Greedy vector quantization

Harald Luschgy, Gilles Pagès

Greedy vector quantization

Harald Luschgy [*] and Gilles Pagès [†]

Abstract

We investigate the greedy version of the L^p-optimal vector quantization problem for an $\mathbb{R}^d$-valued random vector $X \in L^p$. We show the existence of a sequence $(a_N)_{N\geq 1}$ such that a_N minimizes $a \mapsto \big\| \min_{1\leq i\leq N-1} |X-a_i| \wedge |X-a| \big\|_{L^p}$ (L^p-mean quantization error at level N induced by $(a_1,\ldots,a_{N-1},a)$). We show that this sequence produces L^p-rate optimal N-tuples $a^{(N)} = (a_1,\ldots,a_{_N})$ (*i.e.* the L^p-mean quantization error at level N induced by $a^{(N)}$ goes to 0 at rate $N^{-\frac{1}{d}}$). Greedy optimal sequences also satisfy, under natural additional assumptions, the distortion mismatch property: the N-tuples $a^{(N)}$ remain rate optimal with respect to the L^q-norms, $p \leq q < p+d$. Finally, we propose optimization methods to compute greedy sequences, adapted from usual Lloyd's I and Competitive Learning Vector Quantization procedures, either in their deterministic (implementable when $d=1$) or stochastic versions.

Keywords : Optimal Vector Quantization ; greedy optimization ; distortion mismatch ; Lloyd's I procedure ; Competitive Learning Vector Quantization.

2010 AMS Classification: 60G15, 60G35, 41A25.

1 Introduction and definition of greedy quantization sequences

Let $p\in(0,+\infty)$. We consider $X:(\Omega,\mathcal{A},\mathbb{P})\to\mathbb{R}^d$ an L^p-integrable random vector. For every $\Gamma\subset\mathbb{R}^d$, we define the L^p-mean quantization error induced by Γ as the L^p-mean of the distance of the random vector X to the subset Γ (with respect to a norm $|\,.\,|$ on $\mathbb{R}^d$)

$$e_p(\Gamma,X) = \big\| d(X,\Gamma)\big\|_p.$$

This quantity is always finite when $X\in L^p(\mathbb{P})$ since $e_p(\Gamma,X)\leq \|X\|_p + \min_{a\in\Gamma}|a| < +\infty$ owing to Minkowski's inequality. The usual L^p-optimal quantization problem *at level* $N\geq 1$ is to solve the following minimization problem

$$e_{p,N}(X) = \min_{\Gamma\subset\mathbb{R}^d, |\Gamma|\leq N} e_p(\Gamma,X)$$

where $|\Gamma|$ denotes the cardinality of the subset Γ, sometimes called *grid* in Numerical Probability or *codebook* in Signal processing. The use of "min" instead of "inf" is justified by the fact (see Proposition 4.12 in [19], p.47 or [26]) that this infimum is always attained by an *optimal quantization* grid $\Gamma^{(N)}$ (of full size N if the support of the distribution $\mu=\mathbb{P}_{_X}$ of X has at least N elements).

The most celebrated result in Optimal (Vector) Quantization Theory is undoubtedly Zador's Theorem which rules the *sharp* asymptotic rate of convergence of $e_{p,N}(X)$ as the *quantization level* N (or grid size) goes to infinity.

[*]Universität Trier, FB IV-Mathematik, D-54286 Trier, Germany. E-mail: `luschgy@uni-trier.de`

[†]Laboratoire de Probabilités et Modèles Aléatoires, UMR 7599, UPMC, case 188, 4, pl. Jussieu, F-75252 Paris Cedex 5, France. E-mail: `gilles.pages@upmc.fr`

Theorem 1.1 ((Zador's Theorem), see [19], Theorem 6.4, p.78, see also [21]). *If $\mathbb{E}|X|^p < +\infty$ and $\mu = \mathbb{P}_X = \varphi.\lambda_d + \nu$ where ν is a singular Borel measure with respect to the Lebesgue measure λ_d on $\mathbb{R}^d$. Then*

$$\liminf_N N^{\frac{1}{d}} e_{p,N}(X) \geq \widetilde{J}_{p,d} \left(\int_{\mathbb{R}^d} \varphi^{\frac{d}{p+d}} d\lambda_d \right)^{\frac{1}{p}+\frac{1}{d}}$$

where $\widetilde{J}_{p,d}$ is the sharp limit for the uniform distribution $U([0,1]^d)$ over the unit hypercube which satisfies

$$\widetilde{J}_{p,d} = \inf_N N^{\frac{1}{d}} e_{p,N}\big(U([0,1]^d)\big) \in (0,+\infty).$$

If furthermore $\mathbb{E}|X|^{p+\delta} < +\infty$ or some $\delta > 0$, then

$$\lim_N N^{\frac{1}{d}} e_{p,N}(X) = \widetilde{J}_{p,d} \left(\int_{\mathbb{R}^d} \varphi^{\frac{d}{p+d}} d\lambda_d \right)^{\frac{1}{p}+\frac{1}{d}}.$$

The starting idea of *greedy quantization* is to determine a *sequence* $(a_N)_{N\geq 1}$ of points of $\mathbb{R}^d$ which is recursively optimal *step by step* or *level by level* with respect to the L^p-mean quantization criterion. We mean that, if we set $a^{(N)} = \{a_1, \ldots, a_N\}$, $N \geq 1$, and $a^{(0)} = \emptyset$, then

$$\forall\, N \geq 0, \quad a_{N+1} \in \operatorname{argmin}_{\xi\in\mathbb{R}^d} e_p(a^{(N)} \cup \{\xi\}, X). \tag{1.1}$$

Note that a_1 is simply an L^p-median of (the distribution of) X and that, when $p > 1$, a strict convexity argument implies the uniqueness of this L^p-median. This idea to design not only optimal N-tuples but an optimal sequence which, hopefully, will produce N-tuples with a rate optimal behavior as $N \to +\infty$ is very natural and can be compared to sequences with low discrepancy in Quasi-Monte Carlo methods.

In fact, such sequences have already been investigated in an L^1 setting for compactly supported random vectors X as a model of short term experiment planning *vs* long term experiment planning represented by regular optimal quantization at a given level N (see [10]). Our aim in this paper is to solve this *greedy* optimization problem for as general as possible distributions $\mu = \mathbb{P}_X$ and in any L^p-space, $p \in (0,+\infty)$, in two directions: first establish the existence of such *L^p-optimal greedy sequences* and then evaluate their rate of decay of $e_p(a^{(N)}, X)$ to 0 as the quantization level N goes to infinity.

A possible wider field of applications is to substitute such sequences to optimal$N$-quantizers in the quantization based numerical schemes that have been developed in the early 2000's. In these procedures optimal quantizations used as a spatial discretization method that "fits" optimal the distribution of interest at each time step. Among theses application, often in connection with Finance but also with reliability, we may mention Numerical integration (see [26, 30]), Optimal Stopping Theory (pricing of American style or callable derivatives, see [2, 3, 4]), Stochastic control of diffusions and portfolio optimization, (see [27, 33, 12] or control of PDMP([1], for reliability (see [8, 9]), non-linear filtering and stochastic volatility models ([27]), discretization of $BSDE$s and Stochastic PDEs [18]. See also the review papers [28, 31] for more details and the references therein. In most of these applications, up to some variant, an $\mathbb{R}^d$-valued discrete time Markov chain $(X_k)_{0\leq k\leq n}$ is approximated path wise and in distribution by its quantized approximation sequence $(\widehat{X}_k^{\Gamma_k})_{0\leq k\leq n}$ living on a *quantization tree* made up by the *optimal quantization grids* Γ_k (of varying sizes N_k) and the transitions matrices $\pi^k = \mathcal{L}\big(\widehat{X}^{\Gamma_{k+1}} \,|\, \widehat{X}^{\Gamma_k}\big)$ which discretize the Markov dynamics of the chain. The quantization based scheme turns out to be in many cases spatial discretization of a (Backward) Dynamical Programming principle. Given the common sizes of the grids in these implemented procedures (N_k is often greater than 1 000) and the number n of time steps ($n \geq 10$ and sometimes equal to 100) the storing of

[1] Piecewise Constant Deterministic Markov Processes introduced by M. Davis in [13].

this quantization tree may exceed the storage capacity of the computing device. Using the induced grids $a^{(N_0)}, a^{(N_1)} \ldots, a^{(N_n)}$ induced by a greedy optimal sequence $(a_N)_{N\geq 1}$ will dramatically reduce this drawback, provided that, on the other hand, their rate of decay of their mean quantization rates remain comparable to those of optimal quantizers.

The paper is organized as follows: in Section 2, the existence of (L^p,μ)-optimal greedy sequences and their first properties are established for general and Euclidean norms. In Section 3, (L^p,μ)-optimal greedy sequences are shown to be rate optimal in terms of mean quantization error, compared to sequences of $L^p(\mu)$-optimal N-quantizers. We also solve - positively – the so-called *distortion mismatch problem* *i.e.* the property that the above rate optimal decay property remains true for the $L^q(\mu)$-mean quantization error when $q \in [p, p+d)$ in a d-dimensional setting (and sometimes for $q = p+d$). In Section 4, easy-to-check criterions, mostly borrowed from [21], are adapted to our greedy framework. Section 5 is devoted to some further questions about the asymptotic behaviour of L^p-greedy sequences, compared to L^p-optimal N-quantizers or non-greedy L^p-rate optimal sequences. In Section 6, we propose numerical procedures to compute quadratic optimal greedy sequences in both 1 and higher dimensional settings, either by deterministic means or by simulation. Finally, we propose in Section 7, when X is uniformly distributed on the unit hypercube $[0,1]^d$, a comparison between optimal greedy sequences and the *sequences with low discrepancy* popularized by the Quasi-Monte Carlo method.

NOTATIONS: • $\mathbb{N}^* = \{1,2,\ldots\}$ the set of positive integers.

• $|\,.\,|$ denotes any norm on $\mathbb{R}^d$ (except specific mention). For every subset $A \subset \mathbb{R}^d$ and $\xi \in \mathbb{R}^d$, $d(\xi, A) = \inf_{a\in A} |\xi - a|$ (distance of ξ to the set A in $(\mathbb{R}^d, |\,.\,|)$).

2 Existence of optimal greedy quantization sequences

Proposition 2.1. (a) Existence: *If $X \in L^p(\mathbb{P})$, then the sequence of optimization problems (1.1) admits at least one solution $(a_N)_{N\geq 1}$ where a_1 is the L^p-median of the distribution μ. Moreover, the finite sequence $\big(e_p(a^{(n)}, X)\big)_{1\leq n\leq N}$ is (strictly) decreasing as long as $N \leq \mathrm{card}\big(\mathrm{supp}(\mu)\big)$. In particular, $a_n \notin a^{(n-1)}$, for $n\in\{1,\ldots,N\}$.*

*Any such a solution is called an L^p-*optimal greedy quantization sequence.

(b) Local optimality: *As long as $N \leq \mathrm{card}\big(\mathrm{supp}(\mu)\big)$*

$$\mu\big(C_{a_N}(a^{(N)})\big) > 0 \;\; \textit{where } C_{a_N}(a^{(N)}) = \big\{\xi\in\mathbb{R}^d \,|\, |\xi - a_N| < \min_{1\leq i\leq N-1} |\xi - a_i|\big\}$$

and for any Borel set C such that $C_{a_N}(a^{(N)}) \subset C \subset W_{a_N}(a^{(N)}) = \big\{\xi\in\mathbb{R}^d \,|\, |\xi - a_N| \leq \min_{1\leq i\leq N-1} |\xi - a_i|\big\}$, a_N is solution to the local optimization problem

$$a_N \in \mathrm{argmin}_{a\in\mathbb{R}^d}\mathbb{E}\big(|X-a|^p \,|\, X\in C\big).$$

(c) Space filling: *Assume $X \in L^q_{\mathbb{R}^d}(\mathbb{P})$ for some $q \geq p$. Then, any L^p-optimal greedy quantization sequence $(a_N)_{N\geq 1}$ satisfies*

$$\lim_N e_q(a^{(N)}, X) = 0$$

i.e., equivalently, $\displaystyle\lim_{N\to+\infty} \int_{\mathbb{R}^d} \min_{1\leq i\leq N} |\xi - a_i|^q \mu(d\xi) = 0$. *In particular* $\lim_N e_p(a^{(N)}, X) = 0$.

Proof. (a) We proceed by induction. When $N = 1$, the existence of a_1 is obvious once noticed that $\xi \mapsto \mathbb{E}\,|X-\xi|^p$ is continuous and goes to infinity as $|\xi| \to +\infty$. Assume there exists $a_1, \ldots a_N$ such that $e_p(a^{(k)}, X) = \min_{a\in\mathbb{R}^d} e_p(a^{(k-1)} \cup \{a\}, X)$ for every $k\in\{1,\ldots,N\}$.

If $\mathrm{supp}(\mu) \subset \{a_1, \ldots, a_N\}$ then for every $a \in \mathbb{R}^d$, $e_p(a^{(N)} \cup \{a\}, X) = e_p(a^{(N)}, X)$. Otherwise, let $\xi^* \in \mathrm{supp}(\mu) \setminus \{a_1, \ldots, a_N\}$. It is clear that $|\xi - \xi^*| < d(\xi, a^{(N)})$ on the ball $B\big(\xi^*, \frac{1}{4}d(\xi^*, a^{(N)})\big)$ which satisfies $\mu\big(B\big(\xi^*, \frac{1}{4}d(\xi^*, a^{(N)})\big)\big) > 0$. Consequently, $e_p(a^{(N)} \cup \{\xi^*\}, X) < e_p(a^{(N)}, X)$. Now let

$$K^0_{N+1} = \big\{\xi \in \mathbb{R}^d \,|\, e_p(a^{(N)} \cup \{\xi\}, X) \leq e_p(a^{(N)} \cup \{\xi^*\}, X)\big\}.$$

This is a closed non-empty set. Now let $(\xi_k)_{k\geq 1}$ be a sequence of elements of K^0_{N+1} such that $|\xi_k| \to +\infty$. It follows from Fatou's Lemma that

$$\begin{aligned}
\liminf_k e_p(a^{(N)} \cup \{\xi_k\})^p &\geq \int_{\mathbb{R}^d} \liminf_k \big(d(\xi, a^{(N)})^p \wedge |\xi - \xi_k|^p\big)\mu(d\xi) \\
&= \int_{\mathbb{R}^d} d(\xi, a^{(N)})^p \mu(d\xi) \\
&= e_p(a^{(N)}, X)^p > e_p(a^{(N)} \cup \{\xi^*\}, X)^p.
\end{aligned}$$

This yields a contradiction which in turn implies that K^0_{N+1} is a compact set. On the other hand $\xi \mapsto e_p(a^{(N)} \cup \{\xi\}, X)$ is clearly Lipschitz continuous on $\mathbb{R}^d$, hence it attains its minimum on K^0_{N+1} which is clearly its absolute minimum.

(b) If $\mu\big(C_{a_N}(a^{(N)})\big) = 0$, then one checks that

$$e_p(a^{(N-1)}, X) - e_p(a^{(N)}, X) = \int_{C_{a_N}(a^{(N)})} \big(d(\xi, a^{(N-1)})^p - |x - a_N|^p\big)\mu(d\xi) = 0$$

which contradicts the strict decreasing monotony of $e_p(a^{(N)}, X)$. Let $(C_i)_{1\leq i\leq N}$ be a Borel Voronoi partition of $\mathbb{R}^d$ induced by $a^{(N)}$, *i.e.* satisfying $C_i \subset W_{a_i}(a^{(N)}) = \big\{\xi \in \mathbb{R}^d \,|\, |\xi - a_i| = \min_{1\leq j\leq N} |\xi - a_j|\big\}$, and such that $C_N = C$. Assume there exists $b \in C$ such that $\int_C |\xi - a_N|^p \mu(d\xi) > \int_C |\xi - b|^p \mu(d\xi)$. Then

$$\begin{aligned}
e_p(a^{(N)}, X)^p &= \sum_{i=1}^{N-1} \int_{C_i} |\xi - a_i|^p \mu(d\xi) + \int_C |\xi - a_N|^p \mu(d\xi) \\
&\geq \sum_{i=1}^{N-1} \int_{C_i} d\big(\xi, a^{(N-1)} \cup \{b\}\big)^p \mu(d\xi) + \int_C |\xi - a_N|^p \mu(d\xi) \\
&> \sum_{i=1}^{N-1} \int_{C_i} d\big(\xi, a^{(N-1)} \cup \{b\}\big)^p \mu(d\xi) + \int_C |\xi - b|^p \mu(d\xi) \qquad \text{since } \mu(C) > 0 \\
&\geq \sum_{i=1}^{N-1} \int_{C_i} d\big(\xi, a^{(N-1)} \cup \{b\}\big)^p \mu(d\xi) + \int_C d\big(\xi, a^{(N-1)} \cup \{b\}\big)^p \mu(d\xi) \\
&= e_p(a^{(N-1)} \cup \{b\}, X)^p
\end{aligned}$$

which contradicts the minimality of a_N.

(c) Let $p \in (0, +\infty)$. It is clear that, for every $\xi \in \mathbb{R}^d$, $\min_{1\leq i\leq N} |\xi - a_i|$ is non-increasing and converges toward $\inf\limits_{N\geq 1} |\xi - a_N|$ so that by the monotone convergence theorem, one has

$$e_p(a^{(N)}, X)^p \downarrow \ell_\infty := \int_{\mathbb{R}^d} \inf_{i\geq 1} |\xi - a_i|^p \mu(d\xi).$$

Let $a^{(\infty)} = \{a_N,\, N \geq 1\}$. If $\ell_\infty \neq 0$, then there exists $\xi_0 \in \mathrm{supp}(\mu)$ such that $\varepsilon_0 = d(\xi_0, a^{(\infty)}) > 0$. Then, for every $\xi \in B(\xi_0, \frac{\varepsilon_0}{4})$, $d(\xi, a^{(\infty)}) \geq \frac{3}{4}\varepsilon_0$ so that

$$\int_{B(\xi_0, \frac{\varepsilon_0}{4})} d\big(\xi, a^{(\infty)}\big)^p \mu(d\xi) \geq \eta_0 \quad \text{with} \quad \eta_0 = \Big(\frac{3\varepsilon_0}{4}\Big)^p \mu\Big(B\big(\xi_0, \frac{\varepsilon_0}{4}\big)\Big).$$

Now, let N_0 be a positive integer such that,

$$\int_{\mathbb{R}^d} d\big(\xi, a^{(N_0)}\big)^p \mu(d\xi) \leq \ell_\infty + \frac{\eta_0}{2}\Big(1 - \frac{1}{3^p}\Big).$$

We consider the (N_0+1)-quantizer $a^{(N_0)} \cup \{\xi_0\}$. On the one hand,

$$\int_{B(\xi_0, \frac{\varepsilon_0}{4})} d\big(\xi, \{a_1, \ldots, a_{N_0}, \xi_0\}\big)^p \mu(d\xi) \leq \Big(\frac{\varepsilon_0}{4}\Big)^p \mu\Big(B\big(\xi_0, \frac{\varepsilon_0}{4}\big)\Big) = \frac{\eta_0}{3^p}$$

and, on the other hand

$$\begin{aligned}
\int_{{}^cB(\xi_0, \frac{\varepsilon_0}{4})} d\big(\xi, a^{(N_0)} \cup \{\xi_0\}\big)^p \mu(d\xi) &\leq \int_{{}^cB(\xi_0, \frac{\varepsilon_0}{4})} d\big(\xi, a^{(N_0)}\big)^p \mu(d\xi) \\
&\leq \int_{\mathbb{R}^d} d\big(\xi, a^{(N_0)}\big)^p \mu(d\xi) - \int_{B(\xi_0, \frac{\varepsilon_0}{4})} d\big(\xi, a^{(N_0)}\big)^p \mu(d\xi) \\
&\leq \ell_\infty + \frac{\eta_0}{2}\Big(1 - \frac{1}{3^p}\Big) - \eta_0
\end{aligned}$$

so that

$$\int_{\mathbb{R}^d} d\big(\xi, a^{(N_0)} \cup \{\xi_0\}\big)^p \mu(d\xi) \leq \ell_\infty + \frac{\eta_0}{2}\Big(1 - \frac{1}{3^p}\Big) - \eta_0 + \frac{\eta_0}{3^p} < \ell_\infty$$

which yields a contradiction. Hence $\ell_\infty = 0$ which completes the proof for $q = p$.

Finally, we derive from what precedes that $\lim_N \min_{1 \leq i \leq N} |X - a_i| = \inf_{N \geq 1} |X - a_N| = 0$ $\mathbb{P}$-*a.s.*. As $X \in L^q(\mathbb{P})$, $\min_{1 \leq i \leq N} |X - a_i| \leq |X - a_1| \in L^q$, the conclusion follows from the Lebesgue dominated convergence theorem. $\square$

Remark on uniqueness. Uniqueness of L^p-optimal greedy quantization sequence turns out to be quite different problem from its counterpart for regular L^p-optimal quantization. Thus, for 1-dimensional log-concave distributions, it is well-known that uniqueness of L^p-optimal quantizers holds true (up to a reordering of the components in an increasing order) holds true. For L^p-optimal greedy quantization, this uniqueness may fail. Basically, greedy quantization is more influenced by the symmetry properties of the distributions: thus for the $\mathcal{N}(0;1)$-distribution (whose density is log-concave), its is clear that $a_1 = 0$ (unique L^p-median) but then we have that if a_2 is the (unique, see Proposition A.1 in Appendix A) solution to the the problem

$$\min_{a \geq 0} \mathbb{E}\big(|X|^p \wedge |X - a|^p\big) \quad \text{where } X \text{ has distribution } \mu = \mathcal{N}(0;1)$$

then both a_2 and $-a_2$ are solutions to the greedy problem (1.1) at level $N = 2$ by symmetry of (the distribution of) X. In fact, one derives in turn that $(0, a_2, -a_2)$ and $(0, -a_2, a_2)$ are both the first three terms of $(L^p, \mathcal{N}(0;1)$-optimal greedy quantization sequences.

2.1 About L^p-optimal greedy quantization in an Euclidean framework

In this section we assume that $|\,.\,|$ denotes an Euclidean norm on $\mathbb{R}^d$. Let $\overline{\mathcal{H}}_\mu$ be the *closed* convex hull of the support of the distribution μ.

Proposition 2.2. *Let $(\,.|.\,)$ denote the inner product induced by the Euclidean norm. If* $\operatorname{supp}(\mu)$ *contains at least N elements then, the first N elements of any optimal greedy quantization sequence takes values in $\overline{\mathcal{H}}_\mu$. If* $\operatorname{supp}(\mu)$ *is infinite any optimal greedy quantization sequence takes values in $\overline{\mathcal{H}}_\mu$.*

Proof. We proceed by induction. Let $a_1 \in \operatorname{argmin}_{a\in\mathbb{R}^d}\mathbb{E}|X-a|^p$ and let $\Pi_1(a_1)$ it projection on $\overline{\mathcal{H}}_\mu$. If $a_1 \neq \pi_1(a_1)$, the pseudo-Pythagorus Theorem implies

$$\forall\,\xi \in \overline{\mathcal{H}}_\mu, \quad |\xi - a_1|^2 \geq |\xi - \pi_1(a_1)|^2 + |a_1 - \pi_1(a_1)|^2$$

so that $\mathbb{E}\,|X-a|^2 \geq \mathbb{E}|X-\pi_1(a_1)|^2 + |a_1-\pi_1(a_1)|^2$ which yields a contradiction to the definition of a_1. Hence $a_1 \in \overline{\mathcal{H}}_\mu$.

Let $a_N \in \operatorname{argmin}_{a\in\mathbb{R}^d} e_p(a^{(N-1)}\cup\{a\}, X)$. It follows from Proposition 2.1(b) that $a_N \in \operatorname{argmin}_{a\in\mathbb{R}^d}\mathbb{E}\big(|X-a|^p\,|\,X\in W_{a_N}\big)$ where $W_{a_N} = \{\xi\in\mathbb{R}^d\,|\,|\xi-a_N| \leq d(\xi, a^{(N-1)})\}$ is a closed (polyhedral) convex set since the norm is Euclidean and has a positive μ-measure. As a consequence $a_N \in \overline{\mathcal{H}}_{\mu(.\,|W_{a_N})}$ where $\mu(.\,|W_{a_N})$ is the conditional distribution of μ given W_{a_N}. One concludes by noting that $\overline{\mathcal{H}}_{\mu(.\,|W_{a_N})} = \overline{\mathcal{H}}_\mu \cap W_{a_N} \subset \overline{\mathcal{H}}_\mu$. $\square$

Remark. Let $p=2$. We know that $\mu(W_{a_N}) > 0$ as soon as $\operatorname{card}(\operatorname{supp}(\mu) \geq N$ from Proposition 2.1(b) and that

$$\operatorname{argmin}_{a\in\mathbb{R}^d}\int_{W_{a_N}} |\xi-a|^2\mu(d\xi) = \left\{\frac{\int_{W_{a_N}}\xi\,\mu(d\xi)}{\mu(W_{a_N})}\right\}$$

i.e.

$$a_N = \frac{\int_{W_{a_N}}\xi\,\mu(d\xi)}{\mu(W_{a_N})} = \mathbb{E}\big(X\,|\,X\in W_{a_N}\big). \tag{2.2}$$

This can be seen as a fixed point formula and is the starting point of stochastic optimization procedure to compute by simulation (of i.i.d. samples of X) of optimal greedy sequences (see Section 6).

3 Greedy quantization is rate optimal

3.1 A general rate optimality result

Following [21], we define for every $b\in(0,+\infty)$ the *b-maximal function* associated to an L^p-optimal greedy quantization sequence $(a_N)_{N\geq1}$ by

$$\forall\,\xi\in\mathbb{R}^d, \quad \Psi_b(\xi) = \sup_{N\geq1}\frac{\lambda_d\big(B(\xi, bd(\xi,a^{(N)}))\big)}{\mu\big(B(\xi, bd(\xi,a^{(N)}))\big)} \in [0,+\infty].$$

It is clear that $\Psi_b(\xi) > 0$ for every $\xi \neq a_1$ (L^p-median).

Note that this notion of b-maximal function (originally introduced in [21]) can be naturally defined with respect to a sequence of grids $(\Gamma_N)_{N\geq1}$ where Γ_N has size N.

The theorem below yields a criterion based on the integrability of the maximal function Ψ_b which implies that an (L^p,μ)-optimal greedy quantization sequence is (L^p,μ)-rate optimal (in the sense of Zador's Theorem). More practical criterions are given further on in Section 4.

Theorem 3.1. *Let* $p\in(0,+\infty)$ *and let* $\mu=\mathbb{P}_X$ *be such that* $\int_{\mathbb{R}^d}|\xi|^p\mu(d\xi)<+\infty$. *Let* $(a_N)_{N\geq 1}$ *be an* L^p*-optimal greedy quantization sequence. Assume that there exists* $b\in(0,\frac{1}{2})$ *such that* $\Psi_b\in L^{\frac{p}{p+d}}(\mu)$. *Then*

$$\limsup_N N^{\frac{1}{d}}e_p(a^{(N)},X)<+\infty. \tag{3.3}$$

Proof. First, note that if μ is a Dirac mass δ_a for some $a\in\mathbb{R}^d$, then $a_1=a$ and $e_p(a^{(N)},X)=0$ for every integer $N\geq 1$. Otherwise, we rely on the following micro-macro inequality established in [21] (see Equation (3.4) in the proof of Theorem 2, with the standard convention $\frac{1}{0}=+\infty$).

$$\forall\,\xi\in\mathbb{R}^d,\qquad d(\xi,a^{(N)})^p\leq\frac{C_{p,b}}{\mu\big(B(\xi,bd(\xi,a^{(N)}))\big)}\Big(e_p(a^{(N)},X)^p-e_p(a^{(N)}\cup\{\xi\},X)^p\Big)$$

where $b\in(0,\frac{1}{2})$ and $C_{p,b}$ is a positive real constant depending on p and b. Then, it follows that

$$e_p(a^{(N)}\cup\{\xi\},X)^p\leq e_p(a^{(N)},X)^p-\frac{1}{C_{p,b}}\frac{\mu\big(B(\xi,bd(\xi,a^{(N)}))\big)}{\lambda_d\big(B(\xi,bd(\xi,a^{(N)}))\big)}b^d d(\xi,a^{(N)})^{p+d}V_d \tag{3.4}$$

where V_d denotes the hyper-volume of the unit ball with respect to the current norm on $\mathbb{R}^d$ *i.e.* $V_d=\lambda_d\big(B_{|\,.\,|}(0;1)\big)$. This implies that

$$e_p(a^{(N)}\cup\{\xi\},X)^p\leq e_p(a^{(N)},X)^p-\frac{1}{\widetilde{C}_{p,b,d}}\frac{1}{\Psi_b(\xi)}d(\xi,a^{(N)})^{p+d} \tag{3.5}$$

where $\widetilde{C}_{p,b,d}=C_{p,b}/(b^dV_d)\in(0,+\infty)$. Note that $\mu(\{a_1\})<1$ since μ is not a Dirac mass, so that

$$\int_{\mathbb{R}^d}\Psi_b^{\frac{p}{p+d}}(\xi)\,\mu(d\xi)>0.$$

Consequently, as $\Psi_b\in L^{\frac{p}{p+d}}(\mu)$, we can define the probability distribution $\nu=\kappa_{b,p,d}\,\Psi_b^{\frac{p}{p+d}}.\mu$ (where $\kappa_{b,p,d}=\Big(\int\Psi_b^{\frac{p}{p+d}}\Big)^{-1}\in(0,+\infty)$ is a normalizing real constant). Then, integrating the above inequality with respect to ν yields

$$\int_{\mathbb{R}^d}e_p\big(a^{(N)}\cup\{\xi\},X\big)^p\nu(d\xi)\leq e_p(a^{(N)},X)^p-\widetilde{C}_{p,b,d}\int_{\mathbb{R}^d}d(\xi,a^{(N)})^{p+d}\frac{\nu(d\xi)}{\Psi_b(\xi)}.$$

Jensen's Inequality applied to the convex function $u\mapsto u^{1+\frac{d}{p}}$ yields

$$\begin{aligned}\int_{\mathbb{R}^d}d(\xi,a^{(N)})^{p+d}\frac{\nu(d\xi)}{\Psi_b(\xi)}&\geq\left(\int_{\mathbb{R}^d}d(\xi,a^{(N)})^p\frac{\nu(d\xi)}{\Psi_b(\xi)^{\frac{p}{p+d}}}\right)^{1+\frac{d}{p}}\\&=\kappa_{b,p,d}^{1+\frac{d}{p}}\left(\int_{\mathbb{R}^d}d(\xi,a^{(N)})^p\mu(d\xi)\right)^{1+\frac{d}{p}}\\&=\kappa_{b,p,d}^{1+\frac{d}{p}}\,e_p\big(a^{(N)},X\big)^{p+d}.\end{aligned}$$

On the other hand, it is clear that

$$e_p\big(a^{(N+1)},X\big)^p\leq\int_{\mathbb{R}^d}\nu(d\xi)e_p\big(a^{(N)}\cup\{\xi\},X)\big)^p$$

so that, finally, if we set $A_{N} = e_p(a^{(N)}, X)^p$, $N \geq 1$, this sequence satisfies for every integer $N \geq 1$, the recursive inequality

$$A_{N+1} \leq A_{N} - \widetilde{\kappa}\, A_{N}^{1+\frac{d}{p}}$$

where $\widetilde{\kappa} = \kappa_{b,p,d}^{1+\frac{d}{p}} \widetilde{C}_{r,b,d}$. The sequence $(A_{N})_{N\geq 1}$ being non-negative, one classically derives the announced conclusion (for a proof, see Lemma B.1 in Appendix B, applied with $\rho = \frac{d}{p}$ and $C = \widetilde{\kappa}$). $\square$

Remark. • A careful reading of the proof shows that, if we define the sequence of functions $\Psi_{b,N}$ by

$$\forall\, \xi \in \mathbb{R}^d, \quad \Psi_{b,N}(\xi) = \frac{\lambda_d\big(B(\xi, bd(\xi, a^{(N)}))\big)}{\mu\big(B(\xi, bd(\xi, a^{(N)}))\big)} \in [0, +\infty],$$

then the theorem holds true under the weaker assumption that there exists an integer $N_0 \geq 1$ such that $\sup_{N \geq N_0} \int_{\mathbb{R}^d} \Psi_{b,N}(\xi)\mu(d\xi) < +\infty$.Unfortunately, this fact seems to be of little practical interest.

• When μ is singular with respect to the Lebesgue measure (no absolutely continuous part), it is likely that, like for standard optimal vector quantization in Zador's Theorem, this rate is not optimal. The natural conjecture should be that greedy quantization sequence(s) go to 0 at the same rate as that obtained for sequences of optimal quantizers which is not $N^{-\frac{1}{d}}$ when the distribution μ is singular (see *e.g.* [19]).

• Since we know that $d(\xi, a^{(N)}) \downarrow 0$ as $N \to +\infty$, $\mu(d\xi)$-$a.s.$, it is clear that if $\mu = \varphi.\lambda_d$ (or even $\mu = \varphi.\lambda_d \overset{\perp}{+} \tilde{\mu}$, to be checked), then by the Lebesgue differentiation theorem

$$\frac{1}{\varphi(\xi)} = \liminf_{N} \frac{\lambda_d\big(B(\xi, bd(\xi, a^{(N)}))\big)}{\mu\big(B(\xi, bd(\xi, a^{(N)}))\big)} \leq \Psi_b(\xi) \quad \mu(d\xi)\text{-}a.s.$$

so that by Fatou's Lemma, the condition $\Psi_b \in L^{\frac{p}{p+d}}(\mu)$ implies

$$\int_{\mathbb{R}^d} \varphi^{\frac{d}{p+d}}(\xi) d\lambda_d(d\xi) < +\infty.$$

So, we retrieve here the statement of Remark 6.3(c), p.79, in [19] which points out that *if* optimal L^p-mean quantization goes to zero at rate $N^{-\frac{1}{d}}$ then the above integral is finite (see also Section 1 in [21]). Of course, as emphasized in Remark 6.3(a) from [19], p.79, the classical condition under which Zador's Theorem holds, namely $\mathbb{E}|X|^{p+\delta} = \int_{\mathbb{R}^d} |\xi|^{p+\delta}\mu(d\xi) < +\infty$ for a $\delta > 0$, implies the finiteness of this integral owing to an appropriate application of Hölder's inequality. The above result suggests a hopefully nonempty question: since L^p-rate optimality for greedy sequence (and consequently for true L^p-optimal quantizers) holds as soon as $X \in L^p(\mathbb{P})$ and $\psi_b(X) \in L^{\frac{p}{p+d}}(\mathbb{P})$ for a $b \in (0, \frac{1}{2})$, are such conditions achievable when $\mathbb{E}|X|^{p+\delta} = +\infty$ for every $\delta > 0$.

3.2 Distortion mismatch for optimal greedy quantization sequences

In this section we address the problem of distortion mismatch originally investigated in [21] for sequences of optimal N-quantizers.

If $q \in (0, p]$ and $X \in L^p(\mathbb{P})$ any optimal greedy sequence $(a_{N})_{N\geq 1}$ remains L^q-rate optimal for the L^q-norm owing to the monotony of the L^q-norm as function of q. But the challenging question for distortion mismatch starts with the case $q > p$. It is solved in the proposition below, still relying on an integrability assumption on the b-maximal function(s) Ψ_b. For more practical criterions we again refer to Section 4.

Proposition 3.1. *Let $q \in (p, +\infty)$ and let $X \in L^p(\mathbb{P})$ with distribution $\mu = \mathbb{P}_X$. Assume that the maximal function $\Psi_b \in L^{\frac{q}{p+d}}(\mu)$ for some $b \in (0, \frac{1}{2})$. Let $(a_N)_{N\geq 1}$ be an L^p-optimal greedy sequence.*

Then $X \in L^q(\mathbb{P})$ and

$$\limsup_N N^{\frac{1}{d}} e_q(a^{(N)}, X) < +\infty.$$

Remarks. When $\operatorname{supp}(\mu)$ is not compact it is hopeless to have results for $q > p + d$ since it has been shown in [21] (Theorem 10 and Equation (2.7)) that the L^q-rate optimality of a sequence $(a_N)_{N\geq 1}$ would imply when $\mu = \varphi.\lambda_d$ that

$$\int_{\varphi>0} \varphi^{-\frac{q}{p+d}}(\xi)\mu(d\xi) = \int_{\varphi>0} \varphi^{1-\frac{q}{p+d}}(\xi)\lambda_d(d\xi) < +\infty.$$

However when μ has a compact support, we will see in Proposition 4.2(c) that L^q-rate optimality can be preserved under appropriate integrability assumptions.

Proof. First, note that if μ is a Dirac mass δ_a for some $a \in \mathbb{R}^d$, then $a_1 = a$ and $e_q(a^{(N)}, X) = 0$ for every integer $N \geq 1$. Otherwise, it follows from Equation (3.4) rewritten in a reverse way that

$$\forall\, \xi \in \mathbb{R}^d, \quad d(\xi, a^{(N)})^q \leq C_{b,d,p,q} \left(e_p(a^{(N)}, X)^p - e_p(a^{(N)} \cup \{\xi\}, X)^p\right)^{\frac{q}{p+d}} \Psi_b(\xi)^{\frac{q}{p+d}}(\xi).$$

Now, we note that

$$\forall\, \xi \in \mathbb{R}^d, \quad e_p(a^{(N)} \cup \{\xi\}, X)^p \geq e_p(a^{(N+1)}, X)^p$$

by definition of the sequence $(a_N)_{N\geq 1}$ so that

$$\forall\, \xi \in \mathbb{R}^d, \quad d(\xi, a^{(N)})^q \leq C_{b,d,p,q} \left(e_p(a^{(N)}, X)^p - e_p(a^{(N+1)}, X)^p\right)^{\frac{q}{p+d}} \Psi_b(\xi)^{\frac{q}{p+d}}(\xi).$$

Integrating with respect to μ yields

$$e_q(a^{(N)}, X)^q \leq C_{b,d,p,q} \left(e_p(a^{(N)}, X)^p - e_p(a^{(N+1)}, X)^p\right)^{\frac{q}{p+d}} \int_{\mathbb{R}^d} \Psi_b(\xi)^{\frac{q}{p+d}}(\xi)\mu(d\xi).$$

We know that $\int_{\mathbb{R}^d} \Psi_b(\xi)^{\frac{q}{p+d}}(\xi)\mu(d\xi) \in (0, +\infty)$ owing to the assumption made on μ and ψ_b. Hence

$$e_q(a^{(N)}, X)^q \leq \widetilde{C}_{b,d,p,q} \left(e_p(a^{(N)}, X)^p - e_p(a^{(N+1)}, X)^p\right)^{\frac{q}{p+d}}$$

where $\widetilde{C}_{b,d,p,q} = C_{b,d,p,q} \int_{\mathbb{R}^d} \Psi_b(\xi)^{\frac{q}{p+d}}(\xi)\mu(d\xi)$. Equivalently

$$e_q(a^{(N)}, X)^{p+d} \leq \widetilde{C}_{b,d,p,q}^{\frac{p+d}{q}} \left(e_p(a^{(N)}, X)^p - e_p(a^{(N+1)}, X)^p\right). \tag{3.6}$$

Summing over k between N and $2N-1$ yields

$$\begin{aligned}
\sum_{k=N}^{2N-1} e_q(a^{(k)}, X)^{p+d} &\leq \widetilde{C}_{b,d,p,q}^{\frac{p+d}{q}} \Big(e_p(a^{(N)}, X)^p - e_p(a^{(2N)}, X)^p\Big) \\
&\leq \widetilde{C}_{b,d,p,q}^{\frac{p+d}{q}} \Big(e_p(a^{(N)}, X)^p - e_{p,2N}(X)^p\Big)
\end{aligned}$$

where $e_{p,N}(X) = \inf_{\Gamma \subset \mathbb{R}^d, \operatorname{card}(\Gamma) \leq N} e_p(\Gamma, X)$ denotes the minimal L^p-mean quantization error.

It is clear that $\Psi_b \in L^{\frac{p}{p+d}}(\mu)$ since $p < q$ and $\Psi_b \in L^{\frac{q}{p+d}}(\mu)$. Consequently, combining the result established in Theorem 3.1 and Zador's Theorem on the sharp rate of decay of $e_{p,N}(X)$ in the scale $N^{-\frac{1}{d}}$ (see Theorem 6.2, p.78 in [19]), implies the existence of a positive real constant $\widetilde{C}'_{b,d,p,q} \in (0, +\infty)$ such that

$$\frac{1}{N}\sum_{k=N}^{2N-1} e_q(a^{(k)}, X)^{p+d} \le \widetilde{C}'_{b,d,p,q} N^{-(1+\frac{p}{d})}.$$

The convexity of $u \mapsto u^{p+d}$ then implies that

$$\left(\frac{1}{N}\sum_{k=N}^{2N-1} e_q(a^{(k)}, X)\right)^{p+d} \le \widetilde{C}'_{b,d,p,q} N^{-\frac{p+d}{d}}.$$

On the other hand the sequence $\big(e_q(a^{(N)}, X)\big)_{N\ge 1}$ is clearly non-decreasing since $(d(\xi, a^{(N)})^q)_{N\ge 1}$ is non-decreasing for every $\xi \in \mathbb{R}^d$. This implies that

$$e_q(a^{(2N)}, X) \le \frac{1}{N}\sum_{k=N}^{2N-1} e_q(a^{(k)}, X) \le \kappa N^{-\frac{1}{d}}.$$

Finally, still using that $N \mapsto e_q(a^{(N)}, X)$ is non-increasing, the announced result holds with $\kappa = \kappa_{b,d,p,q,X} = \max\big(\widetilde{C}^{\frac{1}{q}}_{b,d,p,q}, \|X - a_1\|_q\big)$. $\square$

4 Practical criterions for the integrability of the maximal function

These criterions are mainly borrowed form [21] where they have been established for the first time in order to solve the mismatch problem for optimal quantization

Compact case and $q < p + d$. The compact case relies on the following lemma which allows for non convex support for the distribution μ.

Lemma 4.1 (see Lemma 1 in [21]). *If $X \in L^p(\mathbb{P})$ has a distribution μ and $(\Gamma_N)_{N\ge 1}$ is a sequence of N-quantizers such that $\int_{\mathbb{R}^d} d(\xi, \Gamma_N)^p \mu(d\xi) \to 0$ then the maximal functions Ψ_b associated to $(\Gamma_N)_{N\ge 1}$ lies in $L^r_{loc}(\mu)$ for every $r \in (0, 1)$ i.e.*

$$\forall\, r \in (0,1),\ \forall\, b \in (0, +\infty),\ \forall R \in (0, +\infty),\quad \int_{\{|\xi|\le R\}} \psi_b(\xi)^r \mu(d\xi) < +\infty.$$

By combining this result (applied with $r = \frac{q}{p+d}$) with Proposition 2.1(b), we derive the following result which extends the one established in [10] for absolutely continuous distributions with convex support on $\mathbb{R}^d$. Note that the proof of the above lemma is not elementary, especially when $\mathrm{supp}(\mu)$ is not convex, and relies on the Besicovitch covering theorem.

Proposition 4.1 (Compact support). *If X has a distribution μ with compact support, then any L^p-optimal greedy quantization sequence $(a_N)_{N\ge 1}$ is L^q-rate optimal for every $q \in (0, p + d)$ i.e. satisfies*

$$\limsup_N N^{\frac{1}{d}} e_q(X, a^{(N)}) < +\infty.$$

Compact case and $q \geq p+d$. Results can be derived for $q > p+d$ when μ is absolutely continuous and has a compact support. They rely on the following Lemma (see Lemma 2 in [21]).

Lemma 4.2. *Assume $\mu = \varphi.\lambda_d$, $\mathbb{E}|X|^p < +\infty$, $\mathrm{supp}(\mu)$ is the finite union of closed convex sets and $\lambda_{d|\mathrm{supp}(\mu)}$ is absolutely continuous with respect to μ.*

Let $(\Gamma_N)_{N\geq 1}$ be a sequence of quantization grids satisfying $e_p(\Gamma_N, X) \to 0$ as $N \to +\infty$. Then, for every $q \in (1, +\infty]$, the associated maximal functions Ψ_b lies in $L^q_{loc}(\mu)$ iff $\frac{1}{\varphi} \in L^q_{loc}(\mu)$.

As a consequence of this lemma, we derive the following proposition which deals with the cases $q > p+d$ (in (a)) and $q = p+d$ (in (b)).

Proposition 4.2. (a) *Let $\mu = \varphi.\lambda_d$ be like in the preceding lemma and let $(a_N)_{N\geq 1}$ is an L^p-optimal greedy quantization sequence for μ. Let $q > d+p$. If*

$$\int_{\mathbb{R}^d} \varphi^{-\frac{q}{d+p}}(\xi)\mu(d\xi) = \int_{\{\varphi>0\}} \varphi^{1-\frac{q}{d+p}}(\xi)\lambda_d(d\xi) < +\infty$$

then $(a_N)_{N\geq 1}$ is $L^{q'}$-rate optimal for every $q' \in (0, q]$ i.e.

$$\limsup_N N^{\frac{1}{d}} e_{q'}(X, a^{(N)}) < +\infty.$$

In particular, if $\varphi \geq \varepsilon > 0$ on $\mathrm{supp}(\mu)$, then the above integral criterion is fulfilled.
(b) *Let $q = p+d$. If there exists $\delta > 0$ such that*

$$\int_{\mathbb{R}^d} \varphi^{-(1+\delta)}(\xi)\mu(\xi) = \int_{\{\varphi>0\}} \varphi^{-\delta}(\xi)\lambda_d(d\xi) < +\infty$$

then
$$\limsup_N N^{\frac{1}{d}} e_{p+d}(X, a^{(N)}) < +\infty.$$

Non-compact radial case.

Lemma 4.3 (see Corollary 3 in [21])**.** *If $X \in L^{p+\delta}(\mathbb{P})$ for some $\delta > 0$ with an essentially radial distribution $\mu(d\xi) = \varphi(\xi)\lambda_d(d\xi)$ in the sense that*

$$\varphi = h(|\,.\,|_0) \;\text{ on }\; B_{|\,.\,|_0}(0,R)^c \text{ with } h : (R, +\infty) \to \mathbb{R}_+, \text{ non-increasing and } |\,.\,|_0 \text{ any norm on } \mathbb{R}^d. \tag{4.7}$$

Let $(\Gamma_N)_{N\geq 1}$ be a sequence of N-quantizers such that $e_q(\Gamma_N, X) \to 0$. If there exists a real constant $c > 1$ such that

$$\int_{\mathbb{R}^d} \varphi(c\,\xi)^{-\frac{q}{p+d}}\mu(d\xi) = \int_{\mathbb{R}^d} \varphi(c\,\xi)^{-\frac{q}{p+d}}\varphi(\xi)d\xi < +\infty \tag{4.8}$$

then $\Psi_b \in L^{\frac{q}{p+d}}(\mu)$.

In fact, as stated in [21], Corollary 3 is written to be used only with L^p-optimal quantizers so the above formulation includes minor modifications. Combining this lemma with Proposition 2.1(b) and Theorem 3.1 yields the following proposition.

Proposition 4.3 (Non-compact support with radial density)**.** *If $X \in L^{p+\delta}(\mathbb{P})$ for some $\delta > 0$ with an essentially radial distribution in the sense of (4.7). If, furthermore, φ satisfies (4.8), then any L^p-optimal greedy sequence $(a_N)_{N\geq 1}$ is L^q-rate optimal i.e. satisfies*

$$\limsup_N N^{\frac{1}{d}} e_q(X, a^{(N)}) < +\infty.$$

This case includes *e.g.* all the centered *hyper-exponential distributions* of the form $\mu = \varphi.\lambda_d$ with

$$\varphi(\xi) = \kappa_{a,b,c}|\xi|_0^c\, e^{-a|\xi|_0^b},\ \xi\in\mathbb{R}^d,\ a,b>0,\ c>-d$$

and $|\,.\,|_0$ is any norm on $\mathbb{R}^d$ and subsequently all hyper-exponential distributions since L^p-mean-quantization errors is invariant by translation of the random vector X. In particular, this includes all normal and Laplace distributions.

Remark. In one dimension, (4.7) can be replaced *mutatis mutandis* by a one-sided variant: if there exists $R_0, R'_0\in\mathbb{R}$, $R'_0\geq R_0$ such that

$$\operatorname{supp}(\mu)\subset[R_0,+\infty)\ \text{ and }\ f_{|[R'_0,+\infty)}\text{ is non-increasing.} \tag{4.9}$$

This criterion is satisfied by the gamma distributions on $\mathbb{R}_+$ (including the exponential distributions).

Non-compact and possibly non-radial case.

Corollary 4.1. *Assume $\mu = \varphi.\lambda_d$ and $\mathbb{E}\,|X|^{p+\delta}<+\infty$ for some $\delta>0$. Furthermore, assume that* $\operatorname{supp}(\mu)$ *is* peakless *in the following sense*

$$\kappa_\varphi := \inf_{\xi\in\operatorname{supp}(\mu),\,0<\rho\leq 1}\frac{\lambda_d\big(\operatorname{supp}(\mu)\cap B(\xi,\rho)\big)}{\lambda_d\big(B(\xi,\rho)\big)}>0 \tag{4.10}$$

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

$$\forall\,\xi,\,\xi'\in\operatorname{supp}(\mu),\ \|\xi\|\geq M,\ \|\xi'-\xi\|\leq 2\eta\,\|\xi\|\implies \varphi(\xi')\geq C\varphi(\xi)^{1+\varepsilon}. \tag{4.11}$$

Then, for every $q\in(0,\frac{p+d}{1+\varepsilon})$ such that

$$\int_{\mathbb{R}^d}\varphi(\xi)^{-\frac{q(1+\varepsilon)}{p+d}}\mu(d\xi) = \int_{\{\varphi(\xi)>0\}}\varphi(\xi)^{1-\frac{q(1+\varepsilon)}{p+d}}\lambda_d(d\xi)<+\infty$$

(if any), any greedy L^p-optimal sequence $(a_N)_{N\geq1}$ is L^q-rate optimal i.e. satisfies

$$\limsup_N N^{\frac{1}{d}}e_q(X,a^{(N)})<+\infty.$$

In particular, if (4.11) holds either for $\varepsilon=0$ or for every $\varepsilon\in(0,\underline{\varepsilon}]$ ($\underline{\varepsilon}>0$), and if

$$\forall\,q\in(0,p+d),\qquad \int_{\mathbb{R}^d}\varphi(\xi)^{-\frac{q}{p+d}}\mu(d\xi) = \int_{\{\varphi(\xi)>0\}}\varphi(\xi)^{1-\frac{q}{d+p}}\lambda_d(d\xi)<+\infty \tag{4.12}$$

then the above conclusion holds for every $q\in(p,p+d)$.

Note that (if $\lambda_d(\operatorname{supp}(\mu))=+\infty$) Assumption (4.10) 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

$$\kappa'_\varphi := \inf_{\xi\in\operatorname{supp}(\mu)}\left\{\frac{\lambda_d(\operatorname{supp}(\mu)\cap B(\xi,\rho))}{\lambda_d(B(\xi,\rho))},\ 0<\rho\leq c\,\|x\|\right\}>0.$$

5 Further answers and questions about greedy quantization

In this section, we temporarily denote by $\big(a_{N,p}\big)_{N\geq 1}$ the L^p-optimal greedy quantization sequence for the uniform distribution $U([0,1])$ and by $\big(\alpha^{(N),p}\big)_{N\geq 1}$ the resulting sequence of greedy quantizers.

$\triangleright$ **Rate optimality of greedy sequences** It is a straightforward consequence of Zador's Theorem that if the distribution $\mu = \mathbb{P}_X$ of $X\in L^{p+\delta}$, $\delta>0$, has a non-zero absolutely continuous component (*i.e.* $\varphi = \frac{d\mu}{d\lambda_d} \not\equiv 0$) and satisfies the assumptions of Theorem 3.1, then

$$e_p(a^{(N),p}, X) \asymp N^{-\frac{1}{d}}$$

since $e_p(a^{(N),p}, X) \geq e_{p,N}(X)$ and $\liminf_N N^{\frac{1}{d}} e_{p,N}(X) \geq \widetilde{J}_{p,d}\|\varphi\|^{\frac{1}{p}}_{L^{\frac{p}{p+d}}(\lambda_d)} > 0$. (By the way it proves that under the assumption of Theorem 3.1, $\|\varphi\|_{L^{\frac{p}{p+d}}(\lambda_d)} < +\infty$.)

By a similar argument, the same holds true for the distortion mismatch problem under the assumptions of Proposition 3.1.

$\triangleright$ **Can greedy quantization sequence produce asymptotically optimal quantizers?** If μ has an absolutely continuous component with density φ, then any sequence $(\Gamma_n)_{n\geq 1}$ of *asymptotically* (L^p,μ)*-optimal quantization grids* at level $N_n = \mathrm{card}(\Gamma_n) \to +\infty$, satisfies the *empirical measure theorem* (see [19], Theorem 7.5, p.96 and [14] for a slight refinement), namely

$$\frac{1}{N_n}\sum_{a\in\Gamma_n}\delta_a \xrightarrow{(w)} \mu^{(p)} = \frac{\varphi^{\frac{d}{d+p}}}{\int_{\mathbb{R}^d}\varphi^{\frac{d}{d+p}}d\lambda_d}.\lambda_d \quad \text{as} \quad n\to+\infty$$

where $\xrightarrow{(w)}$ denotes the weak convergence of probability measures. Note that when $\mu = U([0,1])$, $\mu^{(p)} = U([0,1])$, for every $p\in(0,+\infty)$.

By *asymptotically* (L^p,μ)*-optimal*, we mean that the (L^p,μ)-mean quantization errors induced by the grids Γ_n satisfy the sharp asymptotics of Zador's Theorem, namely $\lim_n N_n^{\frac{1}{d}} e_p(\Gamma_n, X) = \widetilde{J}_{p,d}\|\varphi\|^{\frac{1}{p}}_{L^{\frac{p}{p+d}}(\lambda_d)}$.

It is pointed out in [10] (Theorem 4.10 and Corollary 4.11) that the quantizers $(a^{(N),p})_{N\geq 1}$ designed from an (L^p,μ)-optimal greedy quantization sequence $(a_{N,p})_{N\geq 1}$ are usually not asymptotically (L^p,μ)-optimal, even up to an extraction. The counter-example is exhibited in the 1-dimensional basic setting of the uniform distribution $U([0,1])$.

The authors first build and analyze an $(L^1, U([0,1]))$-optimal greedy sequence $(a_{N,1})_{N\geq 1}$. Then, they shown that the (tight) sequence of empirical measures $\widetilde{\mu}_N = \frac{1}{N}\sum_{1\leq k\leq N}\delta_{a_{k,1}}$ on $[0,1]$ does not have the uniform distribution $U([0,1])$ (or equivalently the Lebesgue measure $\lambda_{1|[0,1]}$ over $[0,1]$) as a weak limiting distribution. In particular, this implies, owing to the above empirical measure theorem, that

$$\liminf_N e_1\big(a^{(N),1}, U([0,1])\big) > \frac{1}{4} = \widetilde{J}_{1,1}$$

keeping in mind that $\widetilde{J}_{1,1} = \lim_N e_{1,N}\big(U([0,1])\big) = \inf_N e_{1,N}\big(U([0,1])\big)$. Otherwise, by the above empirical mean theorem, there would exists a subsequence $N'\to+\infty$ such that $\widetilde{\mu}_{N'} \overset{(w)}{\to} \mu^{(1)} = \lambda_{1|[0,1]} = U([0,1])$. Equivalently, this reads

$$\liminf_N \frac{e_1\big(a^{(N),1}, U([0,1])\big)}{e_{1,N}\big(U([0,1])\big)} > 1.$$

Numerical tests graphically reproduced in [10] (Figure 1, p.521) suggest that

$$\liminf_N N e_1\big(a^{(N),1}, U([0,1])\big) \approx 1,02 \times \widetilde{J}_{1,1}.$$

Our own numerical tests, based on the algorithms developed in Section 6 in the quadratic case ($p = 2$), implemented with the uniform distribution, the scalar $\mathcal{N}(0,1)$ and bi-variate $\mathcal{N}(0;I_2)$ normal distributions provide similar conclusions (see Section 6 devoted to algorithmic aspects and numerical experiments).

This leads to our first open question: is this a generic situation? Or, to be more precise:

Open question 1: May an optimal (L^p,μ)-greedy sequence $(a_{N,p})_{N\ge 1}$ contain subsequence(s) $\big(a^{(N'),p}\big)_{N\ge 1}$ of asymptotically (L^p,μ)-optimal μ-quantizers?

In fact, we conjecture that the a generic answer is negative. This amounts to proving, still owing to the empirical measure theorem, that for any optimal (L^p,μ)-greedy sequence $(a_{N,p})_{N\ge 1}$

$$\liminf_N \frac{e_p\big(a^{(N),p},\mu\big)}{e_{p,N}\big(\mu\big)} > 1.$$

$\triangleright$ **Are (L^p,μ)-optimal greedy quantization sequence really optimal among (μ-rate optimal) sequences?** Let us have a look at the celebrated dyadic Van der Corput (VdC) sequence, viewed as a quantization sequence. Let us recall that the dyadic VdC sequence is defined by

$$\xi_N = \sum_{k=0}^{r} \frac{n_k}{2^{k+1}} \quad \text{where} \quad N = n_r 2^r + \cdots + n_0,\ n_i \in \{0,1\},\ i = 1,\ldots,r.$$

$\triangleright$ *The L^1-mean quantization problem for the* VdC *sequence.* Elementary computations, not reproduced here, show that

$$\liminf_N N e_1\big(\xi_1,\ldots,\xi_N,[0,1]\big) = \frac{1}{4} = \widetilde{J}_{1,1}$$

and that

$$\limsup_N N e_1\big(\xi_1,\ldots,\xi_N,[0,1]\big) = \frac{9}{32} = \frac{9}{8}\,\widetilde{J}_{1,1}.$$

This $\liminf$ is achieved by the subsequence $N_n = 2^{n-1}$, $n \ge 1$, and the $\limsup$ with subsequence $N_n = \frac{3}{2}.2^n = 3.2^{n-1}$, $n \ge 1$. So we can claim that:

- there exist rate optimal sequences in the sense of (3.3) which are not solutions to the greedy problem (1.1);
- there exist rate optimal sequences $(\xi_N)_{N\ge 1}$ containing subsequence of quantizers $(\xi^{(N')})_{N\ge 1}$ which are asymptotically L^1-rate optimal quantizers: so is the case of the VdC sequence with the above subsequence $N' = 2^{n-1}$.

Figure 1 in [10] also suggests that the L^1-optimal greedy quantization sequence $\big(a_{N,1}\big)_{N\ge 1}$ for the uniform distribution $U([0,1])$ satisfies

$$\limsup_N N e_1\big(a^{(N),1}, U([0,1])\big) \approx 1.09 \times \widetilde{J}_{1,1} \ \text{ and } \ 1.09 < 1.125 = 9/8.$$

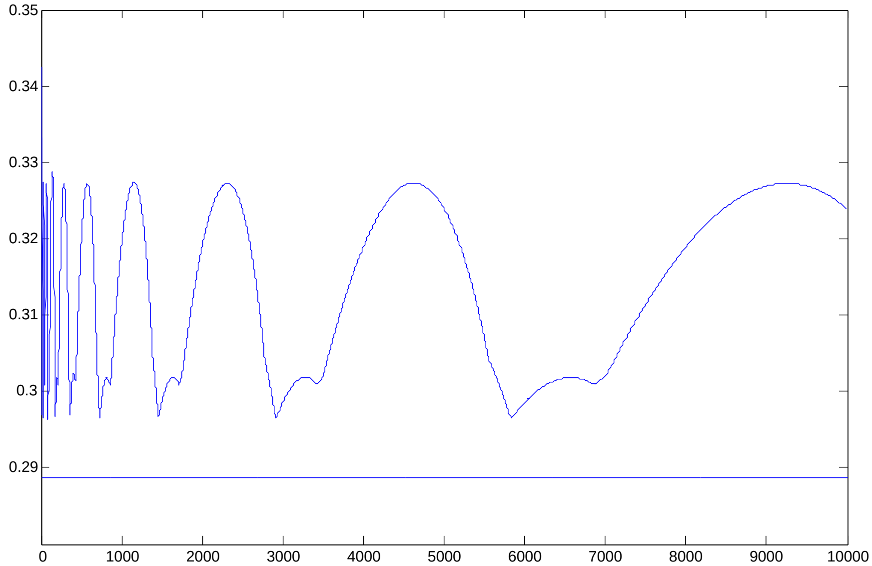

Figure 1: *Graph* $N \mapsto Ne_2\big(a^{(N)}, U([0,1])\big)$, $N = 1, \ldots, 10\,000$.

$\triangleright$ *The L^2-mean quantization problem for the* VdC *sequence.* The same phenomenons are confirmed in the quadratic case since, *mutatis mutandis*,

$$\liminf_N Ne_2\big(\xi_1, \ldots, \xi_N, [0,1]\big) = \frac{1}{2\sqrt{3}} = \widetilde{J}_{2,1} \;\text{ and }\; \limsup_N Ne_2\big(\xi_1, \ldots, \xi_N, [0,1]\big) = \frac{3\sqrt{5}}{4} \times \widetilde{J}_{2,1}$$

where we keep in mind that $\widetilde{J}_{2,1} = \lim_N Ne_{2,N}(U([0,1])) = \inf_N Ne_{2,N}(U([0,1]))$.

On the other hand, in a quadratic framework, using the greedy Lloyd I procedure described and analyzed in the next Section 6.1 (see Equations (6.14) if $d = 1$ and (6.17) if $d \geq 2$), we also observe numerically (see Figure 5) that

$$\liminf_N Ne_2(a^{(N),2}, U([0,1])) \approx 1.02732 \times \widetilde{J}_{2,1} > \widetilde{J}_{2,1}$$

and

$$\limsup_N Ne_2\big(a^{(N),2}, U([0,1])\big) \approx 1.13401 \times \widetilde{J}_{2,1}.$$

As for the lim inf, we verify again that no subsequence of $\big(a^{(N),2}\big)_{N\geq 1}$ can be asymptotically L^2-optimal and, ss for the lim sup, that the quadratic optimal greedy sequence $\big(a_{N,2}\big)_{N\geq 1}$ outperforms the dyadic *VdC* sequence from the lim sup criterion since $1.13401 < \frac{3\sqrt{5}}{4} = 1.67706$.

$\triangleright$ *Concatenated sequences.* From a more general point of view, there is a canonical method to produce for any distribution μ on $(\mathbb{R}^d, \mathcal{B}or(\mathbb{R}^d))$, a μ-rate optimal sequence for (L^p, μ)-quantization by *concatenating* (L^p, μ)-optimal grids of size 2^ℓ. We proceed as follows. Let $(b_N)_{N\geq 1}$ be a sequence made up with (L^p, μ)-optimal quantizers at level 2^ℓ, $\ell = 0, \ldots n-1$ *i.e.* so that

$$\Big\{b_{2^\ell}, \ldots, b_{2^{\ell+1}-1}\Big\} \text{ is an } (L^p, \mu)\text{-optimal quantizer at level } 2^\ell. \tag{5.13}$$

One checks straightforwardly by monotony of the L^p-mean quantization error that, for every $n \geq 1$,

$$e_{2^n-1}(b^{(2^n-1)}, \mu) \leq e_{2^{n-1}}\big(\{b_{2^{n-1}}, \ldots, b_{2^n-1}\}, \mu\big).$$

Hence, for every $N \geq 1$, let be $n = n(N)$ be such that $2^n - 1 \leq N \leq 2^{n+1}$. Then

$$e_p(b^{(N)}, \mu) \leq e_p(\{b_{2^{n-1}}, \dots, b_{2^n-1}\}) = e_{p,2^{n-1}}(\mu)$$

so that

$$\limsup_N N^{\frac{1}{d}} e_p(b^{(N)}, \mu) \leq \limsup_N \left(\frac{N}{2^{n(N)}}\right)^{\frac{1}{d}} \lim_N N^{\frac{1}{d}} e_{p,N}(\mu) = 2^{\frac{1}{d}} \lim_N N^{\frac{1}{d}} e_{p,N}(\mu).$$

$\triangleright$ *First elements of comparison.*

– If $\mu = U([0,1])$ and $p = 1$, one easily checks by induction that the dyadic VdC sequence can be obtained as a properly reordered sequence $(b_N)_{N\geq 1}$ from the L^p-optimal quantizers at level N given by $\left\{\frac{2k-1}{2N},\ 1 \leq k \leq N\right\}$ when $N = 2^n$, $n \geq 0$. In this very situation, the factor $2^{\frac{1}{d}} = 2$ is conservative since it can be replaced when $p = 1$ by $\frac{9}{8} = 1.125$ as seen above.

Anyway, the L^1-optimal greedy quantization sequence keeps the lead, since $\limsup_N \dfrac{e_1(a^{(N),1}, \mu)}{e_{1,N}(\mu)} \approx 1.09 < \dfrac{9}{8} \approx 1.125$.

– If $\mu = U([0,1])$ and $p = 2$, once again, the quadratic optimal greedy quantization sequence again keeps the lead, since

$$\limsup_N \frac{e_2(a^{(N),2}, \mu)}{e_{2,N}(\mu)} \approx 1.13401 < \frac{3\sqrt{5}}{4} \approx 1.67706 < 2.$$

– If $\mu = \mathcal{N}(0; I_2)$ (bivariate normal distribution *i.e.* $d = p = 2$), our own numerical experiments suggest for the third time (see more detailed numerical results in Section 6.2) that a quadratic optimal greedy quantization sequence (or, in practice, the suboptimal sequence resulting from the numerical implementation of the greedy Lloyd I algorithm) has a lower constant than $2^{\frac{1}{d}} \times \lim_N N^{\frac{1}{2}} e_{2,N}\mathcal{N}(0; I_2)$.

All these considerations experiments lead us to formulate a second open question:

Open question 2: Does an (L^p, μ)-optimal greedy quantization $(a_{N,p})_{N\geq 1}$ produce the lowest value for $\limsup_N N^{\frac{1}{d}}\, e_{p,N}\big(a^{(N),p}, \mu\big)$ among all sequences $(a_{N,p})_{N\geq 1}$?

A less ambitious question could be to compare (L^p, μ)-optimal greedy sequences to concatenated sequences (5.13) *i.e.*: "Is the (strict)inequality $\limsup_N \dfrac{e_p\big(a^{(N),p}, \mu\big)}{e_{p,N}(\mu)} < 2^{\frac{1}{d}}$ always satisfied?"

$\triangleright$ **Practical aspects in view of numerics.** From a more applied point of view, it would be of interest to establish for (L^p, μ)-optimal greedy sequences a counterpart of the non-asymptotic Zador Theorem in order to upper-bound the (L^p, μ)-mean quantization error of any greedy optimal sequence (normalized by $N^{-\frac{1}{d}}$) by the $L^{p+\delta}$-pseudo-standard deviation of the distribution μ and a universal constant depending only on d, p and δ. The proof of the non-asymptotic Zador's Theorem (a slight improvement of Pierce's Lemma established *e.g.* in [24]) relies on a random quantization argument involving the random quantizers $(Y^{(N)})_{N\geq 1}$ designed from an i.i.d. sequence $(Y_N)_{N\geq 1}$ with an appropriate distribution ν, such a result is not hopeless.

For numerical purposes, in particular numerical integration or conditional expectation approximation, some reasonably good estimates of $\limsup_N N^{\frac{1}{d}} e_{1,N}(a^{(N)}, \mu)$ in (3.3) would be very useful. This is to be compared to the never ending quest for sequences with low discrepancy with lower constant in the Quasi-Monte Carlo community.

6 Algorithmic aspects in the quadratic case

In this section we assume that $\mathbb{R}^d$ is equipped with the canonical Euclidean norm and that $p = 2$ (purely quadratic setting). So, will simply denote $(a_N)_{N\ge 1}$ quadratic optimal greedy sequences.

Practical computation of an optimal greedy sequence of quantizers relies on obvious variants algorithms ($CLVQ$ and Lloyd) implemented recursively: to switch from N to $N+1$, one first adds a $(N+1)^{th}$ point (sampled from the support of the distribution μ) to the N-tuple $(a_1, \dots a_N)$ computed during the first N^{th} stages of the optimization procedure. This makes the starting $(N+1)$-tuple for the modified $CLVQ$ to Lloyd procedure. Then, one launches one of these two optimization procedures with the following restriction: *all formerly computed components* a_i, $1 \le i \le N-1$ *are kept frozen*, and only the new point is moved following the standard rules. Thus, if implementing a $CLVQ$ like procedure, when the N^{th} component is the "winner" in the competition phase (*i.e.* the N^{th} component is the nearest neighbour to the new input stimulus). As for the (randomized) Lloyd I procedure, the Voronoi cell of the N^{th} component is the only one whose centroid (the N^{th} component) is updated, the other $N-1$ components remaining frozen as well. Let us be more precise.

6.1 The one-dimensional quadratic case

When $d = 1$ and the distribution μ is absolutely continuous with a continuous positive probability density φ on the real line, one can directly consider the counterpart of the historical deterministic Lloyd I procedure and of the gradient descent sometimes known as Forgy's algorithm or k-means. Let us be more specific.

$\triangleright$ **Greedy Lloyd's I procedure** • Assume $a_1, \dots a_{N-1}$ have been computed. Let $a_1^{(N-1)} < \cdots < a_{N-1}^{(N-1)}$ be an increasing reordering of $a_1, \dots, a_{N-1}$.

• Assume the N inter-point local inertia has also been computed, namely

$$\sigma_i^2 := \int_{a_i^{(N-1)}}^{a_{i+\frac12}^{(N-1)}} |a_i^{(N-1)} - \xi|^2 \mu(d\xi) + \int_{a_{i+\frac12}^{(N-1)}}^{a_{i+1}^{(N-1)}} |a_{i+1}^{(N-1)} - \xi|^2 \mu(d\xi),\ \ i = 1, \dots, N$$

where

$$a_0^{(N-1)} = a_{\frac12}^{(N-1)} = -\infty,\ a_{i-\frac12}^{(N-1)} = \frac{a_{i-1}^{(N-1)} + a_i^{(N-1)}}{2},\ \ i = 2, \dots, N-2,\ a_{N-\frac12}^{(N-1)} = a_N^{(N-1)} = +\infty.$$

• Choose an index $i_0 = i_0(N-1)$ such that $\sigma_{i_0}^2 = \max_{0\le i\le N} \sigma_i^2$ (maximal local inertia), then consider $a_0 = a_{N,0} \in (a_{i_0}^{(N-1)}, a_{i_0+1}^{(N-1)})$ and finally define recursively a sequence $a_{[n]} = a_{N,n}$, $n \ge 1$, by

$$a_{[n+1]} = \mathbb{E}\big(X \mid X \in W_{N,[n]}\big) = \frac{K_\mu\big(\frac{a_{i_0+1}^{(N-1)} + a_{[n]}}{2}\big) - K_\mu\big(\frac{a_{i_0}^{(N-1)} + a_{[n]}}{2}\big)}{F_\mu\big(\frac{a_{i_0+1}^{(N-1)} + a_{[n]}}{2}\big) - F_\mu\big(\frac{a_{i_0}^{(N-1)} + a_{[n]}}{2}\big)},\ \ n \ge 0, \tag{6.14}$$

where $F_\mu(x) = \mu((-\infty, x])$ is the cumulative distribution function of μ and K_μ its cumulative first moment function defined by

$$K_\mu(x) = \int_{(-\infty,x]} \xi\,\mu(d\xi),\ \ x \in \mathbb{R}.$$

It follows form an easy induction that, at every step $n \ge 0$ of the procedure, $a_{N,[n]} \in W_{N,[n]} \subset (a_{i_0}^{(N-1)}, a_{i_0+1}^{(N-1)})$ so that the procedure is well-defined.

Proposition 6.1. *If μ is* strongly unimodal *in the sense that $\mu = \varphi.\lambda_1$ with $\varphi : \mathbb{R} \to \mathbb{R}$* log-*concave, then $a_{N,[n]}$ converges toward the unique solution $a_{N,\infty} \in (a^{(N-1)}_{i_0}, a^{(N-1)}_{i_0+1})$ of the fixed point equation*

$$a_{_N} = \mathbb{E}\big(X \,|\, X \in W_N\big) \tag{6.15}$$

where $W_N \subset (a^{(N-1)}_{i_0}, a^{(N-1)}_{i_0+1})$ is the closed Voronoi cell of $a_{_N}$ in $a^{(N-1)} \cup \{a_{_N}\}$.

The detailed proof is postponed to the Appendix A.1. But we can already mention that it relies on classical arguments called upon in the proofs of the convergence of the standard Lloyd I procedure (and the uniqueness of the possible stationary limiting point, see [22, 7]).

Remarks. • The computation of the integrals involved in the algorithm can be performed by higher order quadrature formulas, or *e.g.* in the case where $\mu = \mathcal{N}(0;1)$ using the closed form for $\int_{-\infty}^{x} \xi e^{-\frac{\xi^2}{2}} \frac{d\xi}{\sqrt{2\pi}} = -\frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}$ and high accuracy approximations for its cumulative distribution function Φ_0, using *e.g.* continuous fractions expansions (see [1]).

• The log-concave assumption which implies the uniqueness of the fixed point for Equation (6.14), is satisfied by many usual families of distributions on the real line like *e.g.* the normal distributions $\mathcal{N}(m;\sigma^2)$, the exponential and Laplace distributions, the $\gamma(\alpha,\beta)$-distributions, $\alpha \geq 1$, $\beta > 0$, are strongly unimodal. On the other hand, the Pareto distributions are not strongly unimodal though uniqueness holds true (see [16]).

▷ **Greedy Forgy's algorithm (Newton zero search algorithm)** This procedure is defined recursively by

$$a_{[n+1]} = a_{[n]} - \Big(\gamma_{n+1} \wedge \frac{1}{\rho(a_{[n]})}\Big) \int_{\frac{a^{(N-1)}_{i_0}+a_{[n]}}{2}}^{\frac{a^{(N-1)}_{i_0+1}+a_{[n]}}{2}} \big(a_{[n]} - \xi\big)\mu(d\xi) \tag{6.16}$$

where $\gamma_{n+1} \in (0,1)$ goes to 0 as $n \to +\infty$, $\sum_n \gamma_n = +\infty$ and

$$\rho(a) = \mu\Big(\Big[\frac{a^{(N-1)}_{i_0}+a}{2}, \frac{a^{(N-1)}_{i_0+1}+a}{2}\Big]\Big) + \frac{a - a^{(N-1)}_{i_0}}{2} f\Big(\frac{a + a^{(N-1)}_{i_0}}{2}\Big) + \frac{a^{(N-1)}_{i_0+1} - a}{2} f\Big(\frac{a + a^{(N-1)}_{i_0+1}}{2}\Big) > 0$$

is the second derivative of the a function $a \mapsto \mathbb{E}\big(\min |X - a_i|^2 \wedge |X - a|^2\big)$.

Note that, owing to the thresholding of $1/\rho(a_{N,[n]})$ by $\gamma_{n+1} \in (0,1)$, this procedure lives in the interval $(a^{(N-1)}_{i_0}, a^{(N-1)}_{i_0+1})$ which makes it well-defined and consistent for every n.

When μ is not absolutely continuous, one can implement the same procedure by removing the term involving the second derivative with a step γ_n satisfying the standard *decreasing step assumption* ($\sum_n \gamma_n = +\infty$ and $\sum_n \gamma_n^2 < +\infty$), provided one can compute the μ-integrals of interest.

▷ **Numerical illustration with the $\mathcal{N}(0;1)$ distribution** To compute a quadratic optimal greedy sequence of the normal distribution $\mu = \mathcal{N}(0;1)$, we will take advantage of its symmetry. To this end we consider the distribution $\tilde{\mu} = \mu(\,.\,|\,\mathbb{R}_+)$ (μ conditioned to stay non-negative) which is clearly strongly unimodal and we compute by induction its quadratic optimal greedy sequence $(\widetilde{a}_{_N})_{N\geq 1}$ by the greedy Lloyd I procedure (6.14) with the convention that the origin 0 is a fixed but *active* point for this slight variant . By active we mean that 0 has its own Voronoi cell in $\mathbb{R}_+$ or, equivalently, that we implement our computation starting at $\widetilde{a}_0 = 0$ at $N = 0$.

As a second step, it is straightforward that the sequence defined by

$$a_0 = 0,\ a_{_{2N-1}} = \widetilde{a}_{_N},\ a_{_{2N}} = -\widetilde{a}_{_N},\ N \geq 1,$$

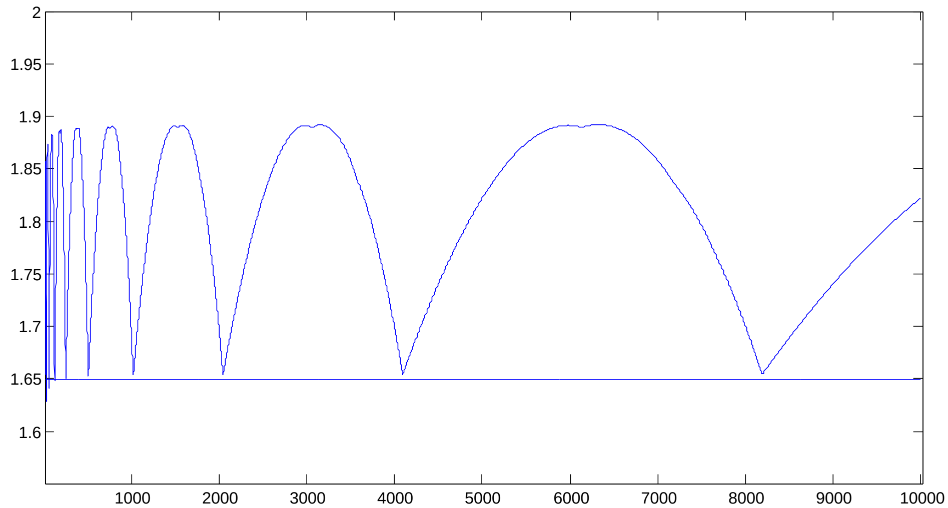

Figure 2: *Graph* $N \mapsto (2N-1)e_2\big(a^{(2N-1)}, \mathcal{N}(0;1)\big)$, $N = 4, \ldots, 10\,000$.

is a quadratic optimal greedy sequence.

We reproduce in Figure 2 the graph $N \mapsto (2N-1)e_2\big(a^{(2N-1)}, \mu\big)$, $N = 4, \ldots, 2^{10} = 10\,000$, where $\mu = \mathcal{N}(0;1)$.

Note that $\limsup_N Ne_2\big(a^{(N)}, \mu\big) = \limsup_N (2N-1)\, e_2\big(a^{(2N-1)}, \mu\big)$ since $e_2\big(a^{(N)}, \mu\big) \downarrow 0$ as $N \to +\infty$. As a consequence, we derive that

$$\liminf_N Ne_2\big(a^{(N)}, \mathcal{N}(0;1)\big) \approx 1.6534\cdots > \sqrt{\frac{3}{2}}\,\pi^{\frac{1}{4}} = \lim_N Ne_{2,N}\big(\mathcal{N}(0;1)\big)$$

since $\sqrt{\frac{3}{2}}\,\pi^{\frac{1}{4}} \approx 1.63055$. (The real constant in the right hand side of the inequality easily follows from Zador's Theorem). Note that, for the values $N = 2^n$, $0 \le n \le 7$, we observe that $Ne_2\big(a^{(N)}, \mathcal{N}(0;1)\big) < \sqrt{\frac{3}{2}}\,\pi^{\frac{1}{4}}$ ([2]).

As for the limsup, we observe numerically that

$$\limsup_N Ne_2\big(a^{(N)}, \mathcal{N}(0;1)\big) \approx 1.8921 < 2 \times \sqrt{\frac{3}{2}}\,\pi^{\frac{1}{4}} \approx 3.2611.$$

6.2 The multidimensional quadratic case (higher dimensions)

In higher dimensions, deterministic procedures like deterministic greedy Lloyd's I (fixed point procedure defined by (6.14)) or the greedy Forgy's (recursive zero search defined by (6.16)) algorithms become computationally too demanding due to the repeated computations of integrals on the Voronoi cells of the quantizers. So, it becomes necessary, at least when $d \geq 3$, to switch to stochastic optimization procedures like those described below, which are adaptations of the stochastic procedures introduced to compute true optimal N-quantizers. For more details about these original stochastic optimization procedures, mostly devised in the 1950's, we refer *e.g.* to [5, 30] for $CLVQ$ and [22, 15, 34]

[2]This is consistent in some way with the conjecture that $Ne_{2,N}\big(\mathcal{N}(0;1)\big)$ is increasing toward its limit.

for (randomized) Lloyd's I procedure or more applied textbooks like [17]. From a theoretical point of view, their common feature is that the convergence results ($a.s.$ or in L^p) remain partial, especially little is known when the distribution μ is not compactly supported. So we present below their greedy variants (without rigorous proof as concerns $CLVQ$). From a practical point of view, for both procedures, the computation of integrals on the Voronoi cells is replaced by repeated nearest neighbor searches among the components of the current N-quantizers which make them rather slow. But in our greedy framework, this drawback could be overcome by appropriate localization around the elementary quantizer of interest. But this is beyond the scope of the present work.

$\triangleright$ (Randomized) greedy Lloyd's I like procedure. The *greedy Lloyd I procedure* to compute a_N, assuming that $a^{(N-1)}$ is known, (starting from the mean $a_1 = \mathbb{E}\,X$) can be recursively defined in the quadratic case as follows:

$$a_{N,[n+1]} = \mathbb{E}\big(X \mid X \in W_{N,[n]}\big), \quad a_{N,[0]} \in \mathbb{R}^d \setminus \{a^{(N-1)}\}, \tag{6.17}$$

where $W_{N,[n]}$ is the closed Voronoi cell of $a_{N,[n]}$ with respect to the quantizer $a^{(N-1)} \cup \{a_{N,[n]}\}$. Of course in practice, we stop the Monte Carlo simulation at finite range M_n.

We establish in the proposition below, at least for absolutely continuous distributions with convex support, that

$$\lim_{n\to+\infty} a_{N,[n]} \text{ does exists}$$

under a local finiteness assumption on the possible equilibrium points. Due to the existence of several equilibrium points, especially in higher dimension, this limit may not be the solution to the greedy optimization problem at level N, but only a local minimizer. However, in practice, it turns out to be a good candidate.

Proposition 6.2. *Assume the distribution μ of X is strongly continuous (i.e. assigns no mass to hyperplanes) with a convex support denoted $C_\mu = \mathrm{supp}(\mu)$. Then the above sequence $(a_{N,[n]})_{n\ge 0}$ is bounded and there exists $\ell \in \big[e_2(a^{(N)}), e_2(a^{(N-1)} \cup \{a_{[0]}\}\big)$ such that the set $\mathcal{A}_\infty(a_{[0]})$ of its limiting points is a connected compact subset of the set Λ_ℓ of ℓ-*stationary points *defined by*

$$\Lambda_\ell = \Big\{ a \in \mathbb{R}^d \mid e_{2,N}\big(a^{(N-1)} \cup \{a\}\big) = \ell \;\; \textit{and} \;\; a = \mathbb{E}\big(X \mid X \in W_{N,a}\big) \Big\}$$

where $W_{N,a}$ denotes the closed Voronoi cell of a induced by the N-quantizer $a^{(N-1)} \cup \{a\}$. In particular, $e_2\big(a^{(N-1)} \cup \{a_{[n]}\}, X\big) \to \ell$ as $n \to +\infty$.

*Furthermore, if the ℓ-*stationary set *Λ_ℓ is locally finite (i.e. with a finite trace on compact sets of $\mathbb{R}^d$), then $a_{N,[n]}$ a.s. converges to some point in Λ_ℓ.*

The proof is postponed to Appendix A.2.

The true algorithm to be implemented in practice is a *randomized* version of this procedure where each conditional expectation is computed by Monte Carlo simulation (provided X can be simulated at a reasonable cost): let $(X^m)_{m\ge 1}$ be an i.i.d. sequence of copies of X (with distribution μ) defined on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$. Then, by the Strong Law of Large Numbers,

$$a_{N,[n+1]} = \lim_{M\to+\infty} \frac{\sum_{m=1}^M X^m \mathbf{1}_{\{X^m \in W_{N,[n]}\}}}{\sum_{m=1}^M \mathbf{1}_{\{X^m \in W_{N,[n]}\}}} \quad \mathbb{P}\text{-}a.s.$$

▷ Sequential Competitive Learning Vector Quantization procedure: Let $(\gamma_n)_{n\geq 1}$ be a sequence of $(0,1)$-valued step parameters satisfying a so-called decreasing step assumption: $\sum_n \gamma_n = +\infty$ and $\sum_n \gamma_n^2 < +\infty$. Then set

$$a_{N,[n+1]} = a_{N,[n]} - \gamma_{n+1}\mathbf{1}_{\{|X^{n+1}-a_{N,[n]}|<\min_{a\in a^{(N-1)}}|X^{n+1}-a|\}}\big(a_{N,[n]} - X^{n+1}\big),\ a_{N,[0]}\in\mathbb{R}^d.$$

One may conjecture and experimentally check, at least for distribution with compact convex support,

$$\lim_{n\to+\infty} a_{N,[n]} = a_N.$$

If so is the case, one may apply the so-called *Ruppert-Polyak principle* which states that choosing a "slowly decreasing" step of the form $\gamma_n = \frac{c}{c+n^\alpha}$, $\frac{1}{2}<\alpha<1$, and averaging the procedure by setting

$$\bar{a}_{N,[n]} = \frac{1}{n}\big(a_{[N,0]}+\cdots+a_{[N,n-1]}\big),\ n\geq 1,$$

will speed up the convergence or, to be more precise, will satisfy a Central Limit Theorem at rate $\sqrt{n}$ with the lowest possible asymptotic variance (see *e.g.* [23, 29] for details).

▷ Randomized Greedy Lloyd's I randomized procedure for the bi-variate normal distribution

Let $\mu = \mathcal{N}(0;I_2)$ be the bi-variate normal distribution on the plane. Figure 3 below depicts the graph of $N\mapsto\sqrt{N}e_2\big(a^{(N)},\mu\big)$ for $N=1$ up to $1\,000$ (and Figure 4 depicts $a^{(1000)}$). This suggests that this sequence remains bounded. However, we are not sure with such a rough procedure that the computed sequence $(a_N)_{N\geq1}$ is the optimal greedy one: at each step/level, there are clearly many local parasitic minima and one should add, prior to computing a_N, a pre-processing phase, like in one dimension, in order to choose among the areas defined by the Delaunay triangulation attached to $a^{(N-1)}$, the one which induces the minimal inertia. But this phase is numerically demanding and has not yet been included in the existing script.

The randomized Lloyd 1 has been implemented at each level N with $M=M(N)=1\,000\times N$ *i.i.d.* simulations of the $\mathcal{N}(0;I_2)$ distribution. Following Zador's Theorem

$$\lim_N \sqrt{N}e_{2,N}\big(\mathcal{N}(0;I_2)\big) = 2\sqrt{2\pi}\widetilde{J}_{2,2} = \frac{2}{3}\sqrt{\frac{5\,\pi}{\sqrt{3}}}\approx 2.0077$$

since, owing to [19] (Theorem 8.15, p.120, and Examples 8.12, p.116, devoted to hexagon lattices), $\widetilde{J}_{2,2}=\frac{1}{3}\sqrt{\frac{5}{2\sqrt{3}}}$. The "loss" is less than $10\,\%$. We verify on our own numerical experiments carried out with $N=1000$ that it is likely that

$$\sup_{1\leq N\leq 1000}\sqrt{N}\,e_2\big(a^{(N)},\mathcal{N}(0;I_2)\big)\precsim 2.18.$$

As already mentioned, it suggests again that the greedy quantization sequence outperforms the concatenated sequence (5.13) since $2.18<\sqrt{2}\times\frac{2}{3}\sqrt{\frac{5\pi}{\sqrt{3}}}\approx 2.8392$ (even if one may guess that the factor $2^{\frac{1}{d}}=\sqrt{2}$ is probably too conservative in practice).

7 Greedy quantization versus Quasi-Monte Carlo?

Of course, for every integer $N\geq1$, the *weights* induced by the μ-mass of the Voronoi cells associated to $a^{(N)}$ define canonically a sequence of N-tuples which usually cannot be "arranged" into a sequence,

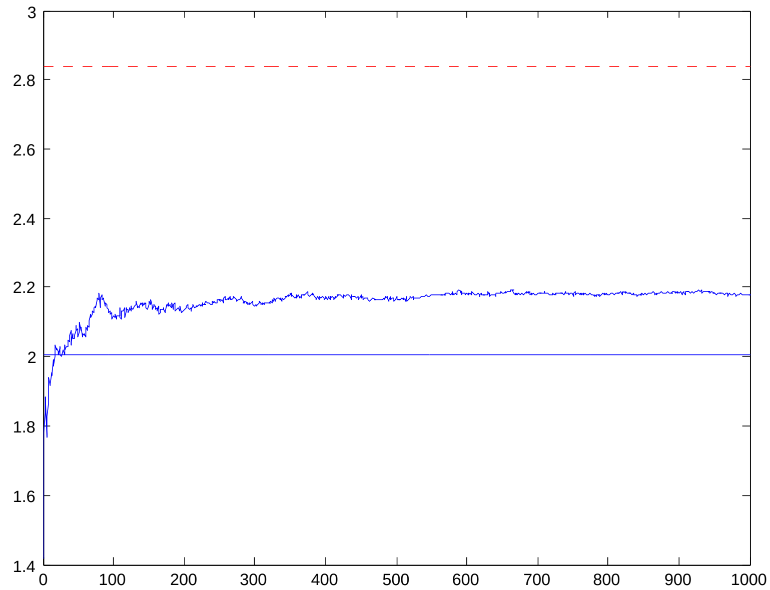

Figure 3: *Graph* $N \mapsto \sqrt{N}\, e_2\big(a^{(N)}, \mathcal{N}(0; I_2)\big)$*,* $N = 1, \ldots, 10^3$*, computed by the randomized greedy Lloyd I procedure* ($M = M(N) = 1\,000 \times N$*,* $N = 1, \ldots, 10^3$)*. Flat solid line (——) depicts Zador's constant* $\widetilde{J}_{2,2} = \frac{2}{3}\sqrt{\frac{5\,\pi}{\sqrt{3}}}$*; flat dashed line (− − −) depicts the natural upper bound for the concatenated sequence.*

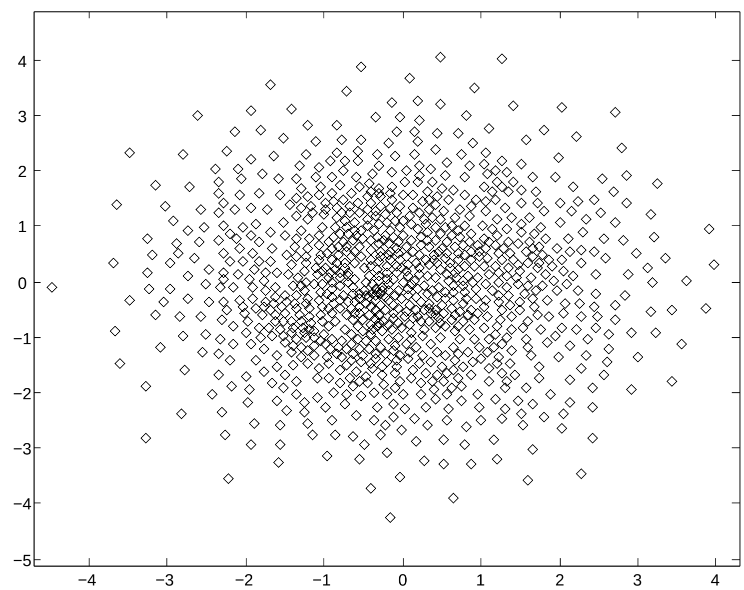

Figure 4: *Greedy quantizer* $a^{(1000)}$ *for the* $\mathcal{N}(0; I_2)$ *distribution computed by the randomized greedy Lloyd I procedure with a simulation of size* $M = 10^6$*.*

even up to a re-scaling. When considering the unit hypercube $[0,1]^d$ as a state space in d dimension, it is easy natural to compare an optimal greedy sequence with respect to the uniform distribution $U([0,1]^d)$ and the so-called *uniformly distributed sequences* usually implemented in the Quasi-Monte Carlo method.

Let us recall that a sequence $(\xi_k)_{k\geq 1}$ is uniformly distributed over $[0,1]^d$ if the empirical measures $\nu_{_N} = \frac{1}{N}\sum_{i=1}^{N}\delta_{\xi_i}$ weakly converges toward the Lebesgue measure λ_d on $[0,1]^d$. In particular this means that for every bounded $d\xi$-$a.s.$ continuous function $f:[0,1]^d\to\mathbb{R}$, $\frac{1}{N}\sum_{i=1}^{N} f(\xi_i)\to\int_{[0,1]^d} f(\xi)d\xi$. This means that the weights associated to a uniformly distributed sequence are by definition all equal to $\frac{1}{N}$ which leads to a simple normalization factor $1/N$. What is the cost induced by these uniform weights $\frac{1}{N}$, compared to the primal weights deduced from the cell (hyper-)volumes of the Voronoi diagram of $\xi_1,\ldots,\xi_{_N}$? The answer is essentially $\log N$ and si provided by Proinov's theorem (see [36]) recalled below which evaluates precisely the convergence rate of empirical measures of uniformly distributed sequences on Lipschitz functions.

In the Quasi-Monte Carlo (QMC) method, the performance of an N-tuple $(\xi_1,\ldots,\xi_{_N})\in([0,1]^d)^N$ is measured by the Kolmogorov-Smirnov distance between the extended cumulative distribution function of its empirical measure $\frac{1}{N}\sum_{1\leq i\leq N}\delta_{\xi_i}$ and the uniform distribution $U([0,1]^d)$, namely the so-called *star discrepancy* defined by

$$D^*_N(\xi_1\ldots,\xi_{_N}) = \sup_{x\in[0,1]^d}\left|\frac{1}{N}\sum_{i=1}^{N}\mathbf{1}_{\{\xi_i\in[\![0,u]\!]\}} - \lambda_d\big([\![0,u]\!]\big)\right| \tag{7.18}$$

where $[\![0,u]\!] = \prod_{\ell=1}^{d}[0,u^\ell]$, $u=(u^1,\ldots,u^d)$.

Several sequences $\xi=(\xi_N)_{N\geq1}$ have been exhibited (see [25]) whose star discrepancy at the origin satisfies for a real constant $C(\xi)\in(0,+\infty)$,

$$\forall\, N\geq 1,\quad D^*_N(\xi_1\ldots,\xi_{_N})\leq C(\xi)\frac{(1+\log N)^d}{N}. \tag{7.19}$$

Among them one can cite the p-adic $VdC(p)$ sequences ($p\geq 2$ in 1-dimension) and, when $d\geq 2$, the Halton sequences (whose i^{th} component is the $VdC(p_i)$ sequence where the "basis" p_i, $i=1,\ldots,d$, are the first d prime numbers), the Faure sequences, the Sobol' sequences (a unifying framework has been developed by Niederreiter, see *e.g.* [25]). For definitions of these sequences and numerical tests on various problems we refer to [6, 32]. Although such a rate has never been proved to be the lowest possible, this opinion is commonly shared by the QMC community (however see again [25] or [29] for a review of existing lower bounds).

The striking fact with these sequences satisfying (7.19), called *sequences with low discrepancy*, is that when they are implemented on the class of functions with *finite variation* on $[0,1]^d$ the Koksma-Hlawka inequality implies that, for every such function $f:[0,1]^d\to\mathbb{R}$

$$\left|\frac{1}{N}\sum_{k=1}^{N}f(\xi_k)-\int_{[0,1]^d}fd\lambda_d\right|\leq V(f)D^*_N(\xi_1\ldots,\xi_{_N}) \tag{7.20}$$

where $V(f)$ denotes the variation of the function f. So it induces for this specific class of functions a rate of numerical integration of order $O\Big(\frac{(\log N)^d}{N}\Big)$. In one dimension ($d=1$), However, the above notion of finite variation coincides with the standard definition of finite variation in real analysis.

When $d \geq 2$, several definitions can be given, the most popular being the finite variation in the Hardy & Krause sense (as described *e.g.* in [25]). Another slightly less general – but more elementary – being the finite variation in the signed measure sense developed in [6] (see also [29]). Unfortunately, as the dimension d increases, the set of functions with finite variation (in any of the above senses) becomes somewhat "sparse" among the set of all real-valued Borel functions defined on $[0,1]^d$. So this striking behavior may be considered as not significant when dealing with practical simulation problems. However to carry out a comparison, we need to evaluate their performances the same significant functional space, namely that of Lipschitz continuous. The Proinov theorem below provides an answer.

Theorem 7.1 (Proinov [36]). *Assume $\mathbb{R}^d$ is equipped with the ℓ^∞-norm $|(\xi^1,\dots,\xi^d)|_\infty = \max_{1\leq i\leq d}|\xi^i|$. For every continuous function $f:[0,1]^d \to \mathbb{R}$, we define uniform continuity modulus of f (with range $\delta \in [0,1]$) by*

$$w(f,\delta) := \sup_{\xi,\xi'\in[0,1]^d,\,|\xi-\xi'|_\infty\leq\delta} |f(\xi)-f(\xi')|.$$

(a) Let $(\xi_1,\dots,\xi_N)\in([0,1]^d)^N$. For every continuous function $f:[0,1]^d\to\mathbb{R}$,

$$\left|\int_{[0,1]^d} f(x)dx - \frac{1}{N}\sum_{i=1}^N f(\xi_i)\right| \leq C_d\, w\big(f, D^*_N(\xi_1,\dots,\xi_N)^{\frac{1}{d}}\big)$$

where $C_d\in(0,\infty)$ is a universal optimal real constant only depending on d. In particular, if the function $f:[0,1]^d\to\mathbb{R}$ is ℓ^∞-Lipschitz continuous with coefficient $[f]_{\rm Lip} := \sup_{x,y\in[0,1]^d}\frac{|f(x)-f(y)|}{|x-y|_\infty}$, then

$$\left|\int_{[0,1]^d} f(x)dx - \frac{1}{N}\sum_{i=1}^N f(\xi_i)\right| \leq C_d\,[f]_{\rm Lip} D^*_N(\xi_1,\dots,\xi_N)^{\frac{1}{d}}.$$

If $d=1$, $C_d=1$ and if $d\geq 2$, $C_d\in[1,4]$.

(b) In particular if $(\xi_N)_{N\geq 1}$ is a sequence with low discrepancy in the above sense, then

$$\left|\int_{[0,1]^d} f(x)dx - \frac{1}{N}\sum_{i=1}^N f(\xi_i)\right| \leq C_d\,[f]_{\rm Lip} C(\xi)\frac{1+\log N}{N^{\frac{1}{d}}}.$$

Corollary 7.1. *(a) For every N-tuple $(\xi_1,\dots,\xi_N)\in([0,1]^d)^N$*

$$e_1\big(\xi_1,\dots,\xi_N, U([0,1]^d)\big) \leq C_d D^*_N(\xi_1,\dots,\xi_N)^{\frac{1}{d}}.$$

(b) In particular, when $d=1$, $e_1\big(\xi_1,\dots,\xi_N,U([0,1])\big)\leq D^*_N(\xi_1,\dots,\xi_N)$.

Proof (of (b)). Assume $d=1$. The function $f_\xi : u\mapsto \min_{1\leq i\leq N}|u-\xi_i|$ defined on $[0,1]$ is 1-Lipschitz continuous, hence has finite variation with $V(f_\xi)=1$. Then Koksma-Hlawka Inequality (7.20) or Proinov's error bound in (a) both imply that

$$\begin{aligned} e_1\big(\xi_1,\dots,\xi_N,U([0,1])\big) &= \left|\frac{1}{N}\sum_{i=1}^N f_\xi(\xi_i) - e_1\big(\xi_1,\dots,\xi_N,U([0,1])\big)\right| \\ &= D^*_N(\xi_1,\dots,\xi_N). \quad \square \end{aligned}$$

The above claim (b) and the corollary both emphasize the fact that considering uniform weights $\frac{1}{N}$ induces the loss of a $\log N$ factor compared to an optimal (or simply rate optimal) greedy sequence for optimal quantization since, for such an $(L^1, U([0,1]))$ greedy optimal sequence $a = (a_N)_{N\geq 1}$, one has

$$\left|\int_{[0,1]^d} f(x)dx - \sum_{i=1}^{N} w_i^{(N)} f(a_i)\right| \leq \kappa(a)\,[f]_{\rm Lip}\frac{1}{N^{\frac{1}{d}}}$$

where the N-tuple $(w_i^{(N)})_{1\leq i\leq N}$ is vector of hyper-volumes (Lebesgue measure) of the Voronoi cells attached to $a^{(N)}$. Of course the practical implementation of such greedy sequences remains more demanding since one needs to have access to these N-tuples of weights.

However, by contrast, optimal quantization based cubature formulas turn out to be efficient (accurate) for much lower values of N than sequences with low discrepancy (see *e.g.* the numerical experiment carried out in [30] dealing with the pricing of European derivatives).

References

[1] Abramovicz, M., Stegun I.A. (1964). *Handbook of Mathematical Functions*, National Bureau of Standards, Washington, 1046 pp.

[2] Bally V., Pagès G. (2003). A quantization algorithm for solving discrete time multidimensional optimal stopping problems, *Bernoulli*, **9**(6):1003-1049.

[3] Bally V., Pagès G. (2003). Error analysis of the quantization algorithm for obstacle problems, *Stochastic Process. Appl.*, **106**(1):1-40.

[4] V. Bally, G. Pagès and J. Printems (2005): A quantization method for pricing and hedging multi-dimensional American style options, *Mathematical Finance*, **15**(1):119-168.

[5] Benveniste A., Métivier M. and Priouret P. (1987). *Algorithmes adaptatifs and approximations stochastiques.* Masson, Paris, 367 pp. English updated translation by Wilson S.S. (2012). *Adaptive Algorithms and Stochastic Approximations*, Springer-Verlag Berlin and Heidelberg, xi+364 pp.

[6] Bouleau N., Lépingle D. (1994). *Numerical methods for stochastic processes*, Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics. A Wiley-Interscience Publication. John Wiley & Sons, Inc., New York, 359 pp.

[7] Bouton C., Pagès G. (1993). Self-organization and *a.s.* convergence of the one-dimensional Kohonen algorithm with non-uniformly distributed stimuli. *Stochastic Process. Appl.*, **47**(2):249-274.

[8] Brandejsky A., de Saporta B. and Dufour F. (2013). Optimal stopping for partially observed piecewise-deterministic Markov processes, *Stochastic Process. Appl.*, **123**(8):3201-3238.

[9] Brandejsky A., de Saporta B. and Dufour F. (2012). Numerical method for impulse control of piecewise deterministic Markov processes, *Automatica*, **48**(5):779-793.

[10] Brancolini A., Butazzo G., Santambrogio F. and Stepanov E. (2009). Long-term planning *versus* short term planning in the asymptotical location problem, *ESAIM: COCV*, **15**:509-524.

[11] Cardot H., Cénac P. and Monnez J.-M. (2012). A fast and recursive algorithm for clustering large dataset with k-medians, *Computational Statistics & Data Analysis*, **56**(6):1431-1449.

[12] Corsi M., Pham H. and Runggaldier W. (2008). Numerical approximation by quantization of control problems in finance under partial observation, chapter from *Mathematical Modeling and Numerical Methods in Finance*, (special volume, A. Bensoussan, Q. Zhang guest eds.), *Handbook of Numerical Analysis* **15** (P.G. Ciarlet Editor), North Holland, 325-360.

[13] Davis M. (1993). *Markov Models and Optimization*, Monographs on Statistics & Applied Probability, **49**, Chapman & Hall/CRC, 308 pp.

[14] Delattre, S., Graf S., Luschgy H. and Pagès G. (2004). Quantization of probability distributions under norm-based distortion measures, *Statist. Decisions* , **22**(4):261-282.

[15] Du Q., Faber V. and Gunzburger M. (1999): Centroidal Voronoi tessellations: Applications and algorithms, *SIAM Review*, **41**:637-676.

[16] Fort J.-C., Pagès G. (2002). Asymptotics of optimal quantizers for some scalar distributions. *J. Comput. Appl. Math.* **146**(2):253-275.

[17] Gersho A., Gray R.M. (1991). *Vector Quantization and Signal Compression*, Springer International Series in Engineering and Computer Science, Springer, **159**, Berlin, 732 pp.

[18] Gobet E., Pagès G., Pham H. and Printems J. (2007). Discretization and simulation of the Zakai Equation, *SIAM J. on Numerical Analysis*, **44**(6):2505-2538.

[19] Graf S., Luschgy H. (2000). *Foundations of Quantization for Probability Distributions*. Lecture Notes in Math. 1730. Springer, Berlin, 203 pp.

[20] Graf S., Luschgy H. and Pagès G. (2007). Optimal quantizers for Radon random vectors in a Banach space, *J. of Approximation*, **144**:27-53.

[21] Graf S., Luschgy H. and Pagès G. (2008). Distortion mismatch in the quantization of probability measures, *ESAIM P&S*, **12**:127-154.

[22] Kieffer J.C. (1982). Exponential rate of convergence for Lloyd's method I, *IEEE Trans. on Inform. Theory*, Special issue on quantization, **28**(2):205-210.

[23] Luschgy, H. (2012). *Martingale in diskreter Zeit*, Theorie und Anwendungen Reihe: Springer-Lehrbuch Masterclass, Springer, Berlin, 452 pp.

[24] Luschgy H. and Pagès G. (2008). Functional quantization rate and mean regularity of processes with an application to Lévy processes, *Annals of Applied Probability*, **18**(2):427-469.

[25] Niederreiter H. (1992). *Random Number Generation and Quasi-Monte Carlo Methods*, CBMS-NSF regional conference series in Applied Mathematics, SIAM, Philadelphia, 241 pp..

[26] Pagès G. (1998). A space vector quantization method for numerical integration, *J. Computational and Applied Mathematics*, **89**:1-38 (Extended version of "Voronoi Tessellation, space quantization algorithms and numerical integration", in: M. Verleysen (Ed.), Proceedings of the ESANN' 93, Bruxelles, Quorum Editions, (1993), 221-228.

[27] Pagès G., Pham H. (2005). Optimal quantization methods for nonlinear filtering with discrete-time observations, *Bernoulli*, **11**(5):893-932.

[28] Pagès G., Pham H. and Printems J. (2005). Optimal quantization methods and applications to numerical problems in finance, *Handbook on Numerical Methods in Finance* (S. Rachev, ed.), Birkhauser, Boston, 253-298.

[29] Pagès G. *Introduction to Numerical Probability and Applications to Finance*, to appear, coll. Universitext, Springer.

[30] Pagès G., Printems J. (2005). Functional quantization for numerics with an application to option pricing, *Monte Carlo Methods & Applications J.*, **11**(4):407-446.

[31] Pagès G., Printems J. (2009). Optimal quantization for finance: from random vectors to stochastic processes, chapter from *Mathematical Modeling and Numerical Methods in Finance* (special volume, A. Bensoussan, Q. Zhang guest eds.), coll. *Handbook of Numerical Analysis* (P.-G. Ciarlet Editor) **15**, North Holland, 595-649.

[32] Pagès G., Xiao Y.J. (1997). Sequences with low discrepancy and pseudo-random numbers : theoretical results and numerical tests, *Journal of Statistical Computation and Simulation*, **56**:163-183.

[33] Pham H., Runggaldier W., Sellami A. (2004). Approximation by quantization of the filter process and applications to optimal stopping problems under partial observation, *Monte Carlo Methods and Applications*, **11**(1):57-81, DOI: 10.1515/1569396054027283.

[34] Pagès G., Yu J. (2013). Pointwise convergence of the Lloyd algorithm in higher dimension, Technical report PMA 1604.

[35] Pollard D. (1982). A central limit theorem for k-means clustering, *Ann. Probab.*, **10**, 919-926.

[36] Proinov P.D. (1988). Discrepancy and integration of continuous functions, *J. of Approx. Theory*, **52**:121-131.

A Appendix: Greedy Lloyd's I procedure

A.1 The one-dimensional greedy Lloyd I procedure

The first is to establish the uniqueness of the equilibrium point a_N satisfying (6.15) and the convergence of the Lloyd I procedure *at level* N toward this point, but with the significant additional constraint that the endpoints of the (closed convex) support of the strongly unimodal distribution μ are *active* (though fixed). By active we mean that, when finite, they have there own Voronoi area. To be more precise we will show the following proposition

Proposition A.1. *Let μ be a distribution on the real line with* log-concave *density : varphi (i.e. strongly unimodal) with a finite second moment. Then the support $I = \overline{\{\varphi > 0\}})$ is closed interval with endpoints $a, b \in \overline{\mathbb{R}}$. If a or b are finite, one may assume without loss of generality that $\varphi(a)$ or $\varphi(b) > 0$ (so that I is closed). Then the quantization problem at level N with active finite endpoints (if any) reads*

$$\min_{x\in I}\left[\varphi(x) := \mathbb{E}\left(|X-a|\wedge|X-b|^2\wedge|X-x|^2\right)\right]$$

(note that when a or b are infinite, the corresponding terms in the above expectation can be omitted).
(a) *The function φ is differentiable on I with a derivative given, for every $x\in I$, by*

$$\varphi'(x) = \frac{1}{2}\int_{\frac{a+x}{2}}^{\frac{x+b}{2}}(x-\xi)\mu(d\xi).$$

Furthermore $\operatorname{argmin}_I \mathcal{G}$ *is reduced to a single (stationary) point x^* satisfying $\varphi'(x^*) = 0$ i.e.*

$$x^* = \Phi(x^*) \quad \text{where} \quad \Phi(x) = \frac{K_\mu\left(\frac{b+x}{2}\right) - K_\mu\left(\frac{a+x}{2}\right)}{F_\mu\left(\frac{b+x}{2}\right) - F_\mu\left(\frac{a+x}{2}\right)}$$

and F_μ and K_μ denote the cumulative distribution and first moment functions of the distribution μ respectively.
(b) *The* greedy Lloyd I *procedure defined by*

$$x_{n+1} = \Phi(x_n),\ x_0\in I$$

converges toward x^*

This result can be seen as a variant of the Lloyd procedure at levels N ($N = 1$ up tp 3), depending on the finiteness of the endpoints of the interval I.

Proof. First note that, when both endpoints are infinite and cannot be active, the above statement becomes trivial since $\mathcal{G}(x) = \mathbb{E}|X-x|^2$ which attains its minimum at $x^* = \mathbb{E}\,X$, whereas the Lloyd I procedure reads $x_1 = \mathbb{E}\,X$, $n \geq 1$, whatever the starting point x_0 is.

Otherwise, if a or b are finite, we may assume, up to a symmetry-translation, that $a = 0$ and $b\in(0,+\infty]$.
(a) Elementary computations show that, for every $x\in I$,

$$\mathcal{G}(x) = \frac{1}{2}\left[\int_0^{\frac{x}{2}}\xi^2\mu(d\xi) + \int_{\frac{x}{2}}^{\frac{x+b}{2}}(x-\xi)^2\mu(d\xi) + \int_{\frac{x+b}{2}}^{+\infty}(x-\xi)^2\mu(d\xi)\mathbf{1}_{\{b<+\infty\}}\right]$$

$$\mathcal{G}'(x) = \int_{\frac{x}{2}}^{\frac{x+b}{2}}(x-\xi)\mu(d\xi)$$

and

$$\mathcal{G}''(x) = F_\mu\big(\frac{x+b}{2}\big) - F_\mu\big(\frac{x}{2}\big) - \Big(\frac{x+b}{2} - x\Big) f\big(\frac{x+b}{2}\big) - \Big(x - \frac{x}{2}\Big) f\big(\frac{x}{2}\big).$$

In what follows we focus on the case $b < +\infty$. The case $b = +\infty$ can be handled likewise (in fact in an easier way).

Note that $\mathcal{G}'(0) = -\int_0^{\frac{1}{2}} \xi\mu(d\xi) < 0$ and $\mathcal{G}'(b) = \int_{\frac{1}{2}}^{1}(b-\xi)\mu(d\xi) > 0$ so that $\mathcal{G}$ has at least one zero on $(0, b)$. (When $b = +\infty$, the existence follows form the fact that $\mathcal{G}$ does attain a minimum on $(0, +\infty)$ since $\lim_{x\to+\infty} \mathcal{G}(x) = +\infty$.)

Set $y_1 = \frac{x}{2}$ and $y_2 = \frac{x+b}{2}$. If we assume that x is a solution to $x = \Phi(x)$ (or equivalently to the stationary point equation $\mathcal{G}'(x) = 0$), we can plug this expression for x into the above equation for $\mathcal{G}''(x)$ so that $\mathcal{G}''(x)$ can be expressed as a function of y_1 and y_2 ias follows:

$$\mathcal{G}''(x) = \frac{\widetilde{\Phi}(y_1, y_2)}{F_\mu(y_2) - F_\mu(y_1)},\; y_1 < y_2,\; y_1, y_2 \in I,$$

with

$$\widetilde{\Phi}(y_1, y_2) = \big(F_\mu(y_2) - F_\mu(y_1)\big)^2 + \big(K_\mu(y_2) - K_\mu(y_1)\Big)\big(\varphi(y_2) - \varphi(y_1)\big) - \big(F_\mu(y_2) - F_\mu(y_1)\big)\big(y_2\,\varphi(y_2) - y_1\,\varphi(y_1)\big).$$

Now we consider y_1 and y_2 as free variables living in I such that $y_1 \le y_2$. First we note that $\widetilde{\Phi}(y, y) = 0$. Then, denoting by φ'_r the right derivative of the log-concave function φ, we compute the following two (right) partial derivatives of $\widetilde{\Phi}$:

$$\left(\frac{\partial\widetilde{\Phi}}{\partial y_1}\right)_r(y_1, y_2) \;=\; \big(F_\mu(y_2) - F_\mu(y_1)\big)\big(y_1\varphi'_r(y_1) - \varphi(y_1)\big) + (y_2 - y_1)\varphi(y_1)\varphi(y_2) - \varphi'_r(y_1)\big(K_\mu(y_2) - K_\mu(y_1)\big)$$

and

$$\left(\frac{\partial^2\widetilde{\Phi}}{\partial y_1\partial y_2}\right)_r(y_1, y_2) \;=\; (y_2 - y_1)\Big(\varphi(y_1)\varphi'_r(y_2) - \varphi'_r(y_1)\varphi(y_2)\Big).$$

As $(\log\varphi)'_r = \frac{\varphi'_r}{\varphi}$ is non-increasing in I, it follows that $\frac{\partial^2\widetilde{\Phi}}{\partial y_1\partial y_2}(y_1, y_2) < 0$ if $y_1 < y_2$ so that $y_2 \mapsto \left(\frac{\partial\widetilde{\Phi}}{\partial y_1}\right)_r(y_1, y_2)$ is (strictly) decreasing on $[y_1, b)$ which in turn implies it is positive on (y_1, b). This shows that $\Phi(y_1, y_2) > 0$ for every $y_1, y_2 \in I$, $y_1 < y_2$. As a consequence, any stationary point x satisfies $\mathcal{G}''(x) < 0$ *i.e.* is a strict local minimum of $\mathcal{G}$. This implies uniqueness of the solution to the equation $\mathcal{G}'(x) = 0$ by an elementary one dimensional "mountain pass" argument.

(b) In this second claim, we use again a random variable X with distribution μ. By Proposition 2.1(b), we know that if $W_{[n]} = [\frac{x_n}{2}, \frac{x_n+b}{2}]$ denotes the closed Voronoi cell of x_n with respect to $\{0, x_n, b\}$ (if b is finite, or $\{0, x_n\}$ otherwise) then

$$\mathbb{E}_\mu\big(|X - x_{n+1}|^2\mathbf{1}_{\{X\in W_{[n]}\}}\big) \le \mathbb{E}_\mu\big(|X - x_n|^2\mathbf{1}_{\{X\in W_{[n]}\}}\big).$$

with equality iff $x_{n+1} = x_n$ which is equivalent to $x_n = x^*$ (see claim (a) above). Decomposing $|X - x_{n+1}|^2$ on the Voronoi partition $[0, \frac{x_n}{2}] \cup [\frac{x_n}{2}, \frac{x_n+b}{2}] \cup [\frac{x_n+b}{2}, b]$ of I, one derives that $\mathcal{G}(x_{n+1}) < \mathcal{G}(x_n)$ as soon as $x_n \neq x^*$. the function $\mathcal{G}$ being non-negative $\mathcal{G}(x_n) \to \ell$ as $n \to \infty$.

When b is finite the sequence $(x_n)_{n\ge0}$ is trivially bounded. When $b = +\infty$, assume there exists a subsequence $x_{n'} \to +\infty$. By combining the above monotony property and Fatou's Lemma, we get

$$\mathbb{E}|X|^2 \wedge |X - x_0|^2 = \mathcal{G}(x_0) \ge \liminf_n \mathcal{G}(x_{n'}) \ge \mathbb{E}\,X^2$$

which implies that $X^2 \le (X - x_0)^2$ $\mathbb{P}$-$a.s.$ This is clearly not satisfied on the event $\{X \in [0, \frac{x_0}{2})\}$ which has positive probability. Consequently, $(x_n)_n$ is always bounded.

Then let $x_\infty = \lim_{n\to+\infty} x_{n'}$ be a limiting value of the I-valued sequence $(x_n)_{n\ge0}$. Up to a new extraction, still denoted (n'), one may assume that $x_{n'+1}$ converges toward a limiting value x'_∞ as well. Passing to the limit owing to continuity we get

$$x'_\infty = \int_{\frac{x_\infty}{2}}^{\frac{x_\infty+b}{2}} \xi\mu(d\xi).$$

One shows as above that, except if $x'_\infty = x_\infty = x^*$, $\mathcal{G}(x'_\infty) < \mathcal{G}(x_\infty)$ which cannot be true since the sequence $(\mathcal{G}(x_n))_{n\geq 0}$ converges to ℓ^2. Consequently, $x^* = x_\infty$ is the only possible limiting value for the bounded sequence $(x_n)_{n\geq 0}$ *i.e.* its limit. $\square$

Proof of Proposition 6.1. The result follows by applying the above result to the procedure on the interval $\left[\dfrac{a^{(N-1)}_{i_0}}{2}, \dfrac{a^{(N-1)}_{i_0+1}}{2}\right]$ of maximal inertia. $\square$

Remark. If we choose $a_{[0]}$ inside an interval which *has not* the highest local inertia, the procedure will still converge since we never use this fact throughout the proof of the convergence. The resulting limit will live in the same interval as the starting value since the algorithm leaves each interval stable by an obvious convexity argument. So the greedy Lloyd I procedure yields potentially $N+1$ "candidates" corresponding to each possible starting interval, but only one (issued from the interval with the highest local inertia) is solution to the greedy optimal quantization problem.

A.2 The multi-dimensional greedy Lloyd I procedure (proof of Proposition 6.2)

We assume in this section that μ has a convex support $C_\mu = \mathrm{supp}(\mu)$ and that $d \geq 2$. Note that in such a framework there is a major topological difference with the 1-dimensional case: a convex set not reduced to a single point remains pathwise connected when one point of its points is removed. Owing to that property, it is easy to show that the algorithm may visit with positive probability the whole support of C_μ (to be precise any nonempty open set of C_μ). Moreover, the points can no longer be naturally ordered like in 1-dimension.

To alleviate notations, we denote by $\mathcal{G}$ the $\mathbb{R}_+$-valued function $a \mapsto e_2\big(a^{(N-1)} \cup \{a\}\big)^2$ defined on C_μ by

$$\mathcal{G}(a) = e_2\big(a^{(N-1)} \cup \{a\}\big)^2 = \mathbb{E}\Big(d\big(X, a^{(N-1)} \cup \{a\}\big)^2\Big).$$

Let $a_{[0]} \in C_\mu \setminus a^{(N-1)}$. Lloyd's I procedure is defined by induction by Equation (6.17), namely

$$a_{[n+1]} = \mathbb{E}\big(X \,|\, X \in W_{N,[n]}\big) \in C_\mu$$

where $W_{N,[n]}$ denotes the (closed) Voronoi cell of $a_{[n]}$ induced by $a^{(N-1)} \cup \{a_{[n]}\}$.
Step 1: It follows from Proposition 2.1(b) that, as son as $a_{[n]}$ is not stationary, *i.e.* $a_{[n]} \neq \mathbb{E}\big(X \,|\, X \in W_{N,[n]}\big)$, one has

$$\mathbb{E}\big(|X - a_{[n+1]}|^2 \mathbf{1}_{\{X \in W_{N,[n]}\}}\big) < \mathbb{E}\big(|X - a_{[n]}|^2 \mathbf{1}_{\{X \in W_{N,[n]}\}}\big).$$

hence

$$\begin{aligned}
\mathcal{G}(a_{[n+1]}) &= \mathbb{E}\big(d(X, a^{(N-1)} \cup \{a_{[n+1]}\})^2\big) \\
&\leq \mathbb{E}\big(d(X, a^{(N-1)})^2 \mathbf{1}_{\{X \notin W_{N,[n]}\}}\big) + \mathbb{E}\big(|X - a_{[n+1]}|^2 \mathbf{1}_{\{X \in W_{N,[n]}\}}\big) \\
&< \mathbb{E}\big(d(X, a^{(N-1)})^2 \mathbf{1}_{\{X \notin W_{N,[n]}\}}\big) + \mathbb{E}\big(|X - a_{[n]}|^2 \mathbf{1}_{\{X \in W_{N,[n]}\}}\big) \\
&= \mathcal{G}(a_{[n]}).
\end{aligned}$$

hence, the (non-increasing, non-negative) sequence $(\mathcal{G}(a_{[n]}))_n$ converges to a finite limit $\ell^2 \in \mathbb{R}_+$ as $n \to +\infty$. The fact that $\ell \in \big(e_2(a^{(N)}, e_2(a^{(N-1)} \cup \{a_{[0]}\}\big)$ is obvious form what precedes.
Step 2: Assume there exists a subsequence $(a_{[n']})$ such that $|a_{[n']}| \to +\infty$ as $n \to +\infty$. Combining the above monotony of the sequence $(\mathcal{G}(a_{[n]}))_{n\geq 0}$ and Fatou's Lemma yields

$$\mathbb{E}\, d(X, a^{(N-1)})^2 \leq \liminf_n \mathbb{E}\, d(X, a^{(N-1)} \cup \{\alpha_{[n']}\})^2 \leq \liminf_n \mathcal{G}(a_{[n]}) \leq \mathcal{G}(a_{[0]})$$

But, as $a_{[0]} \in C_\mu \setminus a^{(N-1)}$, we know from Proposition 2.1(a) that $\mathcal{G}(a_{[0]}) < \mathbb{E}\, d(X, a^{(N-1)})^2$ which yields a contradiction.

Step 3: Let $a_{[\infty]}$ be a limiting value of the bounded sequence $(a_{[n]})_n$ (*i.e.* the limit of a subsequence). Up to a new extraction, we may also assume that $a_{[n+1]} \to a'_{[\infty]}$. Since $\mathcal{G}(a_{[n]})$ is non-decreasing, $\mathcal{G}(a_{[\infty]})$ and

$\mathcal{G}(a'_{[\infty]}) \leq \mathcal{G}(a_{[0]}) < e_2(a^{(N-1)}, \mu)$ so that we $a'_{[\infty]}, a'_{[\infty]} \notin a^{(N-1)}$. The distribution μ being strongly continuous (*i.e.* assigning no mass to hyperplanes), one shows *e.g.* by following the lines of the proof of Lemma 2.3 in [34] that

$$\mathbb{E}\big(X \,|\, W \in W_{N,[n]}\big) \to \mathbb{E}\big(X \,|\, X \in W_{N,[\infty]}\big)$$

where $W_{N,[\infty]}$ denotes the closed Voronoi cell of $a_{[\infty]}$ induced by $a^{(N-1)} \cup \{a_{[\infty]}\}$. Consequently

$$a'_{\infty} = \mathbb{E}\big(X \,|\, X \in W_{N,[\infty]}\big).$$

If $a_{[\infty]} \neq a_{[\infty]}$ then $\mathcal{G}(a'_{[\infty]}) < \mathcal{G}(a_{[\infty]})$ which is in a contradiction with the fact $(\mathcal{G}(a_{[n]}))_n$ converges to a finite limit $\ell^2 \in \mathbb{R}_+$ as $n \to +\infty$. Hence $a'_{[\infty]} = a_{[\infty]}$ which shows that one the one hand that $a_{[n+1]} - a_{[n]} \to 0$ as $n \to +\infty$ and that any limiting value of $(a_{[n]})_n$ is a stationary point in the sense that $a_\infty = \mathbb{E}\big(X \,|\, X \in W_{N,[\infty]}\big)$. The conclusion follows by standard topological arguments on convergence of sequences. □

B Appendix: A technical result on sequences

Lemma B.1. *Let $(A_N)_{N\geq 1}$ be a sequence of non-negative real numbers and let $\rho \in (0, +\infty)$ such that*

$$\forall\, N \geq 1,\ A_{N+1} \leq A_N - C\, A_N^{1+\rho}$$

for some real constant $C > 0$. Then there exists a real constant $K > 0$ such that

$$\forall\, N \geq 1, \quad A_N \leq K N^{-\frac{1}{\rho}}.$$

Proof. We may assume that $A_N > 0$ for every $N \geq 1$, it follows from the inequality satisfies by the sequence $(A_N)_{N\geq 1}$ that for very $N \geq 1$,

$$\frac{1}{A_{N+1}^\rho} \geq \frac{1}{A_N^\rho} \frac{1}{(1 - CA_N^\rho)^\rho} \geq \frac{1}{A_N^\rho}(1 + CA_N^\rho)^\rho$$

Now, there exists $u_0 = u_0(\rho)$ such that for every $u \in [0, u_0]$, $(1+u)^\rho \geq 1 + \frac{\rho}{2}u$. It is clear from the assumptions that $A_N \downarrow 0$, hence, there exists a large enough integer N_0 such that for every $N \geq N_0$,

$$\frac{1}{A_{N+1}^\rho} \geq \frac{1}{A_N^\rho} + \frac{C\rho}{2}$$

which in turn implies that

$$\frac{1}{A_N^\rho} \geq \frac{1}{A_{N_0}^\rho} + \frac{C\rho}{2}(N - N_0) \geq \frac{C\rho}{2}(N - N_0)$$

so that, for every $N > N_0$,

$$A_N \leq \left(\frac{2}{C\rho}\right)^{\frac{1}{\rho}} \frac{1}{(N - N_0)^{\frac{1}{\rho}}}.$$

This completes the proof. □