Universal -rate-optimality of -optimal quantizers by dilatation and contraction

Abass Sagna

Universal L^s-rate-optimality of L^r-optimal quantizers by dilatation and contraction

Abass Sagna

Laboratoire de probabilités et modèles aléatoires, UMR7599, Université Pierre et Marie Curie, Case 188, 4 place Jussieu, 75252 Cedex 05, Paris, France. E-mail address: sagna@ccr.jussieu.fr.

Abstract

We investigate in this paper the properties of some dilatations or contractions of a sequence $(\alpha_n)_{n\geq 1}$ of L^r-optimal quantizers of an $\mathbb{R}^d$-valued random vector $X \in L^r(\mathbb{P})$ defined in the probability space $(\Omega, \mathcal{A}, \mathbb{P})$ with distribution $\mathbb{P}_X = P$. To be precise, we investigate the L^s-quantization rate of sequences $\alpha_n^{\theta,\mu} = \mu + \theta(\alpha_n - \mu) = \{\mu + \theta(a - \mu),\ a \in \alpha_n\}$ when $\theta \in \mathbb{R}_+^\star, \mu \in \mathbb{R}, s \in (0, r)$ or $s \in (r, +\infty)$ and $X \in L^s(\mathbb{P})$. We show that for a wide family of distributions, one may always find parameters (θ, μ) such that $(\alpha_n^{\theta,\mu})_{n\geq 1}$ is L^s-rate-optimal. For the Gaussian and the exponential distributions we show the existence of a couple $(\theta^\star, \mu^\star)$ such that $(\alpha^{\theta^\star,\mu^\star})_{n\geq 1}$ also satisfies the so-called L^s-empirical measure theorem. Our conjecture, confirmed by numerical experiments, is that such sequences are asymptotically L^s-optimal. In both cases the sequence $(\alpha^{\theta^\star,\mu^\star})_{n\geq 1}$ is incredibly close to L^s-optimality. However we show (see Remark 5.4) that this last sequence is not L^s-optimal (e.g when $s = 2,\ r = 1$) for the exponential distribution.

1 Introduction

Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and let $X : (\Omega, \mathcal{A}, \mathbb{P}) \longrightarrow \mathbb{R}^d$ be a random variable with distribution $\mathbb{P}_X = P$. The $L^r(P)$-optimal quantization problem of size n for P (or X) consists in the study of the best approximation of X by a $\sigma(X)$-measurable random vector taking at most n values. For $X \in L^r(\mathbb{P})$ this leads to the following optimization problem:

$$e_{n,r}(X) = \inf\{\|X - q(X)\|_r, q : \mathbb{R}^d \to \alpha, \alpha \subset \mathbb{R}^d,\ \text{card}(\alpha) \leq n\}.$$

Let $\alpha \subset \mathbb{R}^d$ be a subset (a codebook) of size n. A Borel partition $C_a(\alpha)_{a\in\alpha}$ of $\mathbb{R}^d$ satisfying

$$C_a(\alpha) \subset \{x \in \mathbb{R}^d : |x - a| = \min_{b\in\alpha_n} |x - b|\},$$

where $|\cdot|$ denotes a norm on $\mathbb{R}^d$ is called a Voronoi partition of $\mathbb{R}^d$ (with respect to α and $|\cdot|$). The random variable $\widehat{X}^\alpha$ taking values in the codebook α defined by

$$\widehat{X}^\alpha = \sum_{a\in\alpha} a\mathbf{1}_{\{X\in C_a(\alpha)\}}.$$

This provisional PDF is the accepted version. The article should be cited as: ESAIM: PS, doi: 10.1051/ps:2008008

is called a Voronoi quantization of X. In other words, it is the nearest neighbour projection of X onto the codebook (also called grid) α.

For any Borel function $q : \mathbb{R}^d \to \alpha$,

$$|X - q(X)| \geq \min_{a \in \alpha} d(X, a) = d(X, \alpha) = |X - \widehat{X}^{\alpha}| \quad \mathbb{P} \text{ a.s}$$

so that

$$\begin{aligned} e_{n,r}(X) &= \inf \{ \|X - \widehat{X}^{\alpha}\|_r, \alpha \subset \mathbb{R}^d, \text{ card}(\alpha) \leq n \} \\ &= \inf_{\substack{\alpha \subset \mathbb{R}^d \\ \text{card}(\alpha) \leq n}} \left(\int_{\mathbb{R}^d} d(x, \alpha)^r dP(x) \right)^{1/r}. \end{aligned} \tag{1.1}$$

For all $n \geq 1$, the infimum in (1.1) is reached at one (at least) codebook $\alpha^\star$; $\alpha^\star$ is then called a L^r-optimal n-quantizer. In addition, if $\text{card}(\text{supp}(P)) \geq n$ then $\text{card}(\alpha^\star) = n$ (see [5] or [9]). Moreover the quantization error, $e_{n,r}(X)$, decreases to zero as n goes to infinity and the so-called Zador's Theorem gives its convergence rate under a slightly stringent moment assumption on X.

Zador Theorem (see [5]) : Suppose $\mathbb{E}|X|^{r+\eta} < +\infty$ for some $\eta > 0$ and let $P = P_a + P_s$ be the Lebesgue decomposition of P with respect to the Lebesgue measure λ_d, where P_a denotes the absolutely continuous part and P_s the singular part of P. Then

$$\lim_{n \to +\infty} n^{r/d} (e_{n,r}(P))^r = Q_r(P).$$

with

$$Q_r(P) = J_{r,d} \left(\int_{\mathbb{R}^d} f^{\frac{d}{d+r}} d\lambda_d \right)^{\frac{d+r}{d}} = J_{r,d} \, \|f\|_{\frac{d}{d+r}} \in [0, +\infty),$$

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

where $U([0,1]^d)$ denotes the uniform distribution on the set $[0,1]^d$ and $f = \frac{dP_a}{d\lambda_d}$. Furthermore, this theorem naturally suggests to set the following definitions.

A sequence of n-quantizers $(\alpha_n)_{n \geq 1}$ is

- $L^r(P)$**-rate-optimal** (or **rate-optimal** for X) if

$$\limsup_{n \to +\infty} n^{1/d} \left(\int_{\mathbb{R}^d} d(x, \alpha_n)^r dP(x) \right)^{1/r} < +\infty,$$

- **asymptotically** $L^r(P)$**-optimal** if

$$\lim_{n \to +\infty} n^{r/d} \int_{\mathbb{R}^d} d(x, \alpha_n)^r dP(x) = Q_r(P),$$

- $L^r(P)$**-optimal** if for all $n \geq 1$,

$$e^r_{n,r}(P) = \int_{\mathbb{R}^d} d(x, \alpha_n)^r dP(x).$$

Optimal quantizers are used in many fields of application like Signal Processing (discretization of emitted signals) or more recently, Numerical Probability where they provide some simples quadrature formulae for the computation of expectations and conditional expectations. This approach has been extensively developed in Finance for the pricing of American options, swing options (commodities), portfolio management (stochastic control) or stochastic volatility estimation (non linear filtration); we refer for example to [10] for applications to stochastic control problems. The errors bounds in these quadrature formulae are always based on the mean quantization error $\|X - \widehat{X}^{\alpha}\|_s$ where α is an optimal L^r-quantizer, usually with $r \leq s$.

Motivated by this problem, the asymptotic behavior of the L^s-mean quantization error of a sequence of L^r-quantizers has been extensively investigated in [6]. A lower bound has been established which shows that if the distribution P is unbounded in the sense that the density function $f = \frac{dP}{d\lambda_d}$ satisfies $\lambda_d(f > 0) = +\infty$ then for any sequence $(\alpha_n)_{n\geq 1}$ of asymptotically L^r-optimal quantizers, $\liminf_n n^{\frac{1}{d}} \|X - \widehat{X}^{\alpha_n}\|_s = +\infty,\ \forall s > r + d$.

On the other hand, under natural assumptions in the tail of the distribution P, it is shown in [6] that for any sequence of L^r-optimal quantizers, $\forall s \in (0, r+d),\ \limsup_n n^{\frac{1}{d}} \|X - \widehat{X}^{\alpha_n}\|_s < +\infty\ \ i.e$ $(\alpha_n)_{n\geq 1}$ remains L^s-rate-optimal as long as $s < r + d$.

The aim of this paper is to show that some simple transformation of the L^r-optimal quantizers, namely some dilatation-translation, makes possible to overcome the critical exponent $r + d$: we will establish that for a wide family of distributions, one can always find $\theta \in \mathbb{R}_+^{\star}$ and $\mu \in \mathbb{R}$ (depending on r, s and d but not on n) such that $(\alpha_n^{\theta,\mu})_{n\geq 1}$ is L^s-rate-optimal. From a general upper bounds that we establish for such transformed sequences of quantizers we derive an heuristic to specify some explicit optimal (in a sense which will be elucidated later) scaling parameters (θ, μ) for several families of distributions (Gaussian Vector, exponential and gamma distributions). Some numerical computations carried with the Gaussian and the exponential distributions show that the resulting sequence of quantizers is very close to L^s-optimality.

So, one application could be to use these quantizers to initialize the procedures used for L^s-optimal (and local optimal) quantizers search when $s \neq 2$. Indeed, in the quadratic case, $s = 2$, several stochastic procedures like the Competitive Learning Vector Quantization algorithm or the randomized Lloyd's I procedure have been designed. Both rely on the stationary property: $\widehat{X}^{\alpha} = \mathbb{E}(X|\widehat{X}^{\alpha})$, satisfied by optimal (and locally optimal) quadratic quantizers. In one dimension, Newton's method is used to compute the optimal quadratic quantizers. Thus a whole package of optimal n-quantizers of the $\mathcal{N}(0, I_d)$ distributions are available in the website `www.quantize.maths-fi.com` for $d \in \{1, \cdots, 10\}$ and $n \in \{1, \cdots, 5000\}$. But, when $s \neq 2$, the natural extension of these procedures become more difficult to implement due to some loss of stability. When $s > 2$ the procedures tend to explode more and more often while when $1 \leq s < 2$ the convergence phase becomes chaotic. In particular, the sensibility of the procedure to its initialization increases as s moves away from 2. Thus, initializing theses procedures by the dilated-contracted L^2-optimal (or locally optimal) quantizers would make them more stable and speed up the convergence. This is what we do to carry the L^4-optimal quantizers of the one dimensional Gaussian distribution (used for numerical experiments in Section 5.1.2) by Newton's method. In fact, initializing this procedure to a n-tuple different from the dilated sequence usually makes the hessian matrix of the L^4-quantization error singular (which makes the procedure very unstable), especially when the grid's size becomes large (typically when $n \geq 400$).

The paper is organized as follows. In section 2 we establish a general lower bound for dilated-translated sequences of quantizers. General upper bounds are also established in section 3 for such sequences. In section 4 we provide a necessary and sufficient condition of L^s-rate-optimality for the

dilated-translated sequences. Section 5 deals with some examples of distributions for which we give the set of parameters (θ, μ) such that the dilated-translated sequence is L^s-rate-optimal and try to find the couple (if any) which makes the resulted transformed sequence satisfy the L^s-empirical measure theorem. The last section is devoted to some applications.

NOTATIONS : • Let α_n be a set of n points of $\mathbb{R}^d$. For every $\mu \in \mathbb{R}^d$ and every $\theta > 0$ we denote $\alpha_n^{\theta,\mu} = \mu + \theta(\alpha_n - \mu) = \{\mu + \theta(a - \mu),\ a \in \alpha_n\}$.

• Let $f : \mathbb{R}^d \longrightarrow \mathbb{R}^d$ be a Borel function and let $\mu \in \mathbb{R}^d, \theta > 0$. One notes by $f_{\theta,\mu}$ (or f_θ if $\mu = 0$) the function defined by $f_{\theta,\mu}(x) = f(\mu + \theta(x - \mu)),\ x \in \mathbb{R}^d$.

• If $X \sim P$, $P_{\theta,\mu}$ will stand for the probability measure of the random variable $\frac{X-\mu}{\theta} + \mu,\ \theta > 0, \mu \in \mathbb{R}^d$. In other words, it is the distribution image of P by $x \longmapsto \frac{x-\mu}{\theta} + \mu$. Note that if $P = f \cdot \lambda_d$ then $\frac{dP_{\theta,\mu}}{d\lambda_d} = \theta^d f_{\theta,\mu}$.

• If A is a matrix A' stands for its transpose.

• Set $x = (x_1, \cdots, x_d);\ y = (y_1, \cdots, y_d) \in \mathbb{R}^d$; we denote $[x, y] = [x_1, y_1] \times \cdots \times [x_d, y_d]$.

• Let $|\cdot|$ be a norm on $\mathbb{R}^d$ and let A be a subset of $\mathbb{R}^d$; we denote by $B(x, r)$ the closed ball, centered to x with radius $r > 0$ and by $d(x, A)$ the distance between x and A; both with respect to the norm $|\cdot|$.

2 Lower estimate

Let $r, s > 0$. Consider an asymptotically $L^r(P)$-optimal sequence of quantizers $(\alpha_n)_{n\geq 1}$. For every $\mu \in \mathbb{R}^d$ and any $\theta > 0$, we construct the sequence $(\alpha_n^{\theta,\mu})_{n\geq 1}$ and try to lower bound asymptotically the L^s-quantization error induced by this sequence. This estimation provides a necessary condition of rate-optimality for the sequence $(\alpha_n^{\theta,\mu})_{n\geq 1}$. In the particular case where $\theta = 1$ and $\mu = 0$ we get the same result as in [6].

Theorem 2.1. *Let $r, s \in (0, +\infty)$, and let X be a random variable taking values in $\mathbb{R}^d$ with distribution P such that $P_a = f.\lambda_d \not\equiv 0$. Suppose that $\mathbb{E}|X|^{r+\eta} < \infty$ for some $\eta > 0$. Let $(\alpha_n)_{n\geq 1}$ be an asymptotically $L^r(P)$-optimal sequence of quantizers. Then, for every $\theta > 0$ and for every $\mu \in \mathbb{R}^d$,*

$$\liminf_{n\to+\infty} n^{s/d} \|X - \widehat{X}^{\alpha_n^{\theta,\mu}}\|_s^s \geq Q_{r,s}^{Inf}(P, \theta), \tag{2.1}$$

with

$$Q_{r,s}^{Inf}(P, \theta) = \theta^{s+d} J_{s,d} \left(\int_{\mathbb{R}^d} f^{\frac{d}{d+r}} d\lambda_d \right)^{s/d} \int_{\{f>0\}} f_{\theta,\mu} f^{-\frac{s}{d+r}} d\lambda_d.$$

Proof. Let $m \geq 1$ and

$$f_m^{\theta,\mu} = \sum_{k,l=0}^{m2^m-1} \frac{l}{2^m} \mathbf{1}_{E_k^m \cap G_l^m};$$

with

$$E_k^m = \left\{ \frac{k}{2^m} \leq f < \frac{k+1}{2^m} \right\} \cap B(0, m) \text{ and } G_l^m = \left\{ \frac{l}{2^m} \leq f_{\theta,\mu} < \frac{l+1}{2^m} \right\} \cap B(0, m).$$

The sequence $(f_m^{\theta,\mu})_{m\geq 1}$ is non-decreasing and

$$\lim_{m\to+\infty} f_m^{\theta,\mu} = f_{\theta,\mu} \qquad \lambda_d \text{ p.p.}$$

Let

$$I_m = \{(k,l) \in \{0, \cdots, m2^m - 1\}^2 : \lambda_d(E_k^m) > 0; \lambda_d(G_l^m) > 0\}.$$

For every $(k,l) \in I_m$ there exists compact sets K_k^m and L_l^m such that :

$$K_k^m \subset E_k^m \, , \, L_l^m \subset G_l^m \, , \, \lambda_d(E_k^m \backslash K_k^m) \leq \frac{1}{m^4 2^{2m+1}} \text{ and } \lambda_d(G_l^m \backslash L_l^m) \leq \frac{1}{m^4 2^{2m+1}}.$$

Then

$$\begin{aligned} (E_k^m \cap G_l^m) \backslash (K_k^m \cap L_l^m) &= E_k^m \cap G_l^m \cap ((K_k^m)^c \cup (L_l^m)^c) \\ &\subset (E_k^m \backslash K_k^m) \cup (G_l^m \backslash L_l^m). \end{aligned}$$

Consequently,

$$\begin{aligned} \lambda_d(E_k^m \cap G_l^m \backslash K_k^m \cap L_l^m) &\leq \lambda_d(E_k^m \backslash K_k^m) + \lambda_d(G_l^m \backslash L_l^m) \\ &\leq \frac{1}{m^4 2^{2m+1}} + \frac{1}{m^4 2^{2m+1}} \\ &= \frac{1}{m^4 2^{2m}}. \end{aligned}$$

For every $m \geq 1$ and every $(k,l) \in I_m$, set

$$A_{k,l}^m := K_k^m \cap L_l^m,$$

$$\tilde{f}_m^{\theta,\mu} := \sum_{k,l=0}^{m2^m-1} \frac{l}{2^m} \mathbf{1}_{A_{k,l}^m},$$

and

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

We get

$$\{f_m^{\theta,\mu} \neq \tilde{f}_m^{\theta,\mu}\} \subset \bigcup_{k,l \in \{0,\cdots,m2^m-1\}} \left((E_k^m \cap G_l^m) \backslash A_{k,l}^m\right).$$

Therefore, for every $m \geq 1$,

$$\lambda_d(\{f_m^{\theta,\mu} \neq \tilde{f}_m^{\theta,\mu}\}) \leq \sum_{k,l=0}^{m2^m-1} \frac{1}{m^4 2^{2m}} = \frac{1}{m^2}$$

and finally

$$\sum_{m \geq 1} \mathbf{1}_{\{f_m^{\theta,\mu} \neq \tilde{f}_m^{\theta,\mu}\}} < \infty \qquad \lambda_d \text{ p.p.}$$

As a consequently $\lambda_d(dx)$-p.p, $f_m^{\theta,\mu}(x) = \tilde{f}_m^{\theta,\mu}(x)$ for large enough m. Then $\tilde{f}_m^{\theta,\mu} \overset{\lambda_d \text{ p.p.}}{\longrightarrow} f_{\theta,\mu}$ when $m \to +\infty$. Since in addition $A_{k,l}^m \subset E_k^m \cap G_l^m$ we obtain

$$\tilde{f}_m^{\theta,\mu} \leq f_m^{\theta,\mu} \leq f_{\theta,\mu}. \tag{2.2}$$

Moreover, for every $n \geq 1$,

$$\begin{aligned} n^{s/d} \, \|X - \widehat{X}^{\alpha_n^{\theta,\mu}}\|_s^s &= n^{s/d} \int_{\mathbb{R}^d} d(z, \mu + \theta(\alpha_n - \mu))^s f(z) \lambda_d(dz) \\ &= n^{s/d} \int_{\mathbb{R}^d} \min_{a \in \alpha_n} |z - (\mu + \theta(a - \mu))|^s f(z) \lambda_d(dz) \\ &= \theta^s n^{s/d} \int_{\mathbb{R}^d} \min_{a \in \alpha_n} |(z - \mu)/\theta + \mu - a|^s f(z) \lambda_d(dz). \end{aligned}$$

Making the change of variable $x := (z - \mu)/\theta + \mu$ yields:

$$\begin{aligned} n^{s/d} \, \|X - \widehat{X}^{\alpha_n^{\theta,\mu}}\|_s^s &= \theta^{s+d} n^{s/d} \int_{\mathbb{R}^d} d(x, \alpha_n)^s f_{\theta,\mu}(x) \lambda_d(dx) \\ &\geq \theta^{s+d} n^{s/d} \int_{\mathbb{R}^d} d(x, \alpha_n)^s \tilde{f}_m^{\theta,\mu} \lambda_d(dx) \qquad \text{(by (2.2))} \\ &= \theta^{s+d} n^{s/d} \sum_{k,l=0}^{m 2^m - 1} \frac{l}{2^m} \int_{A_{k,l}^m} d(x, \alpha_n)^s \lambda_d(dx). \end{aligned} \tag{2.3}$$

Let $m \geq 1$ and $(k, l) \in I_m$. Define the closed sets $\tilde{A}_{k,l}^m$ by $\tilde{A}_{k,l}^m = \emptyset$ if $\lambda_d(\tilde{A}_{k,l}^m) = 0$ and otherwise by

$$\tilde{A}_{k,l}^m = \{x \in \mathbb{R}^d : d(x, A_{k,l}^m) \leq \varepsilon_m\}$$

where $\varepsilon_m \in (0, 1]$ is chosen so that

$$\int_{\tilde{A}_{k,l}^m} f^{\frac{d}{d+r}} d\lambda_d \leq \big(1 + 1/m\big) \int_{A_{k,l}^m} f^{\frac{d}{d+r}} d\lambda_d.$$

Since $\tilde{A}_{k,l}^m$ is compact $\big(\tilde{A}_{k,l}^m \subset B(0, m+1) \; \forall (k,l)\big)$, and

$$A_{k,l}^m \subset (\tilde{A}_{k,l}^m)_{\varepsilon_m/2} := \{x \in \mathbb{R}^d : d(x, A_{k,l}^m) \leq \varepsilon_m/2\} = \{x \in \mathbb{R}^d : d(x, (\tilde{A}_{k,l}^m)^c) > \varepsilon_m/2\},$$

there is (ref. [1], *Lemma 4.3*) a finite " firewall" set $\beta_{k,l}^m$ such that $\forall n \geq 1, \; \forall x \in (\tilde{A}_{k,l}^m)_{\varepsilon_m/2}$,

$$d(x, \alpha_n \cup \beta_{k,l}^m) = d(x, (\alpha_n \cup \beta_{k,l}^m) \cap \tilde{A}_{k,l}^m).$$

This last equality holds in particular for every $x \in A_{k,l}^m$ since $A_{k,l}^m \subset (\tilde{A}_{k,l}^m)_{\varepsilon_m/2}$.

Now set $\beta^m = \bigcup_{k,l} \beta_{k,l}^m$ and $n_{k,l}^m = \text{card}((\alpha_n \cup \beta^m) \cap \tilde{A}_{k,l}^m)$. The empirical measure theorem (see (5.3)) yields

$$\limsup_n \frac{\text{card}(\alpha_n \cap \tilde{A}_{k,l}^m)}{n} = \frac{\int_{\alpha_n \cap \tilde{A}_{k,l}^m} f^{\frac{d}{d+r}} d\lambda_d}{\int f^{\frac{d}{d+r}} d\lambda_d} \leq \frac{\int_{\tilde{A}_{k,l}^m} f^{\frac{d}{d+r}} d\lambda_d}{\int f^{\frac{d}{d+r}} d\lambda_d}.$$

Moreover

$$\frac{n_{k,l}^m}{n} \sim \frac{\text{card}(\alpha_n \cap \tilde{A}_{k,l}^m)}{n} \qquad \text{when } \; n \to +\infty$$

then

$$\liminf_{n \to +\infty} \frac{n}{n_{k,l}^m} \geq \frac{\int f^{\frac{d}{d+r}} d\lambda_d}{\int_{\tilde{A}_{k,l}^m} f^{\frac{d}{d+r}} d\lambda_d} \geq \frac{m}{m+1} \frac{\int f^{\frac{d}{d+r}} d\lambda_d}{\int_{A_{k,l}^m} f^{\frac{d}{d+r}} d\lambda_d}. \tag{2.4}$$

On the other hand,

$$\begin{aligned}\int_{A_{k,l}^m} d(x,\alpha_n)^s \lambda_d(dx) &\geq \int_{A_{k,l}^m} d(x,(\alpha_n \cup \beta_{k,l}^m) \cap \tilde{A}_{k,l}^m)^s \lambda_d(dx) \\ &= \lambda_d(A_{k,l}^m) \int d(x,(\alpha_n \cup \beta_{k,l}^m) \cap \tilde{A}_{k,l}^m)^s \mathbf{1}_{A_{k,l}^m}(x) \frac{\lambda_d(dx)}{\lambda_d(A_{k,l}^m)} \\ &\geq \lambda_d(A_{k,l}^m) e^s_{n_{k,l}^m,s}(U(A_{k,l}^m)),\end{aligned}$$

where $U(A) = \mathbf{1}_A/\lambda_d(A)$ denotes the uniform distribution in the Borel set A when $\lambda_d(A) \neq 0$. Then we can write for every $(k,l) \in I_m$,

$$\liminf_{n\to+\infty} n^{s/d} \int_{A_{k,l}^m} d(x,\alpha_n)^s \lambda_d(dx) \geq \lambda_d(A_{k,l}^m) \liminf_n \left(\frac{n}{n_{k,l}^m}\right)^{s/d} \liminf_n n^{s/d} e^s_{n,s}(U(A_{k,l}^m)),$$

since

$$\liminf_n n^{s/d} e^s_{n,s}(U(A_{k,l}^m)) \geq J_{s,d} \cdot \lambda_d(A_{k,l}^m)^{s/d}.$$

Owing to Equation (2.4), one has

$$\liminf_{n\to+\infty} n^{s/d} \int_{A_{k,l}^m} d(x,\alpha_n)^s \lambda_d(dx) \geq \lambda_d(A_{k,l}^m) \left(\frac{m}{m+1} \frac{\int f^{\frac{d}{d+r}} d\lambda_d}{\int_{A_{k,l}^m} f^{\frac{d}{d+r}} d\lambda_d}\right)^{s/d} J_{s,d} \cdot \lambda_d(A_{k,l}^m)^{s/d}.$$

However, on the sets $A_{k,l}^m$, the statement $\frac{1}{f} \geq \left(\frac{k+1}{2^m}\right)^{-1}$ holds since $f < \frac{k+1}{2^m}$ on E_k^m. Hence

$$\liminf_{n\to+\infty} n^{s/d} \int_{A_{k,l}^m} d(x,\alpha_n)^s \lambda_d(dx) \geq J_{s,d} \left(\frac{m+1}{m} \int f^{\frac{d}{d+r}} \lambda_d(dx)\right)^{s/d} \left(\frac{k+1}{2^m}\right)^{-\frac{d}{d+r}\cdot\frac{s}{d}} \lambda_d(A_{k,l}^m).$$

It follows from Equation (2.3) and the super-additivity of the liminf that for every $m \geq 1$,

$$\begin{aligned}\liminf_n n^{s/d} \|X - \widehat{X}^{\alpha_n^{\theta,\mu}}\|_s^s &\geq \theta^{s+d} J_{s,d} \left(\frac{m+1}{m} \int f^{\frac{d}{d+r}} \lambda_d(dx)\right)^{s/d} \sum_{k,l=0}^{m2^m-1} \frac{l}{2^m} \left(\frac{k+1}{2^m}\right)^{-\frac{s}{d+r}} \lambda_d(A_{k,l}^m) \\ &\geq \theta^{s+d} J_{s,d} \left(\frac{m+1}{m} \int f^{\frac{d}{d+r}} \lambda_d(dx)\right)^{s/d} \int_{\{f>0\}} \tilde{f}_m^{\theta,\mu} (\tilde{f}_m + 2^{-m})^{-\frac{s}{d+r}} d\lambda_d.\end{aligned}$$

Finally, applying Fatou's Lemma yields

$$\liminf_{n\to+\infty} n^{s/d} \|X - \widehat{X}^{\alpha_n^{\theta,\mu}}\|_s^s \geq \theta^{s+d} J_{s,d} \left(\int_{\mathbb{R}^d} f^{\frac{d}{d+r}} d\lambda_d\right)^{s/d} \int_{\{f>0\}} f_{\theta,\mu} f^{-\frac{s}{d+r}} d\lambda_d.$$

□

3 Upper estimate

Let $r, s > 0$. Let $(\alpha_n)_{n\geq 1}$ be an (asymptotically) $L^r(P)$ - optimal sequence of quantizers. In this section we will provide some sufficient conditions of $L^s(P)$-rate-optimality for the sequence $(\alpha_n^{\theta,\mu})_{n\geq 1}$.

Definition 3.1. *Let $\theta > 0,\ \mu \in \mathbb{R}^d$ and let P be a probability distribution such that $P = f \cdot \lambda_d$. The couple (θ, μ) is said P-**admissible** if*

$$\{f > 0\} \subset \mu(1-\theta) + \theta\{f > 0\} \qquad \lambda_d\text{-p.p.} \tag{3.1}$$

One remarks that when $\text{supp}(P) = \mathbb{R}^d$ then every couple (θ, μ) is P-admissible. Indeed, every $x \in \mathbb{R}^d$ can be written $x = \mu(1-\theta) + \theta z$ with $z = \frac{x-\mu(1-\theta)}{\theta}$ and $f(z) > 0$.

Theorem 3.1. *Let $r, s \in (0, +\infty), s < r$ and let X be a random variable taking values in $\mathbb{R}^d$ with distribution P such that $P = f \cdot \lambda_d$. Suppose that (θ, μ) is P-admissible for some $\theta > 0; \mu \in \mathbb{R}$, and $\mathbb{E}|X|^{r+\eta} < \infty$, for some $\eta > 0$. Let $(\alpha_n)_{n\geq 1}$ be an asymptotically L^r-optimal sequence of n-quantizers. If*

$$\int_{\{f>0\}} f_{\theta,\mu}^{\frac{r}{r-s}} f^{-\frac{s}{r-s}} d\lambda_d < +\infty \tag{3.2}$$

then, $(\alpha_n^{\theta,\mu})_{n\geq 1}$ is $L^s(P)$-rate-optimal and

$$\limsup_{n\to+\infty} n^{s/d} \left\| X - \widehat{X}^{\alpha_n^{\theta,\mu}} \right\|_s^s \leq \theta^{s+d} \left(Q_r(P)\right)^{s/r} \left(\int_{\{f>0\}} f_{\theta,\mu}^{\frac{r}{r-s}} f^{-\frac{s}{r-s}} d\lambda_d \right)^{1-\frac{s}{r}}. \tag{3.3}$$

Remark 3.1. *Note that if $\theta = 1$ and $\mu = 0$ then*

$$\int_{\{f>0\}} f_{\theta,\mu}^{\frac{r}{r-s}} f^{-\frac{s}{r-s}} d\lambda_d = \int_{\{f>0\}} f^{\frac{r}{r-s}} f^{-\frac{s}{r-s}} d\lambda_d = \int_{\{f>0\}} f d\lambda_d = 1.$$

Which gives the expected result since $\|X - \widehat{X}^{\alpha_n}\|_s \leq \|X - \widehat{X}^{\alpha_n}\|_r$.

Proof. Let P^θ denote the distribution of the random variable θX. P^θ is absolutely continuous with respect to λ_d, with $p.d.f\ g_\theta(x) = \theta^{-d} f(\frac{x}{\theta})$.

For every $n \geq 1$,

$$\begin{aligned} n^{s/d} \left\| X - \widehat{X}^{\alpha_n^{\theta,\mu}} \right\|_s^s &= n^{s/d} \int_{\mathbb{R}^d} d(x, \alpha_n^{\theta,\mu})^s dP(x) \\ &= n^{s/d} \int_{\{f>0\}} \min_{a\in\alpha_n} |x - \mu(1-\theta) - \theta a|^s f(x) d\lambda_d(x). \end{aligned}$$

Making the change of variable $z := x - \mu(1-\theta)$ gives

$$\begin{aligned} n^{s/d} \left\| X - \widehat{X}^{\alpha_n^{\theta,\mu}} \right\|_s^s &= n^{s/d} \int_{\{f>0\}-\mu(1-\theta)} d(z, \theta\alpha_n)^s f(z + \mu(1-\theta)) d\lambda_d(z) \\ &\leq n^{s/d} \int_{\theta\{f>0\}} d(z, \theta\alpha_n)^s f(z + \mu(1-\theta)) g_\theta^{-1}(z) dP^\theta(z) \qquad (3.4) \\ &\leq n^{s/d} \left(\int_{\mathbb{R}^d} d(z, \theta\alpha_n)^r dP^\theta(z) \right)^{s/r} \left(\int_{\theta\{f>0\}} \left(f(z + \mu(1-\theta)) g_\theta^{-1}(z) \right)^{\frac{r}{r-s}} dP^\theta(z) \right)^{\frac{r-s}{r}} \\ &\leq \left(n^{r/d} \|\theta X - \widehat{\theta X}^{\theta\alpha_n}\|_r^r \right)^{s/r} \left(\int_{\theta\{f>0\}} f(z + \mu(1-\theta))^{\frac{r}{r-s}} g_\theta^{-\frac{s}{r-s}}(z) d\lambda_d(z) \right)^{\frac{r-s}{r}} \end{aligned}$$

where we used the P-admissibility of (θ, μ) in the first inequality. The second inequality derives from *Hölder* inequality applied with $p = r/s > 1$ and $q = 1 - s/r$.

Moreover

$$\|\theta X - \widehat{\theta X}^{\theta\alpha_n}\|_r^r = \mathbb{E}\big(\min_{a\in\alpha_n} |\theta X - \theta a|^r\big) = \theta^r \|X - \widehat{X}^{\alpha_n}\|_r^r. \tag{3.5}$$

Then

$$n^{s/d}\,\|X - \widehat{X}^{\alpha_n^{\theta,\mu}}\|_s^s \le \theta^s \left(n^{r/d}\|X - \widehat{X}^{\alpha_n}\|_r^r\right)^{s/r} \left(\int_{\theta\{f>0\}} f(z+\mu(1-\theta))^{\frac{r}{r-s}} g_\theta^{-\frac{s}{r-s}}(z) d\lambda_d(z)\right)^{\frac{r-s}{r}}.$$

Owing to the asymptotically $L^r(P)$-optimality of (α_n) and making again the change of variable $x := z/\theta$ yields

$$\begin{aligned}
\limsup_{n\to+\infty} n^{s/d}\,\|X - \widehat{X}^{\alpha_n^{\theta,\mu}}\|_s^s &\le \theta^s\,(Q_r(P))^{s/r} \left(\theta^{\frac{ds}{r-s}} \int_{\theta\{f>0\}} f(z+\mu(1-\theta)))^{\frac{r}{r-s}} f(z/\theta)^{-\frac{s}{r-s}} d\lambda_d(z)\right)^{\frac{r-s}{r}} \\
&= \theta^s\,(Q_r(P))^{s/r} \left(\theta^{\frac{rd}{r-s}} \int_{\{f>0\}} f_{\theta,\mu}(x)^{\frac{r}{r-s}} f(x)^{-\frac{s}{r-s}} d\lambda_d(x)\right)^{\frac{r-s}{r}} \\
&= \theta^{s+d}\,(Q_r(P))^{s/r} \left(\int_{\{f>0\}} f_{\theta,\mu}(x)^{\frac{r}{r-s}} f(x)^{-\frac{s}{r-s}} d\lambda_d(x)\right)^{\frac{r-s}{r}}.
\end{aligned}$$

□

The next theorem provides a less accurate asymptotic upper bound than the previous one since, beyond the restriction on the distribution of X, we need now the sequence (α_n) to be (exactly) $L^r(P)$-optimal. Before giving the theorem, recall first the following result established in [6] and related to the maximal function $\psi_b : \mathbb{R}^d \longrightarrow \mathbb{R}_+ \cup \{+\infty\}$ defined by

$$\psi_b(x) = \sup_{n\ge1} \frac{\lambda_d(B(x, bd(x,\alpha_n)))}{P(B(x, bd(x,\alpha_n)))}. \tag{3.6}$$

Proposition 3.1. *Let $b \in \big(0, 1/2\big)$, $X \sim P$, with $P_a \neq 0$, such that $\mathbb{E}|X|^{r+\eta} < \infty$, for some $\eta > 0$. Let (α_n) be an $L^r(P)$-optimal sequence of quantizers. Then $\forall x \in \mathbb{R}^d, \forall n \ge 1$,*

$$n^{1/d} d(x, \alpha_n) \le C(b)\psi_b(x)^{1/(d+r)} \tag{3.7}$$

where $C(b)$ denotes a real constant not depending on n.

Theorem 3.2. *Let $r, s \in (0, +\infty)$ and let X be a random variable taking values in $\mathbb{R}^d$ with distribution P such that $P = f \cdot \lambda_d$. Suppose that $\mathbb{E}|X|^{r+\eta} < \infty$ for some $\eta > 0$ and $P_{\theta,\mu} \ll P$ $\big($i.e $P_{\theta,\mu}$ is absolutely continuous with respect to $P\big)$ for some $\theta > 0, \mu \in \mathbb{R}^d$. Let $(\alpha_n)_{n\ge1}$ be an $L^r(P)$- optimal sequence of quantizers and suppose that the maximal function defined previously satisfies*

$$\psi_b^{s/(d+r)} \in\ L^1(P_{\theta,\mu}) \text{ for some } b \in (0, 1/2). \tag{3.8}$$

Then,

$$\limsup_n n^{s/d}\,\|X - \widehat{X}^{\alpha_n^{\theta,\mu}}\|_s^s \le C(b)\,\theta^{s+d} \int_{\{f>0\}} f_{\theta,\mu} f^{-\frac{s}{d+r}} d\lambda_d < +\infty \tag{3.9}$$

where $C(b)$ is a positive real constant not depending on θ, μ and n.

Notice that this theorem does not require (θ, μ) to be P-admissible.

Proof. It follows from the definition of ψ_b that (because f is a limit which is less than the sup)

$$f^{-\frac{s}{d+r}} \leq \psi_b^{\frac{s}{d+r}} \qquad P_{\theta,\mu}\text{-a.s.}$$

Then, under Assumption (3.8),

$$\int f^{-\frac{s}{d+r}} dP_{\theta,\mu} = \int_{\{f>0\}} f_{\theta,\mu} f^{-\frac{s}{d+r}} d\lambda_d < +\infty.$$

For all $n \geq 1$,

$$\begin{aligned} n^{s/d} \left\| X - \widehat{X}^{\alpha_n^{\theta,\mu}} \right\|_s^s &= n^{s/d} \int_{\mathbb{R}^d} d(z, \alpha_n^{\theta,\mu})^s f(z) d\lambda_d(z) \\ &= n^{s/d} \theta^s \int_{\mathbb{R}^d} \min_{a \in \alpha_n} |(z-\mu)/\theta + \mu - a|^s f(z) d\lambda_d(z) \end{aligned}$$

We make the change of variable $x := (z-\mu)/\theta + \mu$. Then

$$\begin{aligned} n^{s/d} \left\| X - \widehat{X}^{\alpha_n^{\theta,\mu}} \right\|_s^s &= n^{s/d} \theta^{s+d} \int_{\mathbb{R}^d} d(x, \alpha_n)^s f(\mu + \theta(x-\mu)) d\lambda_d(x) \\ &= n^{s/d} \theta^s \int_{\mathbb{R}^d} d(x, \alpha_n)^s dP_{\theta,\mu}(x). \end{aligned}$$

Moreover, it is established in [6] that

$$\limsup_n n^{s/d} d(\cdot, \alpha_n)^s \leq C(b) f^{-\frac{s}{d+r}}.$$

Hence, from Inequality (3.7) and under Assumption (3.8), we can apply the Lebesgue dominated convergence theorem to the above inequalities to get

$$\begin{aligned} \limsup_n n^{s/d} \int d(x, \alpha_n)^s dP_{\theta,\mu}(x) &\leq \int \limsup_n n^{s/d} d(x, \alpha_n)^s dP_{\theta,\mu}(x) \\ &\leq C(b) \int f^{-\frac{s}{d+r}} dP_{\theta,\mu}(x). \\ &= \theta^d C(b) \int_{\{f>0\}} f_{\theta,\mu}(x) f^{-\frac{s}{d+r}}(x) d\lambda_d(x). \end{aligned}$$

□

For a given distribution, Assumption (3.8) is not easy to verify. But when $s \neq r + d$, the lemma and criterions below provide a sufficient condition so that Assumption (3.8) is satisfied. In the rest of this section we extend some of the results obtained in [6].

Let $P = f \cdot \lambda_d$ be an absolutely continuous distribution. Let $r, s \in (0, +\infty)$ and (θ, μ) be a P-admissible couple of parameters. We will need the following hypotheses:

(H1) for all $M > 0$,

$$\sup_{z \in B(0,M)} \frac{f(\mu + \theta(z-\mu))}{f(z)} \mathbf{1}_{\{f(z)>0\}} < +\infty. \tag{3.10}$$

(H2) There exists $b \in (0, 1/2),\ M \in (0, +\infty)$ such that

$$\int_{B(0,M)^c} \left(\sup_{t \leq 2b|x|} \frac{\lambda_d(B(x,t))}{P(B(x,t))} \right)^{s/(d+r)} dP_{\theta,\mu} < +\infty. \tag{3.11}$$

(H3) $\lambda_d(\cdot \cap \text{supp}(P)) \ll P$ and $\text{supp}(P)$ is a finite union of closed convex sets.

Lemma 3.1. *Let $P = f \cdot \lambda_d$ and $r > 0$ such that $\int |x|^r P(dx) < +\infty$. Assume $(\alpha_n)_{n\geq 1}$ is a sequence of quantizers such that $\int d(x, \alpha_n)^r dP \to 0$. Let (θ, μ) be a P-admissible couple of parameters for which (**H1**) holds.*

(a) If $p \in (0, 1)$ then for every $b > 0,\ \psi_b^p \in L^1_{loc}(P_{\theta,\mu})$.

*(b) If $p \in (1, +\infty]$ and if furthermore **(H3)** holds then for every $b > 0$,*

$$f^{-p} \in L^1_{loc}(P) \Longrightarrow \psi_b^p \in L^1_{loc}(P_{\theta,\mu}).$$

Proof. It follows from the P-admissibility of (θ, μ) that

$$P_{\theta,\mu}(dz) = \theta^d f(\mu + \theta(z - \mu))\lambda_d(dz) = g_\theta(z)P(dz),$$

where $g_\theta(z) = \theta^d \frac{f(\mu+\theta(z-\mu))}{f(z)} \mathbf{1}_{\{f(z)>0\}}$. Then g_θ is locally bounded by (**H1**).

(a) If $p \in (0, 1)$, it follows from Lemma 1 in [6] that $\psi_b^p \in L^1_{\text{loc}}(P)$. Hence $\psi_b^p \in L^1_{\text{loc}}(P_{\theta,\mu})$ since g_θ is locally bounded.

(b) If $p \in (1, +\infty)$ it follows from Lemma 2 in [6] that if $f^{-p} \in L^1_{\text{loc}}(P)$ then $\psi_b^p \in L^1_{\text{loc}}(P_{\theta,\mu})$ since g_θ is locally bounded. □

Corollary 3.1. *(Distributions with unbounded support) Let $r > 0,\ s \in (0, +\infty),\ s \neq r + d$ and let X be a random variable with probability measure $P = f \cdot \lambda_d$ such that $E|X|^{r+\eta} < +\infty$ for some $\eta > 0$. Let (θ, μ) be P-admissible and suppose that **(H1)**, **(H2)** hold.*

(a) If $s \in (0, r + d)$ then Assumption (3.8) of Theorem 3.2 holds true.

*(b) If $s \in (r + d, +\infty)$, and if furthermore, **(H3)** holds and $f^{-\frac{s}{r+d}} \in L^1_{loc}(P)$ then Assumption (3.8) of Theorem 3.2 holds true.*

Proof. Let $x_0 \in \text{supp}(P)$. We know from [1] that $d(x_0, \alpha_n) \to 0$. Then following the lines of the proof of Corollary 2 in [6] one has for $|x| > N = |x_0| + \sup_{n\geq 1} d(x_0, \alpha_n),\ d(x, \alpha_n) \leq 2|x|$ for every $n \geq 1$. Thus for every $b > 0,\ x \in B(0, N)^c$,

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

Now, coming back to the core of our proof, it follows from **(H2)** that (for b coming from **(H2)**),

$$\int_{B(0,M\vee N)^c} \psi_b^{s/(d+r)} dP_{\theta,\mu} < +\infty.$$

Since

$$\int \psi_b^{s/(d+r)} dP_{\theta,\mu} = \int_{B(0,M\vee N)} \psi_b^{s/(d+r)} dP_{\theta,\mu} + \int_{B(0,M\vee N)^c} \psi_b^{s/(d+r)} dP_{\theta,\mu},$$

it remains to show that the first term in the right hand side of this last equality is finite.

(a) If $s \in (0, r+d)$ it follows from Lemma 3.1, (a) that the first term in the right hand side of the above equality is finite. As a consequence, $\psi^{\frac{s}{r+d}} \in L^1(P_{\theta,\mu})$.

(b) If $s > r+d$, the first term in the right hand side of the above equality still finite owing to Lemma 3.1, (b). Consequently, Assumption (3.8) of Theorem 3.2 holds true provided **(H3)** holds and $f^{-\frac{s}{r+d}} \in L^1_{loc}(P)$. □

We next give two useful criterions ensuring that Hypothesis **(H2)** holds. The first one is useful for distributions with radial tails and the second one for distributions which does not satisfy this last assumption.

Criterion 3.1. *Let $X \sim P$. Suppose that $P = f \cdot \lambda_d$ and $E|X|^{r+\eta} < +\infty$ for some $\eta > 0$.*
(a) Let $r, s > 0$ and $f = h(|\cdot|)$ on $B_{|\cdot|}(0,N)^c$ with $h : (R, +\infty) \to \mathbb{R}_+$, $R \in \mathbb{R}_+$, a decreasing function and $|\cdot|$ any norm on $\mathbb{R}^d$. Suppose that (θ, μ) is a couple of P-admissible parameters such that

$$\int f(cx)^{-\frac{s}{d+r}} dP_{\theta,\mu}(x) < +\infty \tag{3.12}$$

*for some $c > 1$. Then **(H2)** holds.*
*(b) Let $r, s > 0$. Suppose $supp(P) \subset [R_0, +\infty)$ for some $R_0 \in \mathbb{R}$ and $f_{|(R_0', +\infty)}$ is decreasing for $R_0' \geq R_0$. Assume furthermore that (θ, μ) is a couple of P-admissible parameters such that (3.12) is satisfied for some $c > 1$. Then Hypothesis **(H2)** holds.*

Note that (b) follows from (a) for $d = 1$ and that (a) is simply deduced from the proof of Corollary 3 in [6] since it has been shown that for $b \in (0, 1/2)$, $M := N/(1-2b)$ one has for every $x \in B(0,M)^c$,

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

Criterion 3.2. *Let $r, s > 0$, $P = f \cdot \lambda_d$ and $\int |x|^{r+\eta} P(dx) < +\infty$ for some $\eta > 0$. Let (θ, μ) be a P-admissible couple such that*

$$\sup_{z \neq 0} \frac{f(\mu + \theta(z-\mu))}{f(z)} \mathbf{1}_{\{f(z)>0\}} < +\infty. \tag{3.13}$$

Assume furthermore that

$$\inf_{x \in supp(P), \rho > 0} \frac{\lambda_d(supp(P) \cap B(x,\rho))}{\lambda_d(B(x,\rho))} > 0$$

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

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

If

$$\int f(x)^{-\frac{s(1+\varepsilon)}{d+r}} dP(x) < +\infty, \tag{3.14}$$

*then **(H2)** holds. If in particular f satisfies the local growth control assumption for $\varepsilon = 0$ or for every $\varepsilon \in (0, \underline{\varepsilon}]$, with $\underline{\varepsilon} > 0$, and if*

$$\int f(x)^{-\frac{s}{d+r}} dP(x) = \int_{\{f>0\}} f(x)^{1-\frac{s}{d+r}} d\lambda_d(x) < +\infty$$

*then Hypothesis **(H2)** holds.*

Notice that Hypothesis (3.13) can be relaxed if we suppose that $f(x)^{-\frac{s(1+\varepsilon)}{d+r}} \in L^1(P_{\theta,\mu})$ instead of (3.14).

The criterion follows from Corollary 4 in [6].

4 Toward a necessary and sufficient condition for $L^s(P)$-rate-optimality

Let $X \sim P$. Let us make some comments about inequalities (2.1) and (3.9). Note first that the moment assumption $\mathbb{E}|X|^{r+\eta} < +\infty$ for some $\eta > 0$, ensure that $\int_{\mathbb{R}^d} f^{\frac{d}{d+r}} d\lambda_d < +\infty$ (cf [5]). Consequently, if $\int_{\{f>0\}} f_{\theta,\mu} f^{-\frac{s}{d+r}} d\lambda_d = +\infty$ one derives from inequality (2.1) that

$$\lim_{n\to+\infty} n^{s/d} \, \|X - \widehat{X}^{\alpha_n^{\theta,\mu}}\|_s^s = +\infty.$$

Then the sequence $(\alpha_n^{\alpha,\mu})_{n\geq 1}$ is not L^s-rate-optimal.

On the other hand if $\int_{\{f>0\}} f_{\theta,\mu} f^{-\frac{s}{d+r}} d\lambda_d < +\infty$ one derives from Inequality (3.9) that $(\alpha_n^{\theta,\mu})_{n\geq 1}$ is L^s-rate-optimal. This leads to a necessary and sufficient condition so that the sequence $(\alpha_n^{\theta,\mu})_{n\geq 1}$ (in particular the sequence $(\alpha_n)_{n\geq 1}$ by taking $\theta = 1$ and $\mu = 0$) is L^s-rate-optimal.

Remark 4.1. *Let $\mu \in \mathbb{R}^d$, $\theta, r > 0$ and let P be a probability distribution such that $P = f \cdot \lambda_d$. Assume (θ, μ) is P-admissible. Let $(\alpha_n)_{n\geq 1}$ be an $L^r(P)$-optimal sequence of n-quantizers and suppose that Assumption (3.8) of Theorem 3.2 holds true. Then for every $s > 0$,*

$$(\alpha_n^{\theta,\mu})_{n\geq 1} \text{ is } L^s\text{-rate-optimal} \iff \int_{\{f>0\}} f_{\theta,\mu} f^{-\frac{s}{d+r}} d\lambda_d < +\infty. \tag{4.1}$$

Remark 4.2. *If $s < r$ and if $(\alpha_n)_{n\geq 1}$ is asymptotically L^r-optimal, Inequality (3.3) provides a sufficient condition so that the sequence $(\alpha_n^{\theta,\mu})_{n\geq 1}$ is L^s-rate-optimal, which is : $\int_{\{f>0\}} f_{\theta,\mu}^{\frac{r}{r-s}} f^{-\frac{s}{r-s}} d\lambda_d < +\infty$ (always satisfied by $(\alpha_n)_{n\geq 1}$ itself).*

However, it follows from Hölder inequality (applied to $p = \frac{r}{r-s} > 1$ and $q = \frac{r}{s}$) that $\forall s < r$,

$$\begin{aligned}
\int_{\{f>0\}} f_{\theta,\mu} f^{-\frac{s}{d+r}} d\lambda_d &= \int_{\{f>0\}} f_{\theta,\mu} f^{-\frac{s}{r}} f^{\frac{sd}{r(d+r)}} d\lambda_d \\
&\leq \left(\int_{\{f>0\}} f_{\theta,\mu}^{\frac{r}{r-s}} f^{-\frac{s}{r-s}} d\lambda_d \right)^{1-\frac{s}{r}} \left(\int f^{\frac{d}{d+r}} d\lambda_d \right)^{\frac{s}{r}}.
\end{aligned}$$

One deduces that

$$\Big\{(\theta,\mu) \text{ s. t } \int_{\{f>0\}} f_{\theta,\mu}^{\frac{r}{r-s}} f^{-\frac{s}{r-s}} d\lambda_d < +\infty\Big\} \subset \Big\{(\theta,\mu) \text{ s. t } \int_{\{f>0\}} f_{\theta,\mu} f^{-\frac{s}{d+r}} d\lambda_d < +\infty\Big\}. \tag{4.2}$$

As a consequence, if $(\alpha_n)_{n\geq 1}$ is an $L^r(P)$-optimal sequence of quantizers and if assumptions of Theorem 3.2 are fulfilled then for every $s < r$ we will rather use Inequality (3.9) instead of (3.3) to find the couple of parameters (θ, μ) so that the sequence is $L^s(P)$-rate-optimal. But If $(\alpha_n)_{n\geq 1}$ is simply asymptotically $L^r(P)$-optimal, we only have at our disposal Inequality (3.3) to find this set of parameters.

Now, for $s \neq r$, is it possible to find a $\theta = \theta^\star$ for which the sequence $(\alpha_n^{\theta^\star,\mu})_{n\geq 1}$ is asymptotically $L^s(P)$-optimal? (when $s < r$ this is the only question of interest since we know that $(\alpha_n)_{n\geq 1}$ is $L^s(P)$-rate-optimal for every $s < r$).

Let $(\alpha_n)_{n\geq 1}$ be an (asymptotically) $L^r(P)$-optimal sequence of quantizers. For a fixed r, b and μ, we can write from inequalities (3.3) and (3.9) :

$$\limsup_n n^{s/d} \, \|X - \widehat{X}^{\alpha_n^{\theta,\mu}}\|_s^s \leq Q_{r,s}^{\text{Sup}}(P, \theta) \tag{4.3}$$

with

$$Q_{r,s}^{\text{Sup}}(P,\theta) = \begin{cases} \theta^{s+d} \left(Q_r(P)\right)^{s/r} \left(\int_{\{f>0\}} f_{\theta,\mu}^{\frac{r}{r-s}} f^{-\frac{s}{r-s}} d\lambda_d \right)^{1-\frac{s}{r}} & \text{if } s < r \\ \theta^{s+d} C(b) \int_{\{f>0\}} f_{\theta,\mu} f^{-\frac{s}{d+r}} d\lambda_d & \text{if } s > r. \end{cases}$$

One knows that for a given $s > 0$, we have for all $n \geq 1$,

$$e_{n,s}^s(X) \leq \|X - \widehat{X}^{\alpha_n^{\theta,\mu}}\|_s^s.$$

Then for every $\theta > 0$,

$$Q_s(P) \leq Q_{r,s}^{\text{Sup}}(P, \theta).$$

Consequently for a fixed $s > 0$, in order to have the best estimation of Zador's constant in L^s, we must minimize over θ, the quantity $Q_{r,s}^{\text{Sup}}(P, \theta)$. In that way, we may hope to reach the sharp rate of convergence in Zador Theorem and so construct an asymptotically L^s-optimal sequence.

For μ well chosen, the examples below show that, for the Gaussian and the exponential distributions, the minimum $\theta^\star$ exists and the sequence $(\alpha_n^{\theta^\star,\mu})_{n\geq 1}$ satisfies the empirical measure theorem and is suspected to be asymptotically L^s-optimal.

5 Examples of distributions

Let $(\alpha_n)_{n\geq 1}$ be an (asymptotically) $L^r(P)$-optimal sequence of quantizers for a given probability distribution P and consider the sequence $(\alpha_n^{\theta,\mu})_{n\geq 1}$. For a fixed μ and s, we try to solve the following minimization problem

$$\theta^\star = \arg\min_{\theta>0} \left\{ Q_{r,s}^{\text{Sup}}(P,\theta),\ (\alpha_n^{\theta,\mu})_{n\geq 1} \ L^s(P)\text{-rate-optimal} \right\}. \tag{5.1}$$

In all examples, C will denote a generic real constant (not depending on θ) which may change from line to line. The choice of μ depends on the probability measure and it is not clear how to choose it. But Proposition 6.1 morally implies that μ must be chosen so that for every $\theta > 0$, the probability distribution $P_{\theta,\mu}$ lies in the same family of distributions as P so that for the Gamma distribution we will set $\mu = 0$. If for every $\theta > 0, \mu \in \mathbb{R}^d$, $P_{\theta,\mu}$ lies in the same family of distribution as P we will choose μ such that $\theta^\star$ do not depend on μ; which means to put $\mu = \mathbb{E}(X)$ if X is further a symmetric random variable.

5.1 The multivariate Gaussian distribution

5.1.1 Optimal dilatation and contraction

Proposition 5.1. *Let $r, s > 0$ and let $P = \mathcal{N}(m; \Sigma),\ m \in \mathbb{R}^d, \Sigma \in \mathcal{S}^+(d, \mathbb{R})$.*

(a) *If $(\alpha_n)_{n\geq 1}$ is an $L^r(P)$-optimal sequence of quantizers then, for $s \neq r + d$, the sequence $(\alpha_n^{\theta,m})_{n\geq 1}$ is $L^s(P)$-rate-optimal iff $\theta \in \big(\sqrt{s/(d+r)}, +\infty\big)$ and*

$$\theta^\star = \sqrt{(s+d)/(r+d)}\ \in (1, +\infty)$$

is the unique solution of (5.1) on the set $\big(\sqrt{s/(d+r)}, +\infty\big)$.

(b) *If $(\alpha_n)_{n\geq 1}$ is an asymptotically $L^r(P)$-optimal sequence of quantizers then, for $s \in (0, r)$, the sequence $(\alpha_n^{\theta,m})_{n\geq 1}$ is $L^s(P)$-rate-optimal if $\theta \in \big(\sqrt{s/r}, +\infty\big)$ and*

$$\theta^\star = \sqrt{(s+d)/(r+d)}\ \in (0, 1)$$

is the unique solution of (5.1) on the set $\big(\sqrt{s/r}, +\infty\big)$.

Proof. Since the multivariate Gaussian distribution is symmetric and for every $\theta > 0$, $P_{\theta,\mu}$ is also a Gaussian random vector, one sets $\mu = m$. Keep in mind that the probability density function f of P is given for every $x \in \mathbb{R}^d$ by

$$f(x) = \big((2\pi)^d \det \Sigma\big)^{-\frac{1}{2}} e^{-\frac{1}{2}(x-m)'\Sigma^{-1}(x-m)}.$$

Note first that Hypothesis **(H1)** is obviously satisfied from the continuity of $\frac{f(m+\theta(z-m))}{f(z)}\mathbf{1}_{\{f(z)>0\}}$ on every $\bar{B}(0, M),\ M > 0$.

(a) Let $s < d + r$. For every $\theta > 0, \mu \in \mathbb{R}^d$, the couple (θ, μ) is P-admissible and f is radial since $f(x) = \varphi(|x - m|_\Sigma)$ with $\varphi : (0, +\infty) \longmapsto \mathbb{R}_+$ defined by

$$\varphi(\xi) = \big((2\pi)^d \det\Sigma\big)^{-1/2} \exp(-\frac{1}{2}|\xi|^2),\ \text{ with } |x|_\Sigma = |\Sigma^{-\frac{1}{2}}x|.$$

Let $\theta > \sqrt{s/(r+d)}$. Then Assumption (3.12) holds for every $c \in (1, \theta\sqrt{\frac{r+d}{s}})$. Consequently, it follows from Corollary $3.1, (a)$ that Assumption (3.8) of Theorem 3.2 holds.

If $s > d + r$, the required additional hypotheses **(H3)** and $f^{-\frac{s}{r+d}} \in L^1_{loc}(P)$ are clearly satisfied since if $P = f \cdot \lambda_d$ then

$$\lambda_d(\text{supp}(P) \cap \{f = 0\}) = 0 \implies \lambda_d(\cdot \cap \text{supp}(P)) \ll P \tag{5.2}$$

and $f^{-\frac{s}{r+d}}$ is continuous on every $\bar{B}(0, M),\ M > 0$. Then it follows from Corollary $3.1, (b)$ that Assumption (3.8) of Theorem 3.2 holds.

On the other hand

$$\begin{aligned}\int_{\mathbb{R}^d} f_{\theta,m}(x) f(x)^{-\frac{s}{d+r}} dx &= \int_{\mathbb{R}^d} f(m + \theta(x-m)) f(x)^{-\frac{s}{d+r}} dx \\ &= C \int_{\mathbb{R}^d} e^{-\frac{1}{2}(\theta^2 - \frac{s}{d+r})(x-m)'\Sigma^{-1}(x-m)} dx\end{aligned}$$

so that

$$\int_{\mathbb{R}^d} f_{\theta,m}(x) f(x)^{-\frac{s}{d+r}} < +\infty \quad \text{iff} \quad \theta > \sqrt{\frac{s}{d+r}}.$$

Now we are in position to solve the problem (5.1). Let $\theta \in \left(\sqrt{s/(d+r)}, +\infty\right)$,

$$\begin{aligned}
\theta^{s+d} \int_{\mathbb{R}^d} f_{\theta,m}(x) f(x)^{-\frac{s}{d+r}} dx &= \left((2\pi)^d det\ \Sigma\right)^{-\frac{1}{2}(1-\frac{s}{d+r})} \theta^{s+d} \int_{\mathbb{R}^d} e^{-\frac{1}{2}(\theta^2 - \frac{s}{d+r})(x-m)'\Sigma^{-1}(x-m)} dx \\
&= \left((2\pi)^d det\ \Sigma\right)^{-\frac{s}{d+r}} \theta^{s+d} \left(\theta^2 - \frac{s}{d+r}\right)^{-\frac{d}{2}}.
\end{aligned}$$

For $\theta \in \left(\sqrt{s/(d+r)}, +\infty\right)$, we want to minimize the function h defined by

$$h(\theta) = \theta^{s+d} \left(\theta^2 - \frac{s}{d+r}\right)^{-\frac{d}{2}}.$$

The function h is differentiable on $\left(\sqrt{s/(d+r)}, +\infty\right)$ with derivative

$$h'(\theta) = s\theta^{d+s-1} \left(\theta^2 - \frac{s}{d+r}\right)^{-1-d/2} \left(\theta^2 - \frac{s+d}{r+d}\right).$$

One easily checks that h reaches its unique minimum on $\left(\sqrt{s/(d+r)}, +\infty\right)$ at $\theta^\star = \sqrt{(s+d)/(r+d)}$.
(b) Let $s < r$ and consider the inequality (3.3). We get

$$\int f_{\theta,m}^{\frac{r}{r-s}}(x) f^{-\frac{s}{r-s}}(x) dx = C \int_{\mathbb{R}^d} e^{-\frac{1}{2}\frac{r}{r-s}(\theta^2 - \frac{s}{r})(x-m)'\Sigma^{-1}(x-m)} dx.$$

So if $\theta \in \left(\sqrt{s/r}, +\infty\right)$ then $\int f_{\theta,m}^{\frac{r}{r-s}}(x) f^{-\frac{s}{r-s}}(x) dx < +\infty$. This proves the first assertion.

To prove the second assertion, let $\theta \in \left(\sqrt{s/r}, +\infty\right)$. Then

$$\begin{aligned}
\theta^{d+s} \left(\int f_{\theta,m}^{\frac{r}{r-s}}(x) f^{-\frac{s}{r-s}}(x) dx\right)^{1-\frac{s}{r}} &= C\, \theta^{s+d} \left(\int_{\mathbb{R}^d} e^{-\frac{1}{2}\frac{r}{r-s}(\theta^2 - \frac{s}{r})(x-m)'\Sigma^{-1}(x-m)} dx\right)^{1-\frac{s}{r}} \\
&= C\, \theta^{s+d} \left(\theta^2 - \frac{s}{r}\right)^{-\frac{d}{2r}(r-s)}.
\end{aligned}$$

We proceed as before by setting

$$h(\theta) = \theta^\alpha \left(\theta^2 - \frac{s}{r}\right)^\beta, \text{ with } \alpha = d+s \text{ and } \beta = -\frac{d}{2r}(r-s).$$

For all $\theta \in \left(\sqrt{s/r}, +\infty\right)$,

$$h'(\theta) = \theta^{\alpha-1} \left(\theta^2 - \frac{s}{r}\right)^{\beta-1} \left((\alpha+2\beta)\theta^2 - \frac{\alpha s}{r}\right).$$

The sign of h' depends on the sign of $\left((\alpha+2\beta)\theta^2 - \frac{\alpha s}{r}\right)$. Moreover $\alpha + 2\beta = \frac{s}{r}(d+r) > 0$ then h' vanishes at $\theta^\star = \sqrt{(s+d)/(r+d)}$, is negative on the set $\left(\sqrt{s/r}, \theta^\star\right)$ and positive on $\left(\theta^\star, +\infty\right)$. Therefore h reaches its minimum on $\left(\sqrt{s/r}, +\infty\right)$ at the unique point $\theta^\star$. □

Definition 5.1. *A sequence of quantizers $(\beta_n)_{n\geq 1}$ is called a **dilatation** of the sequence $(\alpha_n)_{n\geq 1}$ with **scaling number** θ and **translating number** μ if, for every $n \geq 1,\ \beta_n = \alpha_n^{\theta,\mu}$, with $\theta > 1$. If $\theta < 1$, one defines likewise the **contraction** of the sequence $(\alpha_n)_{n\geq 1}$ with **scaling number** θ and **translating number** μ.*

From this definition follows the following remark.

Remark 5.1. *Let $X \sim \mathcal{N}(m;\Sigma)$.*
If $s < r$ then $\theta^\star < 1$. Hence $(\alpha_n^{\theta^\star,m})_{n\geq 1}$ is a contraction of $(\alpha_n)_{n\geq 1}$ with scaling number $\theta^\star$ and translating number m. On the other hand, if $s > r$, then $\theta^\star > 1$. In this case the sequence $(\alpha_n^{\theta^\star,m})_{n\geq 1}$ is a dilatation of $(\alpha_n)_{n\geq 1}$ with scaling number $\theta^\star$ and translating number m. Also note that $\theta^\star$ does not depend on the covariance matrix Σ.

What we do expect from the resulting sequence $(\alpha_n^{\theta^\star,m})_{n\geq 1}$? Before giving any answer to this question let us recall first the empirical measure theorem (see [5]) which gives the asymptotic distribution of the empirical measure induced by an asymptotically L^r-optimal sequence of quantizers.

Theorem 5.1. *(Empirical measure theorem) Let $X \sim P$, with $P_a \neq 0$, and let $(\alpha_n)_{n\geq 1}$ be an asymptotically $L^r(P)$-optimal sequence of quantizers. Then*

$$\frac{1}{n}\sum_{a\in\alpha_n}\delta_a \xrightarrow{w} P_r \tag{5.3}$$

where $\xrightarrow{w}$ denotes the weak convergence and for every Borel set A of $\mathbb{R}^d$, P_r is defined by

$$P_r(A) = \frac{1}{C_{f,r}}\int_A f(x)^{\frac{d}{d+r}}d\lambda_d(x),\ \textit{with}\quad C_{f,r} = \int_{\mathbb{R}^d} f(x)^{\frac{d}{d+r}}d\lambda_d(x). \tag{5.4}$$

A sequence of quantizers $(\alpha_n)_{n\geq 1}$ will be said to satisfy **the L^r-empirical measure theorem** if (5.3) holds. The next proposition shows that the sequence $(\alpha_n^{\theta^\star,m})_{n\geq 1}$ satisfies the L^s-empirical measure theorem.

Proposition 5.2. *Let $r, s > 0$ and let $P = \mathcal{N}(m;\Sigma)$. Assume $(\alpha_n)_{n\geq 1}$ is asymptotically $L^r(P)$-optimal. Then the sequence $(\alpha_n^{\theta^\star,m})_{n\geq 1}$ (as defined before with $\theta^\star = \sqrt{(s+d)/(r+d)}$) satisfies the L^s-empirical measure theorem.*
In other words, for every $a, b \in \mathbb{R}^d$,

$$\frac{1}{n}\textit{card}(\{x \in \alpha_n^{\theta^\star,m} \cap [a,b]\}) \longrightarrow \frac{1}{C_{f,s}}\int_{[a,b]} f(x)^{\frac{d}{d+s}}dx.$$

Proof. For all $n \geq 1$,

$$\{x \in \alpha_n^{\theta^\star,m} \cap [a,b]\} = \{x \in \alpha_n \cap [(a-m)/\theta^\star + m, (b-m)/\theta^\star] + m\}.$$

Since $(\alpha_n)_{n\geq 1}$ is asymptotically L^r-optimal, applying the empirical measure theorem to the sequence $(\alpha_n)_{n\geq 1}$ yields

$$\frac{1}{n}\text{card}(\{x \in \alpha_n \cap [(a-m)/\theta^\star+m, (b-m)/\theta^\star+m]\}) \longrightarrow \frac{1}{C_{f,r}}\int_{[(a-m)/\theta^\star+m,(b-m)/\theta^\star+m]} f(x)^{\frac{d}{d+r}}dx.$$

It remains to verify that

$$\frac{1}{C_{f,r}}\int_{[(a-m)/\theta^\star+m,(b-m)/\theta^\star+m]} f(x)^{\frac{d}{d+r}}dx = \frac{1}{C_{f,s}}\int_{[a,b]} f(x)^{\frac{d}{d+s}}dx.$$

One knows that

$$f(x) = \big((2\pi)^d \text{det}\,\Sigma\big)^{-\frac{1}{2}} e^{-\frac{1}{2}(x-m)'\Sigma^{-1}(x-m)}$$

and $\big($see $(5.4)\big)$

$$C_{f,r} = \int_{\mathbb{R}^d} f(x)^{\frac{d}{d+r}}dx$$

so that for all $r>0$,

$$C_{f,r} = \big((2\pi)^d \text{det}\,\Sigma\big)^{\frac{r}{2(r+d)}}\left(\frac{d+r}{d}\right)^{\frac{d}{2}}.$$

By making the change of variable $x = m+\theta^\star(z-m)$, one gets :

$$\frac{1}{C_{f,r}}\int_{[(a-m)/\theta^\star+m,(b-m)/\theta^\star+m]} f(z)^{\frac{d}{d+r}}dz \;\;=\;\; \frac{1}{C_{f,r}}(\theta^\star)^{-d}\int_{[a,b]} f((x-m)/\theta^\star+m)^{\frac{d}{d+r}}dx.$$

It is easy to check that

$$\big(f((x-m)/\theta^\star+m\big)^{\frac{d}{d+r}} = \big(f(x)\big)^{\frac{d}{d+s}}\big((2\pi)^d \text{det}\,\Sigma\big)^{-\frac{1}{2}(\frac{d}{d+r}-\frac{d}{d+s})}$$

and that

$$\frac{1}{C_{f,r}}(\theta^\star)^{-d}\big((2\pi)^d \text{det}\Sigma\big)^{-\frac{1}{2}(\frac{d}{d+r}-\frac{d}{d+s})} = \big((2\pi)^d \text{det}\,\Sigma\big)^{-\frac{s}{2(s+d)}}\left(\frac{d+s}{d}\right)^{-\frac{d}{2}}.$$

The last term is simply equal to $\frac{1}{C_{f,s}}$. We then deduce that

$$\frac{1}{C_{f,r}}\int_{[(a-m)/\theta^\star+m,(b-m)/\theta^\star+m]} f(x)^{\frac{d}{d+r}}dx = \frac{1}{C_{f,s}}\int_{[a,b]} f(x)^{\frac{d}{d+s}}dx.$$

□

We have just built a sequence $(\alpha_n^{\theta^\star,m})_{n\geq 1}$ satisfying the empirical measure theorem. The question of interest is now to know whether or not this sequence is asymptotically L^s-optimal. The following proposition shows that the lower bound in (2.1) is in fact reached by considering the sequence $(\alpha_n^{\theta^\star,m})_{n\geq 1}$.

Proposition 5.3. *Let $s>0$ and let $\theta=\theta^\star=\sqrt{(s+d)/(r+d)}$. Then, the constant in the asymptotic lower bound for the L^s-error induced by the sequence $(\alpha_n^{\theta^\star,m})_{n\geq 1}$ (see (2.1)) satisfies :*

$$Q_{r,s}^{Inf}(P,\theta^\star) = Q_s(P). \tag{5.5}$$

Proof. Keep in mind that if $P\sim\mathcal{N}(m;\Sigma)$ then for all $r>0$,

$$\big(Q_r(P)\big)^{1/r} = \big(J_{r,d}\big)^{1/r}\sqrt{2\pi}\left(\frac{d+r}{d}\right)^{\frac{d+r}{2r}}\big(\text{det}\,\Sigma\big)^{\frac{1}{2d}}.$$

We have on one hand

$$\begin{aligned}\left(\int_{\mathbb{R}^d} f^{\frac{d}{d+r}}(x)d(x)\right)^{s/d} &= \left(\left((2\pi)^d \det\Sigma\right)^{-\frac{1}{2}\frac{d}{d+r}} \int_{\mathbb{R}^d} e^{-\frac{1}{2}\frac{d}{d+r}(x-m)'\Sigma^{-1}(x-m)}dx\right)^{s/d} \\ &= \left(\left((2\pi)^d \det\Sigma\right)^{\frac{1}{2}\frac{r}{d+r}} \left(\frac{d+r}{d}\right)^{\frac{d}{2}}\right)^{s/d}\end{aligned}$$

and on the other hand

$$\begin{aligned}\int_{\mathbb{R}^d} f_{\theta^\star,\mu}(x) f^{-\frac{s}{d+r}}(x)d(x) &= \left((2\pi)^d \det\Sigma\right)^{-\frac{1}{2}-\frac{s}{d+r}} \int_{\mathbb{R}^d} e^{-\frac{1}{2}\frac{d}{d+r}(x-m)'\Sigma^{-1}(x-m)}dx \\ &= \left((2\pi)^d \det\Sigma\right)^{-\frac{s}{d+r}} \left(\frac{d+r}{d}\right)^{\frac{d}{2}}.\end{aligned}$$

Combining these two results yields

$$\begin{aligned}Q^{\text{Inf}}_{r,s}(P,\theta^\star) &= (\theta^\star)^{s+d} J_{s,d} \left(\int_{\mathbb{R}^d} f^{\frac{d}{d+r}} d\lambda_d\right)^{s/d} \int_{\mathbb{R}^d} f_{\theta^\star,\mu} f^{-\frac{s}{d+r}} d\lambda_d \\ &= J_{s,d} \left(\frac{s+d}{r+d}\right)^{\frac{d+s}{2}} \left((2\pi)^d \det\Sigma\right)^{\frac{s}{2d}} \left(\frac{r+d}{d}\right)^{\frac{d+s}{2}} \\ &= J_{s,d} \left(\frac{s+d}{d}\right)^{\frac{d+s}{2}} \left((2\pi)^d \det\Sigma\right)^{\frac{s}{2d}} \\ &= Q_s(P).\end{aligned}$$

□

After some elementary calculations, it follows from the proposition above and inequalities (2.1),(4.3), the corollary below :

Corollary 5.1. *Let $X \sim \mathcal{N}(m;\Sigma)$ and $\theta^\star = \sqrt{(s+d)/(r+d)}$. Then,*

$$Q_s(P)^{1/s} \le \liminf_{n\to\infty} n^{1/d} \left\|X - \widehat{X}^{\alpha_n^{\theta^\star,m}}\right\|_s \le \limsup_{n\to\infty} n^{1/d} \left\|X - \widehat{X}^{\alpha_n^{\theta^\star,m}}\right\|_s \le Q^{Sup}_{r,s}(P,\theta^\star)^{1/s} \quad (5.6)$$

with

$$Q^{Sup}_{r,s}(P,\theta^\star)^{1/s} = \begin{cases} \left(\frac{s+d}{d}\right)^{\frac{s+d}{2s}} J_{r,d}^{\frac{1}{r}} \left((2\pi)^d det\,\Sigma\right)^{\frac{1}{2d}} & \text{if } s < r \\ \left(\frac{s+d}{d}\right)^{\frac{d}{2}} \sqrt{\frac{s+d}{r+d}}\, C(b) \left((2\pi)^d det\,\Sigma\right)^{\frac{1}{2(d+r)}} & \text{if } s > r. \end{cases}$$

Remark 5.2. (a) *If $s > r$, we cannot prove the asymptotically $L^s(P)$-optimality of $(\alpha_n^{\theta^\star,m})_{n\ge1}$ using (3.9) since the constant $C(b)$ is not explicit.*

(b) *When $s < r$, the corollary above shows that the upper bound in (3.3) does not reach the Zador constant. Then our upper estimate does not allow us to show that the sequence $(\alpha_n^{\theta^\star,m})_{n\ge1}$ is asymptotically $L^s(P)$-optimal.*

Moreover, using $H\ddot{o}lder$ inequality (with $p = r/(r-s)$ and $q = r/s$), we have for every $\theta > 0$,

$$\begin{aligned}\int_{\mathbb{R}^d} f_{\theta,\mu}(x) f^{-\frac{s}{d+r}}(x)d\lambda_d(x) &= \int_{\mathbb{R}^d} f_{\theta,\mu}(x) f^{-s/r}(x) f^{\frac{sd}{r(d+r)}}(x)d\lambda_d(x) \\ &\le \left(\int_{\mathbb{R}^d} f_{\theta,\mu}^{\frac{r}{r-s}}(x) f^{-\frac{s}{r-s}}(x)d\lambda_d(x)\right)^{\frac{r-s}{r}} \left(\int_{\mathbb{R}^d} f^{\frac{d}{d+r}}(x)d\lambda_d\right)^{\frac{s}{r}}.\end{aligned}$$

and (for $\theta = \theta^\star$)

$$\int_{\mathbb{R}^d} f_{\theta^\star,\mu}(x) f^{-\frac{s}{d+r}}(x) d\lambda_d(x) = \left(\int_{\mathbb{R}^d} f_{\theta^\star,\mu}^{\frac{r}{r-s}}(x) f^{-\frac{s}{r-s}}(x) d\lambda_d(x) \right)^{\frac{r-s}{r}} \left(\int_{\mathbb{R}^d} f^{\frac{d}{d+r}}(x) d\lambda_d \right)^{\frac{s}{r}}. \tag{5.7}$$

Hence, according to (5.5), one gets for every $s < r$,

$$(\theta^\star)^{s+d} J_{s,d} \left(\int_{\mathbb{R}^d} f_{\theta^\star,\mu}^{\frac{r}{r-s}}(x) f^{-\frac{s}{r-s}}(x) d\lambda_d(x) \right)^{\frac{r-s}{r}} \|f\|_{\frac{d}{d+r}}^{s/r} = Q_s(P). \tag{5.8}$$

Thus, to reach the Zador constant in (3.3) we must rather have $J_{s,d}$ instead of $J_{r,d}$ (which will be coherent since for all $s < r,\ J_{s,d}^{1/s} \leq J_{r,d}^{1/r}$), that is,

$$\limsup_{n\to\infty} n^{1/d} \|X - \widehat{X}^{\alpha_n^{\theta,\mu}}\|_s \leq \theta^{s+d} J_{s,d} \left(\int_{\mathbb{R}^d} f_{\theta,\mu}^{\frac{r}{r-s}}(x) f^{-\frac{s}{r-s}}(x) d\lambda_d(x) \right)^{\frac{r-s}{r}} \|f\|_{\frac{d}{d+r}}^{s/r}.$$

5.1.2 Numerical experiments

For numerical example, supppose that $d = 1$ and $r \in \{1, 2, 4\}$. Let $X \sim \mathcal{N}(0, 1)$ and, for a fixed n, let $\alpha_n^{(r)} = \{\alpha_{n1}^{(r)}, \cdots, \alpha_{nn}^{(r)}\}$ be the L^r-optimal grid of size n (obtained by a Newton-Raphson zero search). For every $n \in \{20, 50, \cdots, 900\}$ and for $(s, r) = (1, 2)$ and $(4, 2)$, we make a linear regression of $\alpha_n^{(r)}$ onto $\alpha_n^{(s)}$:

$$\alpha_{ni}^{(s)} \simeq \hat{a}_{sr} \alpha_{ni}^{(r)} + \hat{b}_{sr}, \quad i = 1, \cdots, n.$$

Table 1 provides the regression coefficients we obtain for different values of n. We note that when n increases, the coefficients $\hat{a}_{sr}$ tend to the value $\sqrt{(s+1)/(r+1)} = \theta^\star$ whereas the coefficients $\hat{b}_{sr}$ almost vanish. For example, for $n = 900$ and for $(r, s) = (2, 1)$ (resp. $(2, 4)$) we get $\hat{a}_{sr} = 0.8170251$ (resp. 1.2900417). The expected values are $\sqrt{2/3} = 0.8164966$ (resp. $\sqrt{5/3} = 1.2909944$). The absolute errors are then 5.285×10^{-4} $\big($resp. $9.527 \times 10^{-4}\big)$. We remark that the error mainly stems from the tail of the distribution.

n	$\hat{a}_{12}$	$\hat{b}_{12}$	ϵ		$\hat{a}_{42}$	$\hat{b}_{42}$	ϵ
20	0.8250096	1.826E-14	0.0003025		1.2761027	- 3.650E-12	0.0008607
50	0.8211387	- 1.021E-13	0.0006870		1.2828110	3.733E-10	0.0020110
100	0.8193424	8.693E-14	0.0009909		1.2859567	4.059E-09	0.0029445
300	0.8177506	- 1.045E-11	0.0013601		1.2887640	0.0000004	0.0041021
700	0.8171428	- 7.219E-11	0.0015111		1.2898393	- 0.0000089	0.0048006
800	0.8170775	- 6.725E-11	0.0015247		1.2900041	0.0000216	0.0040577
900	0.8170251	4.564E-11	0.0015346		1.2900417	- 0.0000141	0.0048182

Table 1: Regression coefficients for the Gaussian.

The previous numerical results, in addition to Equation (5.5), strongly suggest that the sequence $(\alpha_n^{\theta^\star,m})_{n\geq 1}$ is in fact asymptotically $L^s(P)$-optimal. This leads to the following conjecture.

Conjecture 1. *Let $P \sim \mathcal{N}(m; \Sigma)$ and let $(\alpha_n)_{n\geq 1}$ be an $L^r(P)$-optimal sequence of quantizers. Then, for every $s > 0$, the sequence $(\alpha_n^{\theta^\star,m})_{n\geq 1}$ (with $\theta^\star = \sqrt{(s+d)/(r+d)}$) is asymptotically $L^s(P)$-optimal.*

5.2 Exponential distribution

5.2.1 Optimal dilatation and contraction

Proposition 5.4. *Let $r, s > 0$ and X be an exponentially distributed random variable with rate parameter $\lambda > 0$. Set $\mu = 0$.*

(a) *If $(\alpha_n)_{n\geq 1}$ is an $L^r(P)$-optimal sequence of quantizers then, for $s \neq r+1$, the sequence $(\alpha_n^{\theta,0})_{n\geq 1}$ is L^s-rate-optimal iff $\theta \in \big(s/(r+1), +\infty\big)$ and*

$$\theta^\star = (s+1)/(r+1)$$

is the unique solution of (5.1) on the set $\big(s/(r+1), +\infty\big)$.

(b) *If $(\alpha_n)_{n\geq 1}$ is an asymptotically $L^r(P)$-optimal sequence of quantizers then, for $s \in (0, r)$, the sequence $(\alpha_n^{\theta,0})_{n\geq 1}$ is L^s-rate-optimal for all $\theta \in \big(s/r, +\infty\big)$ and*

$$\theta^\star = (s+1)/(r+1)$$

is the unique solution of (5.1) on $\big(s/r, +\infty\big)$.

***Proof*. (a)** Let $s < r+1$. For all $\theta > 0, \mu \in \mathbb{R}^d$, the couple (θ, μ) is P-admissible and the function f is decreasing on $(0, +\infty)$. For $\theta > s/(r+1)$, Assumption (3.12) holds true for every $c \in \big(1, \theta(1+r)/s\big)$. Moreover, Hypothesis **(H1)** is clearly satisfied. Consequently, it follows from Corollary $3.1, (a)$ that Assumption (3.8) holds true.

If $s > r+1$, Assumption (3.8) still holds since the additional assumptions **(H3)** and $f^{-\frac{s}{r+1}} \in L^1_{loc}(P)$ required to apply the corollary $3.1, (b)$ are satisfied.

On the other hand, one has

$$\int_{\mathbb{R}} f(\theta x) f(x)^{-s/(r+1)} dx = C \int_0^{+\infty} e^{-\lambda(\theta - s/(r+1))x} dx < +\infty \Longleftrightarrow \theta > s/(r+1).$$

Now, let us solve the problem (5.1). For all $\theta > s/(r+1)$,

$$\begin{aligned} h(\theta) := \theta^{s+1} \int_{\mathbb{R}} f(\theta x) f(x)^{-\frac{s}{r+1}} dx &= C\, \theta^{s+1} \int_0^{+\infty} e^{-\lambda(\theta - \frac{s}{r+1})x} dx \\ &= C\, \theta^{s+1} \left(\theta - \frac{s}{r+1}\right)^{-1}. \end{aligned}$$

Consequently

$$h'(\theta) = Cs\, \theta^s \left(\theta - \frac{s}{r+1}\right)^{-2} \left(\theta - \frac{s+1}{r+1}\right).$$

Hence, h reaches its unique minimum on $\big(s/(r+1), +\infty\big)$ at $\theta^\star = (s+1)/(r+1)$.
(b) Let $s < r$. We have

$$\int_{\mathbb{R}} f^{\frac{r}{r-s}}(\theta x) f^{-\frac{s}{r-s}}(x) dx = C \int_{\mathbb{R}_+} e^{-x\frac{\lambda}{r-s}(r\theta - s)} dx.$$

Then, for all $\theta > s/r$, $\int_{\mathbb{R}} f^{\frac{r}{r-s}}(\theta x) f^{-\frac{s}{r-s}}(x) dx < +\infty$. Which gives the first assertion.

For every $\theta > s/r$,

$$\begin{aligned}\theta^{s+1}\left(\int_{\mathbb{R}} f_{\theta,\mu}^{\frac{r}{r-s}}(x) f^{-\frac{s}{r-s}}(x)dx\right)^{1-\frac{s}{r}} &= C\,\theta^{s+1}\left(\int_{\mathbb{R}_+} e^{-x\frac{\lambda}{r-s}(r\theta-s)}dx\right)^{\frac{r-s}{r}} \\ &= C\,\theta^{s+1}\left(r\theta-s\right)^{\frac{s-r}{r}}.\end{aligned}$$

We easily check that the function $h(\theta) = \theta^{s+1}\left(r\theta - s\right)^{\frac{s-r}{r}}$ reaches its minimum on $\big(s/r, +\infty\big)$ at the unique point $\theta^\star = (s+1)/(r+1)$. □

Remark 5.3. *Let $X \sim \mathcal{E}(\lambda)$. If $s < r$, then $\theta^\star = (s+1)/(r+1) < 1$. As a consequence, the sequence $(\alpha_n^{\theta^\star,0})_{n\geq 1}$ is a contraction of $(\alpha_n)_{n\geq 1}$ with scaling number $\theta^\star$. On the other hand, if $s > r$, then $\theta^\star > 1$ and then $(\alpha_n^{\theta^\star,0})_{n\geq 1}$ is a dilatation of $(\alpha_n)_{n\geq 1}$ with scaling number $\theta^\star$. Note that $\theta^\star$ does not depend on the rate parameter λ of the exponential distribution.*

One shows below that the sequence $(\alpha_n^{\theta^\star,0})_{n\geq 1}$, with $\theta^\star = (1+s)/(1+r)$, satisfies the L^s-empirical measure theorem.

Proposition 5.5. *Let $r, s > 0$ and let X be an exponentially distributed random variable with rate parameter $\lambda > 0$. Assume $(\alpha_n)_{n\geq 1}$ is an asymptotically L^r-optimal sequence of quantizers for X and let $(\alpha_n^{\theta^\star,0})_{n\geq 1}$ be defined as before, with $\theta^\star = (s+1)/(r+1)$. Then, the sequence $(\alpha_n^{\theta^\star,0})$ satisfies the L^s-empirical measure theorem.*

Proof. Since $(\alpha_n^{\theta^\star,0})_{n\geq 1} = (\theta^\star \alpha_n)_{n\geq 1}$ it amounts to show that for every $a, b \in \mathbb{R}_+$

$$\frac{\text{card}(\alpha_n \cap [a/\theta^\star, b/\theta^\star])}{n} \longrightarrow \frac{1}{C_{f,s}}\int_a^b f(x)^{\frac{1}{1+s}}dx$$

i.e that for all $a, b \in \mathbb{R}_+$,

$$\frac{1}{C_{f,r}}\frac{1}{\theta^\star}\int_a^b f(x/\theta^\star)^{\frac{1}{1+r}}dx = \frac{1}{C_{f,s}}\int_a^b f(x)^{\frac{1}{1+s}}dx.$$

Elementary computations show that $\forall\, r > 0$,

$$C_{f,r} = \lambda^{-\frac{r}{1+r}}(1+r).$$

so that

$$\begin{aligned}\frac{1}{C_{f,r}}\frac{1}{\theta^\star}\int_a^b f(x/\theta^\star)^{\frac{1}{1+r}}dx &= \frac{1}{C_{f,r}}\frac{1+r}{1+s}\int_a^b \left(\lambda e^{-x\lambda\frac{1+r}{1+s}}\right)^{\frac{1}{1+r}}dx \\ &= \frac{1}{C_{f,r}}\frac{1+r}{1+s}\,\lambda^{\frac{1}{1+r}-\frac{1}{1+s}}\int_a^b \left(\lambda e^{-\lambda x}\right)^{\frac{1}{1+s}}dx \\ &= \frac{1}{C_{f,s}}\int_a^b f(x)^{\frac{1}{1+s}}dx.\end{aligned}$$

□

Once again, the question of interest is to know if the sequence $(\alpha_n^{\theta^\star,0})_{n\geq 1}$ is asymptotically L^s-optimal. The remark 5.2 is also valid for the exponential distribution. Our upper bounds in (3.3) and (3.9) do not allow us to show that $(\theta^\star \alpha_n)$ is asymptotically L^s-optimal because of the corollary below. But the numerical results strongly suggest that it is.

Corollary 5.2. *Let $X \sim \mathscr{E}(\lambda)$ and $\theta^\star = (s+1)/(r+1)$. Then,*

$$Q_s(P)^{1/s} \leq \liminf_{n\to\infty} n^{1/d}\|X - \widehat{X}^{\alpha_n^{\theta^\star,0}}\|_s \leq \limsup_{n\to\infty} n^{1/d}\,\|X - \widehat{X}^{\alpha_n^{\theta^\star,0}}\|_s \leq Q_{r,s}^{Sup}(P,\theta^\star)^{1/s} \quad (5.9)$$

with

$$Q_{r,s}^{Sup}(P,\theta^\star)^{1/s} = \begin{cases} \frac{1}{2\lambda}(s+1)^{1+1/s}(r+1)^{-1/r} & \text{if } s < r \\ (s+1)^{1+1/s}\big((r+1)\lambda^{\frac{1}{1+r}}\big)^{-1} C(b)^{1/s} & \text{if } s > r. \end{cases}$$

Proof. We easily prove, like in proposition 5.2, that $Q_{r,s}^{\text{Inf}}(P,\theta^\star) = Q_s(P)$. The corollary follows then from (2.1) and (4.3) $\big($Note that for every $r>0, \quad J_{r,1} = \frac{1}{(r+1)2^r}\big)$. □

5.2.2 Numerical experiments

We relate first a proposition established in [2] and used in our context to compute the L^r-optimal quantizers for the exponential distribution.

Proposition 5.6. *Let $r > 0$ and let X be an exponentially distributed random variable with scale parameter $\lambda > 0$. Then for every $n \geq 1$, the L^r-optimal quantizer $\alpha_n^{(r)} = (\alpha_{n1}^{(r)}, \cdots, \alpha_{nn}^{(r)})$ is unique and given by*

$$\alpha_{nk}^{(r)} = \frac{a_n^{(r)}}{2} + \sum_{i=n+1-k}^{n-1} a_i^{(r)}, \quad 1 \leq k \leq n, \quad (5.10)$$

where $(a_k^{(r)})_{k\geq 1}$ is a $\mathbb{R}_+$-valued sequence defined by the following implicit recursive equation:

$$a_0^{(r)} := +\infty, \quad \phi_r\big(-a_{k+1}^{(r)}\big) := \phi_r\big(a_k^{(r)}\big),\ k \geq 0$$

with $\phi_r(x) := \int_0^{x/2} |u|^{r-1}\text{sign}(u)e^{-u}du$ (convention : $0^0 = 1$).

Furthermore, the sequence $(a_k^{(r)})_{k\geq 1}$ decreases to zero and for every $k \geq 1$,

$$a_k^{(r)} = \frac{r+1}{k}\big(1 + \frac{c_r}{k} + O(\frac{1}{k^2})\big)$$

for some real constant c_r.

For numerical examples, Table 2 gives the regression coefficients we obtain by regressing the L^2 grids onto the grids we get with the L^1 and L^4 norms, for different values of n. The notations are the same as in the previous example. We note that for large enough n, the coefficients $\hat{a}_{sr}$ tend to $(s+1)/(r+1) = \theta^\star$. For example, if $n = 900$, we get $\hat{a}_{12} = 0.6676880;\ \hat{a}_{42} = 1.6640023$ whereas the expected values are respectively $2/3 = 0.66666667$ and $5/3 = 1.6666667$. The absolute errors are in the order of 10^{-3}. Like the Gaussian case, we remark that the error of the estimation results mainly from the tail of the exponential distribution.

n	$\hat{a}_{12}$	$\hat{b}_{12}$	ϵ		$\hat{a}_{42}$	$\hat{b}_{42}$	ϵ
20	0.6765013	- 0.0104881	0.0019489		1.6396807	0.0288348	3.081E-33
50	0.6726145	- 0.0082123	0.0045310		1.6502245	0.0225246	1.149E-28
100	0.6706176	- 0.0062439	0.0070734		1.6556979	0.0172020	1.573E-27
300	0.6686428	- 0.0036234	0.0114628		1.6611520	0.0100523	1.508E-27
700	0.6677864	- 0.0022222	0.0146186		1.6635261	0.0061356	1.222E-25
800	0.6676880	- 0.0020482	0.0150735		1.6638043	0.0057199	2.020E-26
900	0.6676079	- 0.0019043	0.0154634		1.6640023	0.0053173	9.683E-25

Table 2: Regression coefficients for exponential distribution.

Conjecture 2. *Let X be an exponentially distributed random variable with rate parameter λ and let $(\alpha_n)_{n\geq 1}$ be an L^r-optimal sequence of quantizers for X. Then for $s > 0$ and $\theta^\star = (s+1)/(r+1)$ the sequence $(\alpha_n^{\theta^\star,0})_{n\geq 1}$ is asymptotically L^s-optimal.*

Remark 5.4. *As a matter of fact, the sequence $(\alpha_n^{\theta^\star,0})_{n\geq 1}$ is not (exactly) L^s-optimal for every $s \neq r$. Otherwise, if $\alpha_{nk}^{(s)} = \frac{s+1}{r+1}\alpha_{nk}^{(r)}$ for every $k \geq 1$ then if follows by backward induction that*

$$\forall k \geq 1, \quad a_k^{(s)} = \frac{s+1}{r+1} a_k^{(r)}.$$

However straightforward calculations show e.g that $a_1^{(2)} = 2$ and $a_1^{(1)} = 2\log(2)$ so that

$$a_1^{(2)} \neq \frac{3}{2} a_1^{(1)}.$$

Moreover, these examples could suggest that a contraction (or a dilatation) parameter $\theta^\star$, solution of the minimization problem (5.1), always leads to a sequence of quantizers satisfying the L^s-empirical measure theorem. The following example shows that this can fail.

5.3 Gamma distribution

5.3.1 Optimal dilatation and contraction

Proposition 5.7. *Let $r, s > 0$ and let P be a Gamma distribution with parameters a and λ : $P = \Gamma(a, \lambda),\ a > 0,\ \lambda > 0$.*

(a) *If $(\alpha_n)_{n\geq 1}$ is an $L^r(P)$-optimal sequence of quantizers then, for $s < r+1$, the sequence $(\alpha_n^{\theta,0})_{n\geq 1}$ is L^s-rate-optimal iff $\theta \in \big(s/(r+1), +\infty\big)$ and for all $a > 0$,*

$$\theta^\star = (s+a)/(r+a)$$

is the unique solution of (5.1) on the set $\big(s/(r+1), +\infty\big)$.

(b) *Let $(\alpha_n)_{n\geq 1}$ be an $L^r(P)$-optimal sequence of quantizers then, if $s > r+1$ and if $a \in \big(0, \frac{s+r+1}{s}\big)$, the sequence $(\alpha_n^{\theta,0})_{n\geq 1}$ is L^s-rate-optimal for every $\theta \in \big(s/(r+1), +\infty\big)$ and*

$$\theta^\star = (s+a)/(r+a)$$

is the unique solution of (5.1) on the set $\big(s/(r+1), +\infty\big)$ (Note that the assumptions imply $a \in (0, 2)$).

(c) *Let $(\alpha_n)_{n\geq 1}$ be an asymptotically $L^r(P)$-optimal sequence of quantizers then, if $s < r$, the sequence $(\alpha_n^{\theta,0})_{n\geq 1}$ is L^s-rate-optimal for every $\theta \in \big(s/r, +\infty\big)$ and for all $a > 0$,*

$$\theta^\star = (s+1)/(r+1)$$

is the unique solution of (5.1) *on the set* $\big(s/r, +\infty\big)$.

Proof. We set $\mu = 0$. The density function reads

$$f(x) = \frac{\lambda^a}{\Gamma(a)} x^{a-1} e^{-\lambda x} \mathbf{1}_{\{x>0\}}, \text{ with } \Gamma(a) = \int_0^{+\infty} x^{a-1} e^{-x} dx.$$

(a) and **(b)**. Let $s \in (0, r+1)$ and set $R_0 = \max(0, (a-1)/\lambda)$. The function f is decreasing on $(R_0, +\infty)$ and for every $\theta > 0, \mu \in \mathbb{R}$, the couple (θ, μ) is P-admissible. For $\theta > s/(r+1)$, Assumption (3.12) holds true for every $c \in \big(1, \theta(1+r)/s\big)$. Moreover, Hypothesis **(H1)** clearly holds. Consequently, it follows from Corollary 3.1, (a) that Assumption (3.8) of Theorem 3.2 holds true.

When $s > r+1$, the additional hypothesis $f^{-\frac{s}{r+1}} \in L^1_{loc}(P)$ holds for $a < \frac{r+1}{s} + 1$. Furthermore, it follows from (5.2) that **(H3)** holds. In this case Assumption (3.8) of Theorem 3.2 holds true.

For all $\theta > 0$,

$$\int_{\mathbb{R}} f(\theta x) f(x)^{-\frac{s}{1+r}} dx = \left(\frac{\lambda^a}{\Gamma(a)}\right)^{1-s/(r+1)} \int_0^{+\infty} x^{(a-1)(1-\frac{s}{r+1})} e^{-(\theta - \frac{s}{r+1})\lambda x} dx,$$

it follows that

$$\int_{\mathbb{R}} f(\theta x) f(x)^{-\frac{s}{r+1}} dx < +\infty \quad \text{iff} \quad \theta > s/(r+1) \text{ and } a(r+1-s) + s > 0.$$

For every $\theta > s/(r+1)$, we have

$$\begin{aligned} h(\theta) &:= \theta^{s+1} \int_{\mathbb{R}} f(\theta x) f(x)^{-\frac{s}{1+r}} dx \\ &= \left(\frac{\lambda^a}{\Gamma(a)}\right)^{1-s/(r+1)} \theta^{s+1} \theta^{a-1} \int_0^{+\infty} x^{(a-1)(1-\frac{s}{1+r})} e^{-(\theta - \frac{s}{1+r})\lambda x} dx \\ &= C\, \theta^\gamma \left(\theta - \frac{s}{1+r}\right)^{-\beta} \end{aligned}$$

with

$$\gamma = s + a \text{ and } \beta = (a-1)(1 - s/(r+1)) + 1.$$

The function h is differentiable for all $\theta > s/(1+r)$ and

$$h'(\theta) = C\theta^{\gamma-1} \left(\theta - \frac{s}{1+r}\right)^{-\beta-1} \left((\gamma-\beta)\theta - \frac{s\gamma}{1+r}\right).$$

The minimum of h is then unique on $\big(s/(r+1), +\infty\big)$ and is reached at $\theta^\star$.

Notice that the condition required for $f^{-\frac{s}{r+1}}$ to be in $L^1_{loc}(P)$ is $a < \frac{r+1}{s} + 1$ and for every $s > r+1$ one has $1 + \frac{r+1}{s} < \frac{s}{s-(r+1)}$. Combined to the condition $a(r+1-s) > 0$ yields the given

constrain on a in (b).

(c) Now let $s < r$ and consider Inequality (3.3). One has

$$\int_{\mathbb{R}} f^{\frac{r}{r-s}}(\theta x) f^{-\frac{s}{r-s}}(x)dx = \frac{\lambda^a}{\Gamma(a)} \int_0^{+\infty} x^{a-1} e^{-\frac{\lambda x}{r-s}(r\theta - s)} dx.$$

Therefore $\int_{\mathbb{R}} f^{\frac{r}{r-s}}(\theta x) f^{-\frac{s}{r-s}}(x)dx < +\infty$ iff $\theta > s/r$.

On the other hand, for every $\theta > s/r$,

$$\begin{aligned} \theta^{1+s} \left(\int_{\mathbb{R}} f_{\theta,\mu}^{\frac{r}{r-s}}(x) f^{-\frac{s}{r-s}}(x)dx \right)^{1-\frac{s}{r}} &= C\, \theta^{s+a} \left(\int_0^{+\infty} x^{a-1} e^{-\frac{\lambda x}{r-s}(r\theta - s)} dx \right)^{\frac{r-s}{r}} \\ &= C\, \theta^{s+a} \big(r\theta - s\big)^{a\frac{s-r}{r}}. \end{aligned}$$

Considering the function h defined by $h(\theta) = \theta^{s+a}\,(r\theta - s)^{a\frac{s-r}{r}}$ we show that h reaches its minimum on $\big(s/r, +\infty\big)$ at the unique point $\theta^\star = (s+a)/(r+a)$. □

Remark 5.5. *Let $X \sim \Gamma(a, \lambda)$. If $s < r$, then $\theta^\star = (s+a)/(r+a) < 1$. Then the sequence $(\alpha_n^{\theta^\star,0})_{n\geq 1}$ is a contraction of $(\alpha_n)_{n\geq 1}$ with scaling number $\theta^\star$. On the other hand, if $s > r$, then $\theta^\star > 1$ and the sequence $(\alpha_n^{\theta^\star,0})_{n\geq 1}$ is a dilatation of $(\alpha_n)_{n\geq 1}$ with scaling number $\theta^\star$. Moreover there is no constraint on the parameter a as long as $s < r$. In this case when we set $a = 1$ (exponential distribution with parameter λ) we retrieve the result related to the exponential distribution. We notice that $\theta^\star$ does not depend on the parameter λ. That is expected since $\Gamma(1, \lambda) = \mathscr{E}(\lambda)$ and, in the exponential case we know that the scaling number does not depend on λ.*

Let $\theta^\star = (s+a)/(r+a)$ and consider now the sequence $(\alpha_n^{\theta^\star,0})_{n\geq 1}$ defined as previously. Does this sequence verify the L^s-empirical measure theorem? If $a = 1$ we boil down to the exponential distribution. On the other hand, when $a \neq 1$, one shows below that there exists $a > 1,\ s > 0$ and $r > 0$ such that the sequence $(\alpha_n^{\theta^\star,0})_{n\geq 1}$ does not verify the L^s-empirical measure theorem.

Suppose that $(\alpha_n^{\theta^\star,0})_{n\geq 1}$ satisfies the L^s-empirical measure theorem. Then we must have, for all $u \in \mathbb{R}_+$,

$$\frac{1}{C_{f,r}} \frac{1}{\theta^\star} \int_0^u f(x/\theta^\star)^{\frac{1}{1+r}} dx = \frac{1}{C_{f,s}} \int_0^u f(x)^{\frac{1}{1+s}} dx. \tag{5.11}$$

with $f(x) = \frac{\lambda^a}{\Gamma(a)} x^{a-1} e^{-\lambda x} \mathbf{1}_{\{x>0\}}$ and $C_{f,r} = \int f(x)^{\frac{1}{1+r}} dx$ for all $r > 0$.

However, we have for any $r > 0$,

$$\begin{aligned} C_{f,r} &= \lambda^{\frac{a}{1+r}} \Gamma(a)^{-\frac{1}{1+r}} \int_0^{+\infty} x^{(a-1)/(r+1)} e^{-\frac{\lambda}{1+r}x} dx \\ &= \lambda^{\frac{a}{1+r}} \Gamma(a)^{-\frac{1}{1+r}} \int_0^{+\infty} x^{(r+a)/(r+1)-1} e^{-\frac{\lambda}{1+r}x} dx \\ &= \lambda^{\frac{a}{1+r}} \Gamma(a)^{-\frac{1}{1+r}} \Gamma\left(\frac{r+a}{r+1}\right) \lambda^{-\frac{r+a}{r+1}} \big(r+1\big)^{\frac{r+a}{r+1}} \\ &= \Gamma\left(\frac{r+a}{r+1}\right) \Gamma(a)^{-\frac{1}{1+r}} \lambda^{-\frac{r}{r+1}} \big(r+1\big)^{\frac{r+a}{r+1}}. \end{aligned}$$

Equation (5.11) is written down for all $u \in \mathbb{R}_+$,

$$C(r) \left(\frac{r+a}{s+a}\right)^{\frac{r+a}{r+1}} \int_0^u x^{\frac{a-1}{r+1}} e^{-\frac{\lambda(r+a)}{(r+1)(s+a)}x} dx = C(s) \int_0^u x^{\frac{a-1}{s+1}} e^{-\frac{\lambda}{s+1}x} dx$$

with $C(r)=\Gamma\left(\frac{r+a}{r+1}\right)^{-1}\lambda^{\frac{r+a}{r+1}}\big(r+1\big)^{-\frac{r+a}{r+1}}, \qquad \forall\, r>0.$

Let $m\in\mathbb{N}$ and $\alpha>0$. We show by induction that, for $u>0$,

$$\int_0^u x^n e^{-\alpha x}dx=-\left(\frac{1}{\alpha}u^n+\frac{n}{\alpha^2}u^{n-1}+\frac{n(n-1)}{\alpha^3}u^{n-2}+\cdots+\frac{n!}{\alpha^n}u+\frac{n!}{\alpha^{n+1}}\right)e^{-\alpha u}+\frac{n!}{\alpha^{n+1}}.$$

Let us consider $a>1$ such that $\frac{a-1}{r+1}$ and $\frac{a-1}{s+1}$ are integers and set $n=\frac{a-1}{r+1}$, $m=\frac{a-1}{s+1}$, $\alpha=\frac{\lambda(r+a)}{(r+1)(s+a)}$ and $\beta=\frac{\lambda}{s+1}$. Equation (5.11) finally reads

$$\begin{aligned}C(r)\ \left(\frac{r+a}{s+a}\right)^{\frac{r+a}{r+1}}&\left[\left(\frac{1}{\alpha}u^n+\frac{n}{\alpha^2}u^{n-1}+\frac{n(n-1)}{\alpha^3}u^{n-2}+\cdots+\frac{n!}{\alpha^n}u+\frac{n!}{\alpha^{n+1}}\right)e^{-\alpha u}-\frac{n!}{\alpha^{n+1}}\right]\\ =C(s)&\left[\left(\frac{1}{\beta}u^m+\frac{m}{\beta^2}u^{m-1}+\frac{m(m-1)}{\beta^3}u^{m-2}+\cdots+\frac{m!}{\beta^m}u+\frac{m!}{\beta^{m+1}}\right)e^{-\beta u}-\frac{m!}{\beta^{m+1}}\right].\end{aligned}$$

Set $a=7,\ s=1,\ r=2,\ \lambda=1$ and $u=1$. Then $n=2, m=3, \alpha=3/8, \beta=1/2$ and we show after some computations that the sequence $(\alpha_n^{\theta^\star,0})_{n\geq 1}$ does not satisfy the L^s-empirical measure theorem since

$$\frac{185}{128}e^{-3/8}-\frac{79}{48}e^{-1/2}\neq-\frac{511}{512}$$

(one side is rational, the other is not). Hence, we have constructed an $L^s(P)$-rate-optimal sequence which does not satisfy the L^s-empirical measure theorem.

6 Applications

6.1 Application to Lloyd's I algorithm

One of the important issues from a computational point of view is the search of the L^r-optimal quantizers. The quadratic case ($r=2$) is the commonly implemented for applications and various algorithms like the Competitive Learning Vector Quantization (CLVQ) algorithm (see e.g [11]) and "randomized versions" of the Lloyd's algorithms (I and II) (see e.g. [4]) are used. In practice, Lloyd's I algorithm is widely used to compute stationary (or optimal) quantizers because it can be easily implemented. We will use the natural extension of Lloyd's I algorithm to compute the L^r-stationary (optimal) quantizers. In a general framework, L^r-stationary quantizers (α_n) (with $\alpha_n=(\alpha_{n1},\cdots,\alpha_{nn})$) are computed using the L^r-stationary equation $\nabla e_{n,r}(X)=0$. This equation reads for every $r\geq 1$ (see [11])

$$\alpha_{ni}=\frac{\mathbb{E}\left(\mathbf{1}_{X\in C_i(\alpha_n)}|X-\alpha_{ni}|^{r-2}X\right)}{\mathbb{E}\left(\mathbf{1}_{X\in C_i(\alpha_n)}|X-\alpha_{ni}|^{r-2}\right)},\qquad i=1,\cdots,n. \tag{6.1}$$

The Lloyd's I procedure is the fixed point procedure derived from (6.1). Starting with an initial quantizer $\alpha_n^{(0)}$ of size n, one defines recursively a sequence $(\alpha_n^{(l)})_{l=1,\cdots,L}$ of L^r-stationary quantizers (where L corresponds to the number of Lloyd's iterations) by setting for every $l=1,\cdots,L$,

$$\alpha_{ni}^{(l)}=\frac{\mathbb{E}\left(\mathbf{1}_{X\in C_i(\alpha_n)}|X-\alpha_{ni}^{(l-1)}|^{r-2}X\right)}{\mathbb{E}\left(\mathbf{1}_{X\in C_i(\alpha_n)}|X-\alpha_{ni}^{(l-1)}|^{r-2}\right)},\qquad i=1,\cdots,n. \tag{6.2}$$

By "randomized version" of the Lloyd's I procedure we mean that both expectations in (6.2) are computed using a Monte Carlo simulation of size M. However in higher dimensions there are several

L^r-stationary quantizers and the Lloyd's I procedure is somehow a method to compute the "nearest" one (namely the one in the attracting basin of which the procedure has been initialized). This is why the initialization of the procedure at already good grid is a crucial issue to obtain good L^r-quantizers. This leads us to consider the optimally L^r-dilated quadratic quantizers as natural good candidates to initialize the L^r-Lloyd's I procedure. We compare it to a random initialization, which consists on generating a vector of size n distributed as X and multiply it by the same scaled number ($\theta^\star$).

We carried out a numerical test in dimension $d = 2, 3$ with $r = 4$ for the $\mathcal{N}(0; I_d)$ distribution. The Monte Carlo size M is equal to 10^6 and our grid size n moves 10 by 10 from 10 to 100. Numerical results depicted in figure 2 (for $d = 2$) show that the dilated L^2-stationary (optimal) grids are already almost L^4-stationary (and likely almost optimal if we suppose that the L^2-stationary quantizers are) since the initial scaled L^2-stationary grids do not move during the successive Lloyd's iterations. This is also confirmed by Figure 1 (when $d = 2$) and Figure 3 (when $d = 3$) where the logarithm of the L^4 quantization error of the dilated grids remains the same during the successive Lloyd's procedures. The dilated L^2-stationary quantizers initialization seems to be the best choice one can do. The Lloyd's procedure initialization with random grids never leads to lower quantization errors. Moreover it needs several iterations of the procedure, depending to the grid size.

6.2 Application of L^r-quantizers to numerical integration

Let $\beta \geq 1$ and let $X \in L^{\beta+\varepsilon_0}(\mathbb{P})$, $\varepsilon_0 > 0$. Let $f \in \mathrm{Lip}_\beta(\mathbb{R}^d) := \{g : \mathbb{R}^d \to \mathbb{R},\ |g(x) - g(y)| \leq C|x - y|(1 + |x|^{\beta-1} + |y|^{\beta-1})\}$. For any sequence of quantizers $(\alpha_n)_{n\geq 1}$ and any $r \in [1, +\infty]$ we have

$$\begin{aligned}\|f(X) - f(\widehat{X}^{\alpha_n})\|_1 &\leq C\,\mathbb{E}(|X - \widehat{X}^{\alpha_n}|(1 + |X|^{\beta-1} + |\widehat{X}^{\alpha_n}|^{\beta-1})) \\ &\leq C\,\|X - \widehat{X}^{\alpha_n}\|_r(1 + \|X\|^{\beta-1}_{(\beta-1)r'} + \|\widehat{X}^{\alpha_n}\|^{\beta-1}_{(\beta-1)r'}) \qquad (6.3)\end{aligned}$$

by *Hölder*'s inequality with $r' = r/(r - 1)$.

Suppose now that $X \notin L^{\beta+\varepsilon}(\mathbb{P})$, $\forall \varepsilon > \varepsilon_0$. To give a sense to the above inequality as a error bound, we must choose r' such that $(\beta - 1)r' \leq \beta + \varepsilon_0$; which in return impose that $r > \frac{\beta+\varepsilon_0}{1+\varepsilon_0}$. Now $\frac{\beta+\varepsilon_0}{1+\varepsilon_0} > 2$ as soon as $\beta > 2 + \varepsilon_0$. Furthermore if $\liminf_{|x|\to+\infty} \frac{|f(x)|}{|x|^\beta} > 0$ there is no alternative to these constraints. In this situation it is impossible to use quadrature formulae for numerical integration based on quadratic quantizers. However we can use some dilated L^2-optimal (at least stationary) quantizers $\alpha_n^{\theta^\star,\mu}$, for large enough n. Then

$$\mathbb{E}\big(f(\widehat{X}^{\alpha_n^{\theta^\star,\mu}})\big) = \sum_{i=1}^n f(\alpha_{ni}^{\theta^\star,\mu})\mathbb{P}(X \in C_i(\alpha_n^{\theta^\star,\mu}))$$

and

$$|\mathbb{E}f(X) - \mathbb{E}f\big(\widehat{X}^{\alpha_n^{\theta^\star,\mu}}\big)| \leq \|\mathbb{E}f(X) - \mathbb{E}f\big(\widehat{X}^{\alpha_n^{\theta^\star,\mu}}\big)\|_1$$

so that the error bound (6.3) holds true profided f is lipschitz. Such an approach requires to compute the weights $\mathbb{P}(X \in C_i(\alpha_n^{\theta^\star,\mu}))$ associated to the Voronoi cells of $\alpha_n^{\theta^\star,\mu}$. The following easy proposition says how to compute these weights.

Proposition 6.1. *Let $X \sim P$ and $P = f \cdot \lambda_d$. Then, for every $n \geq 1$, we have*

$$\forall i \in \{1, \cdots, n\}, \quad P(C_i(\alpha_n^{\theta,\mu})) = P_{\theta,\mu}(C_i(\alpha_n)). \qquad (6.4)$$

Proof. Let $n \geq 1$ and $\alpha_n = (\alpha_{n1}, \cdots, \alpha_{nn})$. One has

$$P(C_i(\alpha_n^{\theta,\mu})) = \mathbb{P}(X \in C_i(\alpha_n^{\theta,\mu})) = \int_{\{|x-(\mu+\theta(\alpha_{ni}-\mu))|=\min_{j\neq i}|x-(\mu+\theta(\alpha_{nj}-\mu))|\}} f(x)d\lambda_d(x).$$

Making the change of variable $z = \frac{x-\mu}{\theta} + \mu$ yields

$$P(C_i(\alpha_n^{\theta,\mu})) = \theta^d \int_{\{z\in C_i(\alpha_n)\}} f_{\theta,\mu}(z)d\lambda_d(z) = P_{\theta,\mu}(C_i(\alpha_n)).$$

□

When a closed formula (like for the exponential distribution) is not available for the weights of the dilated cells, these weights can be estimated by the Monte Carlo method using the Nearest-Neighbour algorithm. Fast implementations of this algorithm can be find e.g. in [3, 8].

Acknowledgements

I want to express my sincere thanks to my advisor, Professor Gilles Pagès, for all his help and advices during the preparation of this work.

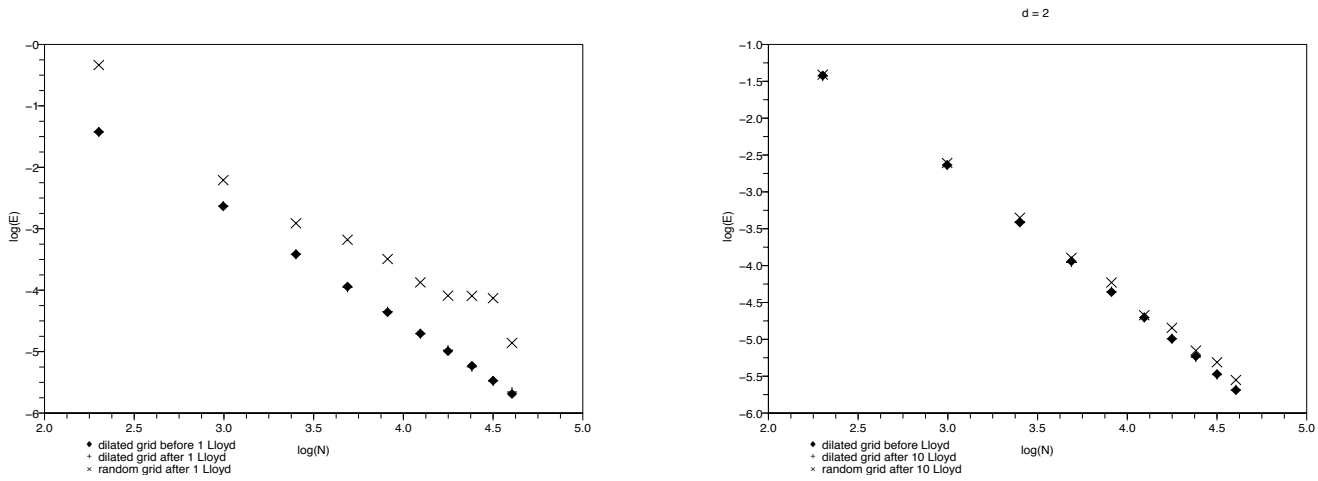

Figure 1: Comparison of the log of the L^4-error (power 4) as function of the log of the grid size after 1 (on the left) and 10 (on the right) Lloyd's iterations.

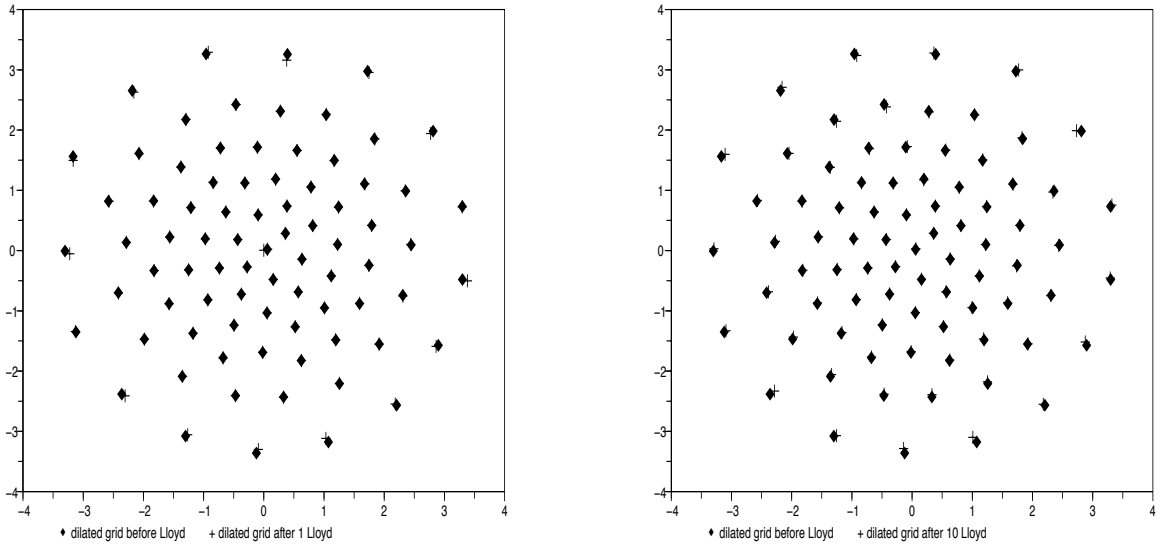

Figure 2: The L^2-dilated grid before and after 1 and 10 Lloyd's iterations; the grid size equals 80.

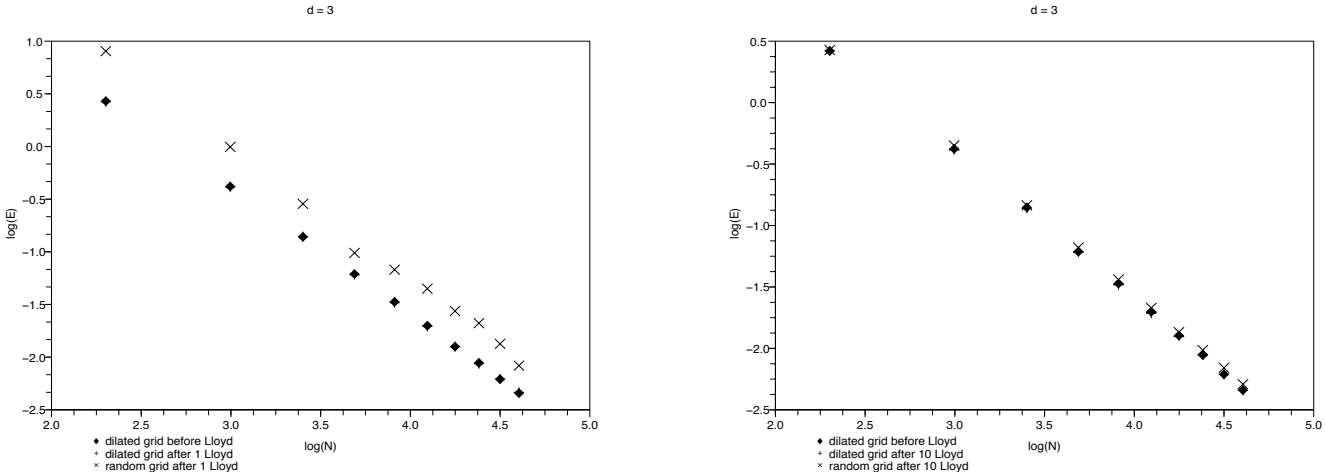

Figure 3: Comparison of the log of the L^4-error (power 4) as function of the log of the grid size after 1 (on the left) and 10 (on the right) Lloyd's iterations.

References

[1] Delattre, S., Graf, S., Luschgy, H., Pagès, G., Quantization of probability distributions under norm-based distribution measures, *Statistics & Decisions*, 22, 261-282 (2004).

[2] Fort, J.C. , Pagès G., Asymptotics of optimal quantizers for some scalar distributions, *Journal of Computational and Applied Mathematics*, 146, 253-275 (2002).

[3] Friedman, J. H., Bentley, J.L., Finkel R.A., An Algorithm for Finding Best Matches in Logarithmic Expected Time, *ACM Transactions on Mathematical Software*, 3(3), 209-226 (1977).

[4] Gersho, A., Gray, R., *Vector Quantization and Signal Compression*, Kluwer, Boston, 732p (6^{th} edition).

[5] Graf, S., Luschgy, H., *Foundations of Quantization for Probability Distributions*, Lecture Notes in Mathematics, Vol 1730, Berlin, 2000.

[6] Graf, S., Luschgy, H., Pagès G., Distorsion mismatch in the quantization of probability measures, *ESAIM PS*, 12, 127-153 (2008).

[7] Matilla, P., *Geometry of Sets and Measures in Euclidean Spaces*, Cambridge Unversity Press.

[8] McNames, J., A Fast Nearest-Neighbor algorithm based on a principal axis search tree, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(9), 964-976 (2001).

[9] Pagès G., Space vector quantization method for numerical integration, *Journal of Computational and Applied Mathematics*, 89, 1-38,1998.

[10] Pagès G., Pham H., Printems J., An Optimal markovian quantization algorithm for multidimensional stochastic control problems, *Stochastics and Dynamics*, 4(4), 501-545 (2004).

[11] Pagès G., Pham H., Printems J., Optimal quantization methods and applications to numerical problems in finance, *Handbook on Numerical Methods in Finance* (S. Rachev, ed.), Birkhauser, Boston, 253-298, 2004.

[12] Pagès, G., Printems, J., Optimal quadratic quantization for numerics: the Gaussian case, *Monte Carlo Methods and Appl.*, 9(2), 135-165.