds true for every  $s \in (0, 2 + d)$ . First, let  $\alpha$  be an optimal quadratic quantizer and  $\widehat{X}^\alpha$  a related Voronoi quantization of  $X$ . For every couple  $(p, q) \in (1, +\infty)^2$  of Hölder conjugate exponents, one has

$$|\mathbb{E} f(X) - \mathbb{E} f(\widehat{X}^\alpha)| \leq \frac{1}{2} \|D^2 f(\widehat{X}^\alpha)\|_p \|X - \widehat{X}^\alpha\|_{2q}^2. \quad (6.1)$$

Note that this error bound follows from the stationarity property of a quadratic quantization that is  $\mathbb{E}(X | \widehat{X}^\alpha) = \widehat{X}^\alpha$  since

$$|\mathbb{E} f(X) - \mathbb{E} f(\widehat{X}^\alpha)| = |\mathbb{E} f(X) - \mathbb{E} f(\widehat{X}^\alpha) - \mathbb{E}(Df(\widehat{X}^\alpha) \cdot (X - \widehat{X}^\alpha))| \leq \frac{1}{2} \mathbb{E}(\|D^2 f(\widehat{X}^\alpha)\| \|X - \widehat{X}^\alpha\|^2).$$

Now, set  $p = \frac{d+2}{d-\eta}$  and  $q = \frac{d+2}{2+\eta}$  in equation (6.1). Then  $(d-\eta)p = d+2$  and  $2q < d+2$  so that,

$$\|D^2 f(\widehat{X}^\alpha)\|_p \leq A'(\mathbb{E} \|\widehat{X}^\alpha\|^{d+2} + 1)^{\frac{1}{p}} \leq A'(\mathbb{E} \|X\|^{d+2} + 1)^{\frac{1}{p}} < +\infty$$

and

$$\|X - \widehat{X}^{\alpha_n}\|_{2q}^2 = O(n^{-\frac{2}{d}}).$$

Consequently the optimal rate

$$|\mathbb{E} f(X) - \mathbb{E}(f(\widehat{X}^{\alpha_n}))| = O(n^{-\frac{2}{d}})$$

for numerical integration by (optimal quadratic) quantization holds for a wide class of (twice differentiable) functions whose growth at infinity can be infinitely faster than quadratic (the quadratic case is obtained for  $q = 1$  and  $p = +\infty$  for functions having a bounded Hessian; then there is no need for the extended quantization rate).

## 6.2. Numerical integration on the Wiener space

Let

$$W = \sum_{n \geq 1} \sqrt{\lambda_n} \xi_n e_n^W$$

be the Karhunen-Loève ( $K$ - $L$ ) expansion of a standard Brownian motion  $(W_t)_{t \in [0, T]}$ .  $(e_n^W, \lambda_n)_{n \geq 1}$  is the  $K$ - $L$  (orthonormal) eigensystem of the (nonnegative trace) covariance operator of the Brownian motion  $C_W(s, t) = s \wedge t$ , indexed so that  $\lambda_n$  decreases (to 0). All these quantities do have explicit expressions (see [9, 12] among others). In that  $K$ - $L$  expansion,  $(\xi_n)_{n \geq 1}$  is then a sequence of i.i.d. random  $\mathcal{N}(0, 1)$ -distributed random variables. One considers a sequence of (scalar) product quantizations  $\widehat{W}^N$  of the Brownian motion  $W$  *i.e.*

$$\widehat{W}^N = \sum_{k=1}^{m_N} \sqrt{\lambda_k} \widehat{\xi}_k^{N_k} e_k^W$$

where for every  $N \geq 1$ ,  $N_1, \dots, N_k \geq 1$ ,  $1 \leq k \leq m_N$  and  $N_1 \times \dots \times N_{m_N} \leq N$ ,  $\widehat{\xi}_k^{N_k}$  is an  $L^2$ -optimal  $N_k$ -quantization of  $\xi_k$  (see [8, 9] for more details). Then denoting  $|\cdot|_{L_T^2}$  the quadratic norm on  $L^2([0, T], dt)$  one has for every  $s \in [2, 3)$ ,

$$\begin{aligned} \|W - \widehat{W}^N\|_s &= \mathbb{E} \left( |W - \widehat{W}|_{L_T^2}^s \right)^{\frac{1}{s}} \\ &= \left( \mathbb{E} |W - \widehat{W}|_{L_T^2}^2 \right)^{\frac{1}{s}} \\ &= \| |W - \widehat{W}|_{L_T^2}^2 \|_{\frac{s}{2}}^{\frac{1}{2}} \\ &= \left\| \sum_{k=1}^{m_N} \lambda_k (\xi_k - \widehat{\xi}_k^{N_k})^2 + \sum_{k \geq m_N+1} \lambda_k \xi_k^2 \right\|_{\frac{s}{2}}^{\frac{1}{2}} \\ &\leq \left( \sum_{k=1}^{m_N} \lambda_k \|(\xi_k - \widehat{\xi}_k^{N_k})\|_{\frac{s}{2}}^2 + \sum_{k \geq m_N+1} \lambda_k \|\xi_k\|_{\frac{s}{2}}^2 \right)^{\frac{1}{2}} \\ &= \left( \sum_{k=1}^{m_N} \lambda_k \|\xi_k - \widehat{\xi}_k^{N_k}\|_s^2 + \sum_{k \geq m_N+1} \lambda_k \|\xi_k\|_s^2 \right)^{\frac{1}{2}} \end{aligned}$$

where we used that  $\frac{s}{2} \geq 1$  so that  $\|\cdot\|_{\frac{s}{2}}$  is a norm. Consequently, using that the normal distribution satisfies Theorem 2 (see Ex. 1), for every  $s \in [2, 3)$  there exists a real constant  $C_s$  such that

$$\|W - \widehat{W}^N\|_2 \leq \|W - \widehat{W}^N\|_s \leq C_s \left( \sum_{k=1}^{m_N} \frac{\lambda_k}{N_k^2} + \sum_{k \geq m_N+1} \lambda_k \right)^{\frac{1}{2}} = C_s \|W - \widehat{W}^N\|_2. \quad (6.2)$$

Consequently, if furthermore  $(\widehat{W}^N)_{N \geq 1}$  is a rate optimal sequence for quadratic quantization, then

$$\|W - \widehat{W}^N\|_s \asymp \|W - \widehat{W}^N\|_2 \sim c_W (\log N)^{-\frac{1}{2}}. \quad (6.3)$$

Note that the result clearly holds for  $s \in (0, 2]$  (simply using the monotony of the  $L^s$ -norm). Furthermore, except for the final rate  $O((\log N)^{-\frac{1}{2}})$  for which we refer to [9], the Brownian motion plays no specific rôle among Gaussian processes: the inequalities in (6.2) extend to the product quantization of any Gaussian process  $(X_t)_{t \in [0, T]}$ .

In fact, as emphasized in [9], optimal scalar product quantization is rate optimal but it is also possible to produce some rate optimal sequences based on  $d$ -dimensional block product quantizations. In fact it is shown in [9] that as  $d$  grows one can obtain the sharp convergence rate of high-resolution quantization (it is even possible to choose this dimension  $d = d(N)$  as a function of  $N$  to achieve this sharp rate). From a numerical viewpoint producing these  $d$ -dimensional marginal blocks is more demanding than producing scalar optimal quantizations, but some issues are in progress on that topic.

The point here is that, if one considers some  $d$ -dimensional marginal blocks to produce a rate optimal sequence of quantizations of the Brownian motion, such a sequence will satisfy (6.3) for every  $s \in (0, d + 2)$  instead of  $(0, 3)$ . This seems to be an interesting feature of  $d$ -dimensional block quantization.

The practical application to numerical integration on the Wiener space is formally similar to that on  $\mathbb{R}^d$  except that one considers some functionals  $F$  on  $(L^2([0, T], dt), |\cdot|_{L^2_T})$  instead of functions  $f$  defined on  $\mathbb{R}^d$ . Then, one approximates  $\mathbb{E}F(W)$  by  $\mathbb{E}F(\widehat{W}^N)$  having in mind that a closed formula exists as well for the distribution of  $\widehat{W}^N$  (*the weights*, see [12]) which makes the computation of  $\mathbb{E}F(\widehat{W}^N)$  fully tractable (with negligible cost). Then, if the Hessian  $D^2F$  of  $F$  satisfies

$$\forall \omega \in L^2([0, T], dt), \quad \|D^2F(\omega)\| \leq A(|\omega|_{L^2_T}^r + 1)$$

for some positive real constant  $A$ ,  $r > 0$  similar computations as those carried out in the  $d$ -dimensional case implies that

$$|\mathbb{E}F(W) - \mathbb{E}F(\widehat{W}^N)| \leq \frac{C_{F,W}}{\log N}.$$

Without the extended quantization rate (6.3) such a rate only holds for functionals  $F$  with a bounded Hessian. For further applications (like Romberg extrapolation) and numerical experiments devoted to path-dependent option pricing, we refer to [12].

## APPENDIX

This appendix presents a proof of Inequality (5.2) which is the starting point of the proof of Proposition 2. We closely follow the proof of Theorem 4 in [5]. This part of the proof is reproduced for the reader's convenience.

Set  $\Delta\alpha_n^k = \alpha_n^k - \alpha_n^{k-1}$ . After an appropriate change of variable the  $L^r$ -stationarity equation (see [5])

$$r \int_{\frac{\alpha_n^k + \alpha_n^{k-1}}{2}}^{\frac{\alpha_n^{k+1} + \alpha_n^k}{2}} |a_n^k - x|^{r-1} \text{sign}(a_n^k - x) f(x) dx = 0, \quad 1 \leq k \leq n,$$

reads:

$$\int_{-\frac{\Delta\alpha_n^k}{2}}^{\frac{\Delta\alpha_n^{k+1}}{2}} v^{r-1} \text{sign}(v) f(\alpha_n^k + v) dv = 0, \quad 1 \leq k \leq n.$$

Introducing  $f(\alpha_n^k)$  we obtain:

$$\begin{aligned} & \int_{-\frac{\Delta\alpha_n^k}{2}}^0 (-v)^{r-1} (f(\alpha_n^k + v) - f(\alpha_n^k)) dv + f(\alpha_n^k) \frac{(\Delta\alpha_n^k)^r}{r2^r} \\ &= \int_0^{\frac{\Delta\alpha_n^{k+1}}{2}} v^{r-1} (f(\alpha_n^k + v) - f(\alpha_n^k)) dv + f(\alpha_n^k) \frac{(\Delta\alpha_n^{k+1})^r}{r2^r}. \end{aligned}$$

Setting  $v := \frac{\Delta\alpha_n^k}{2}u$  and  $v := \frac{\Delta\alpha_n^{k+1}}{2}u$  respectively we have, since  $f(\alpha_n^k) > 0$ ,

$$\begin{aligned} & (\Delta\alpha_n^k)^r \left( 1 + r \int_0^1 u^{r-1} \frac{f(\alpha_n^k - \frac{\Delta\alpha_n^k}{2}u) - f(\alpha_n^k)}{f(\alpha_n^k)} du \right) \\ &= (\Delta\alpha_n^{k+1})^r \left( 1 + r \int_0^1 u^{r-1} \frac{f(\alpha_n^k + \frac{\Delta\alpha_n^{k+1}}{2}u) - f(\alpha_n^k)}{f(\alpha_n^k)} du \right). \end{aligned}$$

Setting  $H(x, y) := \int_0^1 ru^{r-1} \frac{f(x + u \frac{y-x}{2}) - f(x)}{f(x)} du$  finally leads to

$$\left( \frac{\Delta\alpha_n^{k+1}}{\Delta\alpha_n^k} \right)^r = \frac{1 + H(\alpha_n^k, \alpha_n^{k+1})}{1 + H(\alpha_n^k, \alpha_n^{k-1})}. \quad (6.4)$$

Let  $[a, b] \subset (m, M)$  and let  $L_f^{a,b}$  denote the Lipschitz coefficient of  $f$  on  $[a, b]$ .

$$\forall \xi, \xi' \in [a, b], \quad |H(\xi, \xi')| \leq L_f^{a,b} \frac{r}{r+1} \frac{|\xi - \xi'|}{f(\xi)} \leq C_{a,b,f} |\xi - \xi'|$$

since  $f$  is bounded away from 0 on  $[a, b]$ . Consequently:

$$|H(\alpha_n^k, \alpha_n^{k-1})| \leq C_{a,b,f} |\alpha_n^k - \alpha_n^{k-1}| \quad \text{and} \quad |H(\alpha_n^k, \alpha_n^{k+1})| \leq C_{a,b,f} |\alpha_n^k - \alpha_n^{k+1}|,$$

whenever  $\alpha_n^{k\pm 1}$  lie in  $[a, b]$ .

The probability density function  $f$  being bounded away from 0 on  $[a, b]$ ,  $\max_{\{k / a \leq \alpha_n^{k\pm 1} \leq b\}} \max(\Delta\alpha_n^k, \Delta\alpha_n^{k+1})$  goes to zero (see *e.g.* Prop. 2 in [5]), so we can estimate the right hand of (6.4):

$$\frac{1 + H(\alpha_n^k, \alpha_n^{k-1})}{1 + H(\alpha_n^k, \alpha_n^{k+1})} = \exp(H(\alpha_n^k, \alpha_n^{k-1}) - H(\alpha_n^k, \alpha_n^{k+1})) + O(\max(\Delta\alpha_n^k, \Delta\alpha_n^{k+1})^2)$$

where

$$\max_{\{k / a \leq \alpha_n^{k\pm 1} \leq b\}} |O(\max(\Delta\alpha_n^k, \Delta\alpha_n^{k+1})^2)| \leq C \max_{\{k / a \leq \alpha_n^{k\pm 1} \leq b\}} \max(\Delta\alpha_n^k, \Delta\alpha_n^{k+1})^2 \xrightarrow{n \rightarrow \infty} 0. \quad (6.5)$$

Now, we have  $\sum_{\{k / a \leq \alpha_n^{k \pm 1} \leq b\}} \Delta \alpha_n^k \leq b - a$ , so that:

$$\sum_{\{k / a \leq \alpha_n^{k \pm 1} \leq b\}} (\Delta \alpha_n^k)^2 \leq (b - a) \sup_{\{k / a \leq \alpha_n^{k \pm 1} \leq b\}} (\Delta \alpha_n^k) = (b - a)o(1). \quad (6.6)$$

Now set  $h_\ell := \exp(H(\alpha_n^{k+\ell}, \alpha_n^{k+\ell+1}) - H(\alpha_n^{k+\ell}, \alpha_n^{k+\ell-1}))$  and  $\delta_\ell := O(\max(\Delta \alpha_n^{k+\ell+1}, \Delta \alpha_n^{k+\ell})^r)$ . Note that, as long as  $\alpha_n^k$  and  $\alpha_n^{\ell+p} \in [a, b]$ ,  $\min_{1 \leq \ell \leq p-1} h_\ell$  is bounded away from 0 by a real constant  $\underline{h} > 0$  not depending on  $n$  since the function  $H$  is bounded away from  $-\infty$  on  $[a, b]^2$ . Hence

$$\left( \frac{\Delta \alpha_n^{k+p}}{\Delta \alpha_n^{k+1}} \right)^r \leq \prod_{\ell=1}^{p-1} (h_\ell + |\delta_\ell|) = \left( \prod_{\ell=1}^{p-1} h_\ell \right) \prod_{\ell=1}^{p-1} \left( 1 + \frac{\delta_\ell}{h_\ell} \right) \leq \left( \prod_{\ell=1}^{p-1} h_\ell \right) \exp \left( \sum_{\ell=1}^{p-1} \frac{\delta_\ell}{h_\ell} \right) \leq \left( \prod_{\ell=1}^{p-1} h_\ell \right) e^{(b-a)o(1)}.$$

One obtains a lower bound the same way round. Assume that  $n$  is large enough so that  $\max\{\delta_\ell, a \leq \alpha_n^{\ell \pm 1} \leq b\} \leq \underline{h}/2$ . Using that  $\ln(1 - u) \geq -2u$  if  $0 \leq u \leq 1/2$ ,

$$\left( \frac{\Delta \alpha_n^{k+p}}{\Delta \alpha_n^{k+1}} \right)^r \geq \left( \prod_{\ell=1}^{p-1} h_\ell \right) \prod_{\ell=1}^{p-1} \left( 1 - \frac{\delta_\ell}{h_\ell} \right) \geq \left( \prod_{\ell=1}^{p-1} h_\ell \right) \exp \left( -2 \sum_{\ell=1}^{p-1} \frac{\delta_\ell}{h_\ell} \right) \geq \left( \prod_{\ell=1}^{p-1} h_\ell \right) e^{(b-a)o(1)}.$$

Thus it yields that for large enough  $n$  and every  $\alpha_n^{\ell+p}, \alpha_n^\ell \in [a, b]$ :

$$\exp \left( \sum_{\ell=1}^{p-1} \ln(h_\ell) - (b - a)|o(1)| \right) \leq \left( \frac{\Delta \alpha_n^{k+p}}{\Delta \alpha_n^{k+1}} \right)^r \leq \exp \left( \sum_{\ell=1}^{p-1} \ln(h_\ell) + (b - a)|o(1)| \right). \quad (6.7)$$

On the other hand, using once again that  $f$  is bounded away from 0 on  $[a, b]$ , yields

$$\begin{aligned} \left| \sum_{\ell=1}^{p-1} \ln(h_\ell) \right| &= \left| \sum_{\ell=1}^{p-1} r \int_0^1 u^{r-1} \frac{f(\alpha_n^{k+\ell} - \frac{\Delta \alpha_n^{k+\ell}}{2} u) - f(\alpha_n^{k+\ell} + \frac{\Delta \alpha_n^{k+\ell+1}}{2} u)}{f(\alpha_n^{k+\ell})} du \right| \\ &\leq L_f^{a,b} \frac{r}{r+1} \sum_{\ell=1}^{p-1} \left| \frac{\Delta \alpha_n^{k+\ell+1} - \Delta \alpha_n^{k+\ell}}{2f(\alpha_n^{k+\ell})} \right| \leq C(b - a). \end{aligned} \quad (6.8)$$

Combined with Inequality (6.7), this provides

$$\max_{\{k / a \leq x_{i-1}^{(n)} \leq \alpha_n^k \leq b\}} \Delta \alpha_n^k \leq C \min_{\{k / a \leq \alpha_n^{k-1} \leq \alpha_n^k \leq b\}} \Delta \alpha_n^k.$$

## REFERENCES

- [1] V. Bally and C. Pagès, A quantization algorithm for solving discrete time multidimensional optimal stopping problems. *Bernoulli* **9** (2003) 1003–1049.
- [2] V. Bally, C. Pagès and J. Printems, First order schemes in the numerical quantization method. *Mathematical Finance* **13** (2001) 1–16.
- [3] J.A. Bucklew and G.L. Wise, Multidimensional asymptotic quantization theory with  $r$ -th power distortion measure. *IEEE Trans. Inform. Theory*, **28**, Special issue on quantization, A. Gersho & R.M. Grey Eds., (1982) 239–247.
- [4] S. Delattre, S. Graf, H. Luschgy and G. Pagès, Quantization of probability distributions under norm-based distortion measures. *Statist. Decisions* **22** (2004) 261–282.
- [5] S. Delattre, J.C. Fort and G. Pagès, Local distortion and  $\mu$ -mass of the cells of one dimensional asymptotically optimal quantizers. *Comm. Statist. Theory Methods* **33** (2004) 1087–1117.
- [6] S. Graf and H. Luschgy, *Foundations of Quantization for Probability Distributions. Lect. Notes in Math.* **1730**, Springer, Berlin (2000).

- [7] S. Graf and H. Luschgy, Rates of convergence for the empirical quantization error. *Ann. Probab.* **30** (2002) 874–897.
- [8] H. Luschgy and G. Pagès, Functional quantization of stochastic processes. *J. Funct. Anal.* **196** (2002) 486–531.
- [9] H. Luschgy and G. Pagès, Sharp asymptotics of the functional quantization problem for Gaussian processes. *Ann. Probab.* **32** (2004) 1574–1599.
- [10] P. Mattila, *Geometry of Sets and Measures in Euclidean Spaces*. Cambridge University Press (1995).
- [11] G. Pagès, A space vector quantization method for numerical integration. *J. Comput. Appl. Math.* **89** (1997) 1–38.
- [12] G. Pagès and J. Printems, Functional quantization for numerics with an application to option pricing. *Monte Carlo Methods & Applications* **11** (2005) 407–446.
- [13] A. Sellami, *Quantization based filtering method using first order approximation*. Pré-pub. LPMA-1009 (2005). To appear in *SIAM J. Numerical Analysis*.
- [14] P.L. Zador, *Development and evaluation of procedures for quantizing multivariate distributions*. Ph.D. thesis, Stanford University (1963).
- [15] P.L. Zador, Asymptotic quantization error of continuous signals and the quantization dimension. *IEEE Trans. Inform. Theory* **28**, Special issue on quantization, A. Gersho & R.M. Grey Eds. (1982) 139–149.