New Weak Error bounds and expansions for Optimal Quantization

Vincent Lemaire, Thibaut Montes, Gilles Pagès

New Weak Error bounds and expansions for Optimal Quantization

Vincent Lemaire * Thibaut Montes *† Gilles Pagès *

May 1, 2020

Abstract

We propose new weak error bounds and expansion in dimension one for optimal quantization-based cubature formula for different classes of functions, such that piecewise affine functions, Lipschitz convex functions or differentiable function with piecewise-defined locally Lipschitz or α-Hölder derivatives. These new results rest on the local behaviours of optimal quantizers, the L^r-L^s distribution mismatch problem and Zador's Theorem. This new expansion supports the definition of a Richardson-Romberg extrapolation yielding a better rate of convergence for the cubature formula. An extension of this expansion is then proposed in higher dimension for the first time. We then propose a novel variance reduction method for Monte Carlo estimators, based on one dimensional optimal quantizers.

Keywords— Optimal quantization; Numerical integration; Weak error; Romberg extrapolation; Variance reduction; Monte Carlo simulation; Product quantizer.

2010 AMS Classification: 65C05, 60E99, 65C50.

Introduction

Optimal quantization was first introduced in [She97], Sheppard worked on optimal quantization of the uniform distribution on unit hypercubes. It was then extended to more general distributions with applications to Signal transmission at the Bell Laboratory in the 50's (see [GG82]) and then developed as a numerical method in the early 90's, for expectation approximations (see [Pag98]) and later for conditional expectation approximations (see [PPP04, BPP01, BP03, BPP05]).

In modern terms, vector quantization consists in finding the projection for the L^p-Wasserstein distance of a probability measure on $\mathbb{R}^d$ with a finite p-th moment on the convex subset of Γ-supported probability measure, where Γ is a finite subset of $\mathbb{R}^d$ and $0 < p < +\infty$. The aim of Optimal Quantization is to determine the set $\Gamma_N := \{x_1^N, \dots, x_N^N\} \subset \mathbb{R}^d$ with cardinality at most N which minimizes this distance among all such sets Γ. Formally, if we consider a random vector $X \in L^p(\mathbb{P})$, we search for Γ_N, the solution to the following problem

$$\min_{\Gamma_N \subset \mathbb{R}, |\Gamma_N| \leqslant N} \|X - \widehat{X}^{\Gamma_N}\|_p$$

*Sorbonne Université, Laboratoire de Probabilités, Statistique et Modélisation, LPSM, Campus Pierre et Marie Curie, case 158, 4 place Jussieu, F-75252 Paris Cedex 5, France.

†The Independent Calculation Agent, The ICA, 5th Floor, 95 Gresham Street, London.

where $\widehat{X}^{\Gamma_N}$ denotes the projection of X onto Γ_N (often $\widehat{X}^{\Gamma_N}$ is denoted by $\widehat{X}^N$ in order to alleviate the notations). The term $\|X - \widehat{X}^{\Gamma_N}\|_p$ is often referred to as the distortion of order p. The existence of an optimal quantizer at a given level N has been shown in [GL00, Pag98] and in the one-dimensional case if the distribution of X is absolutely continuous with a *log-concave* density then there exists a unique optimal quantizer at level N. In the present paper we will consider one dimensional optimal quantizers. Moreover, we are not only interested by the existence of such a quantizer but also in the asymptotic behaviour of the distortion because it is an important feature for the method in order to determine the level of the error introduced by the approximation. The question concerning the sharp rate of convergence of $\|X - \widehat{X}^N\|_p$ as N goes to infinity is answered by Zador's Theorem. For $X \in L^{p+\delta}(\mathbb{P})$, $\delta > 0$, such that $\mathbb{P}_X(d\xi) = \varphi(\xi) \cdot \lambda(d\xi) + \nu(d\xi)$, where $\nu \perp \lambda$ is the singular component of $\mathbb{P}_X$ with respect to the Lebesgue measure λ on $\mathbb{R}^d$, the rate of convergence is given by

$$\lim_{N \to +\infty} N^{\frac{1}{d}} \|X - \widehat{X}^N\|_p = \widetilde{J}_{p,d} \left[\int_{\mathbb{R}^d} \varphi^{\frac{d}{d+p}} d\lambda_d \right]^{\frac{1}{p} + \frac{1}{d}}$$

where φ is the density of X, λ_d is the Lebesgue measure on $\mathbb{R}^d$ and $\widetilde{J}_{p,d} = \inf_{N \geqslant 1} N^{\frac{1}{d}} \|U - \widehat{U}^N\|_p$, $U \overset{\mathcal{L}}{\sim} \mathcal{U}\big((0,1)^d\big)$. For more insights on the mathematical/probabilistic aspects of Optimal quantization theory, we refer to [GL00, Pag15].

The reason for which we are interested in this optimal quantizer is numerical integration. The discrete feature of the optimal quantizer $\widehat{X}^N$ allows us to define, for every continuous function $f : \mathbb{R}^d \longrightarrow \mathbb{R}$, such that $f(X) \in L^2(\mathbb{P})$, the following quantization-based cubature formula

$$\mathbb{E}\big[f(\widehat{X}^N)\big] = \sum_{i=1}^{N} p_i f(x_i^N)$$

where $p_i = \mathbb{P}(\widehat{X}^N = x_i^N)$. Indeed, as $\widehat{X}^N$ is constructed as the best discrete approximation of X in $L^p(\mathbb{P})$, it is reasonable to approximate $\mathbb{E}\big[f(X)\big]$ by $\mathbb{E}\big[f(\widehat{X}^N)\big]$ which is useful for numerical integrations problems.

The problem of numerical integration appears a lot in applied fields, such as Physics, Computer Sciences or Numerical Probability. For example, in Quantitative Finance, many quantities of interest are of the form

$$\mathbb{E}\big[f(S_t)\big] \qquad \text{for some } t > 0,$$

where $f : \mathbb{R}^d \longrightarrow \mathbb{R}$ is a Borel function and $(S_s)_{s \in [0,t]}$ is a diffusion process solution to a Stochastic Differential Equation (SDE)

$$S_t = S_0 + \int_0^t b(s, S_s) ds + \int_0^t \sigma(s, S_s) dW_s, \qquad S_0 = s_0,$$

where W is a standard Brownian motion living on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ and b and σ are Lipschitz continuous in x uniformly with respect to $s \in [0,t]$, which are the standard assumptions in order to ensure existence and uniqueness of a strong solution to the SDE. Since it is often impossible to compute $\mathbb{E}\big[f(S_t)\big]$ directly, it has been proposed in [Pag98] to compute an optimal quantizer $\widehat{X}^N$ of X where X is a random variable having the same distribution as S_t and to use the previously defined quantization-based cubature formula as an approximation.

Another approach, often used in order to approximate $\mathbb{E}\big[f(X)\big]$, is to perform a Monte Carlo simulation $\widehat{I}_M := \sum_{m=1}^{M} f(X^m)$, where $(X^m)_{m=1,\dots,M}$ is a sequence of independent copies of X. The method's rate of convergence is determined by the strong law of numbers and the central limit theorem, which says that if X is square integrable, then

$$\sqrt{M}\Big(\widehat{I}_M - \mathbb{E}\big[f(X)\big]\Big) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0, \sigma^2_{f(X)}\right) \quad \text{as} \quad M \to +\infty$$

where $\sigma^2_{f(X)} = \mathbb{V}\mathrm{ar}\big(f(X)\big)$. One notices that, for a given M, the limiting factor of the method is $\sigma^2_{f(X)}$. Hence, a lot of methods have been developed in order to reduce the variance term: antithetic variables, control variates, importance sampling, etc. The reader can refer to [Pag18, Gla13] for more details concerning the Monte Carlo methodology and the variance reduction methods.

In this paper we propose a novel variance reduction method of Monte Carlo estimator through quantization. Our method innovates in that it uses a linear combination of one dimensional control variates to reduce the variance of a higher dimensional problem. More precisely, we introduce a quantization-based control variates Ξ^N_k for $k = 1, \dots, d$. If one considers a function $f : \mathbb{R}^d \mapsto \mathbb{R}$, we approximate $\mathbb{E}\big[f(X)\big]$ by

$$\mathbb{E}\big[f(X) - \langle \lambda, \Xi^N \rangle\big]$$

with $\langle \cdot, \cdot \rangle$ the scalar product in $\mathbb{R}^d$ and $(\Xi^N_k)_{k=1,\dots,d} := f_k(X_k) - \mathbb{E}\big[f_k(\widehat{X}^N_k)\big]$, where X_k is the k-th component of X, $\widehat{X}^N_k$ is an optimal quantizer of X_k of size N and $f_k : \mathbb{R} \mapsto \mathbb{R}$ is designed from f. Looking closely at the introduced control variates, one notices that we introduce a bias in the approximation. However, as since it is closely linked to weak error, this bias can be controlled. The present paper focuses on the weak error's rate of convergence.

First, we place ourselves in the case where X is a random variable in dimension one and we consider a quadratic optimal quantizer. We work on the rate of convergence of the weak error induced by the expectation approximation by an optimal quantization-based cubature formula for different classes of functions f

$$\lim_{N \to +\infty} N^\alpha \big| \mathbb{E}\big[f(X)\big] - \mathbb{E}\big[f(\widehat{X}^N)\big] \big| \leqslant C_{f,X} < +\infty.$$

The first classical result concerns Lipschitz continuous functions. Using directly the Lipschitz continuity property of f and Zador's Theorem a rate of order $\alpha = 1$ can be obtained. Moreover, if we consider the supremum among all functions with a Lipschitz constant upper-bounded by 1, then

$$N \sup_{[f]_{Lip} \leqslant 1} \big| \mathbb{E}\big[f(X)\big] - \mathbb{E}\big[f(\widehat{X}^N)\big] \big| = N \|X - \widehat{X}^N\|_1 \leqslant N \|X - \widehat{X}^N\|_2 \xrightarrow{N \to +\infty} C_f < +\infty.$$

A faster rate ($\alpha = 2$) can be attained for differentiable functions with Lipschitz continuous derivative, using a Taylor expansion with integral remainder and the following stationarity property of quadratic optimal quantizers

$$\mathbb{E}\big[X \mid \widehat{X}^N\big] = \widehat{X}^N.$$

Moreover, considering the supremum among all functions where the Lipschitz constant of the derivative is upper-bounded by 1, we have

$$N^2 \sup_{[f']_{Lip} \leqslant 1} \big| \mathbb{E}\big[f(X)\big] - \mathbb{E}\big[f(\widehat{X}^N)\big] \big| = \frac{1}{2} N^2 \|X - \widehat{X}^N\|_2^2 \xrightarrow{N \to +\infty} C_f < +\infty$$

where the limit is given by Zador's Theorem. A detailed summary about this results can be found in [Pag18].

In the first part of this paper, we extend this improved rate ($\alpha = 2$) to classes of less smooths functions in one dimension. These new results enable us to design efficient variance reduction methods in higher dimensional settings with in view applications to option pricing. The new results concerns the following classes of functions

- Lipschitz continuous piecewise affine functions with finitely many breaks of affinity. We use the stationarity property of the optimal quantizer on the cells where there is no break of affinity and then we control the error on the remaining cells using results on the local behaviour of the quantizer.

- Lipschitz continuous convex functions, using local behaviours results on optimal quantizers. We use a representation formula for convex functions as integrals of Ridge functions combined with the local behaviour result in order to control the error again.

- Differentiable functions with piecewise-defined locally Lipschitz derivative. The functions have K breaks of affinity $\{a_1, \dots, a_K\}$, such that $-\infty = a_0 < a_1 < \cdots < a_K < a_{K+1} = +\infty$ and the locally Lipschitz property of the derivative is defined by

$$\forall k = 0, \dots, K, \quad \forall x, y \in (a_k, a_{k+1}) \quad |f'(x) - f'(y)| \leqslant [f']_{k,Lip,loc}|x - y|\big(g_k(x) + g_k(y)\big)$$

  where $g_k : (a_k, a_{k+1}) \to \mathbb{R}_+$ are non-negative Borel functions. We use the locally Lipschitz property of the derivative combined with the L^r-L^s distortion Theorem and Zador's Theorem on the cells where there is no break of affinity and then we control the error on the remaining cells using results on the local behaviour of the quantizer.

- Differentiable functions with piecewise-defined locally α-Hölder derivative. The functions have K breaks of affinity $\{a_1, \dots, a_K\}$, such that $-\infty = a_0 < a_1 < \cdots < a_K < a_{K+1} = +\infty$ and the locally α-Hölder property of the derivative is defined by

$$\forall k = 0, \dots, K, \quad \forall x, y \in (a_k, a_{k+1}), \quad |f'(x) - f'(y)| \leqslant [f']_{k,\alpha,loc}|x - y|^\alpha\big(g_k(x) + g_k(y)\big)$$

  where $g_k : (a_k, a_{k+1}) \to \mathbb{R}_+$ are non-negative Borel functions. For this class of functions, the rate of convergence is of order $1 + \alpha$. The result is obtained using the same ideas as in the locally Lipschitz case.

Hence, for all this classes of functions, except the last one, we have

$$\lim_{N \to +\infty} N^2 \big| \mathbb{E}\big[f(X)\big] - \mathbb{E}\big[f(\widehat{X}^N)\big] \big| \leqslant C_{f,X} < +\infty.$$

In the second part of the paper we deal with the *weak error expansion* of the approximation of $\mathbb{E}\big[f(X)\big]$ by $\mathbb{E}\big[f(\widehat{X}^N)\big]$. First, we place ourselves in the one dimensional case by considering a twice differentiable function $f : \mathbb{R} \mapsto \mathbb{R}$ with a bounded Lipschitz continuous second derivative and $X : (\Omega, \mathcal{A}, \mathbb{P}) \to \mathbb{R}$. Through a second order Taylor expansion and with the help of Corollary 1.8, Theorem 1.13 and the L^r-L^s distortion mismatch Theorem we obtain

$$\mathbb{E}\big[f(X)\big] = \mathbb{E}\big[f(\widehat{X}^N)\big] + \frac{c_2}{N^2} + O(N^{-(2+\beta)})$$

where $\beta \in (0, 1)$. This expression suggests to use a Richardson-Romberg extrapolation in order to *kill* the first term of the expansion which yields

$$\mathbb{E}\big[f(X)\big] = \mathbb{E}\left[\frac{M^2 f(\widehat{X}^M) - N^2 f(\widehat{X}^N)}{M^2 - N^2}\right] + O(N^{-(2+\beta)}).$$

Second, we present a result in higher dimension when considering a twice differentiable function $f : \mathbb{R}^d \mapsto \mathbb{R}$ with a bounded Lipschitz continuous Hessian, $X : (\Omega, \mathcal{A}, \mathbb{P}) \to \mathbb{R}^d$ with independent components $(X_k)_{k=1,\dots,d}$ and $\widehat{X}^N$ a product quantizer of X with d components $(\widehat{X}_k^{N_k})_{k=1,\dots,d}$

such that $N_1 \times \cdots \times N_d \simeq N$. Using product quantizer allows us to rely on the one dimensional results for quadratic optimal quantizers and in that case we have

$$\mathbb{E}\big[f(X)\big] = \mathbb{E}\big[f(\widehat{X}^N)\big] + \sum_{k=1}^{d} \frac{c_k}{N_k^2} + O\bigg(\Big(\min_{k=1:d} N_k\Big)^{-(2+\beta)}\bigg).$$

The paper is organized as follows. First we recall some basic facts and deeper results about optimal quantization in Section 1. In Section 2, we present our new results on weak error for some classes of functions. Then, we see in Section 3 how to derive *weak error expansion* allowing us to specify the right hypothesis under which we can use a Richardson-Romberg extrapolation. Finally, we conclude with some applications. The first one is the introduction of our novel variance reduction involving optimal quantizers. The last one illustrates numerically the results shown in Section 2 and 3, by considering a Black-Scholes model and pricing different types of European Options. We also propose a numerical example for the variance reduction.

1 About optimal quantization $(d = 1)$

Let X be a $\mathbb{R}$-valued random variable with distribution $\mathbb{P}_X$ defined on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ such that $X \in L^2(\mathbb{P})$.

Definition 1.1. Let $\Gamma_N = \{x_1^N, \dots, x_N^N\} \subset \mathbb{R}$ be a subset of size N, called N-quantizer. A Borel partition $\big(C_i(\Gamma_N)\big)_{i=1,\dots,N}$ of $\mathbb{R}$ is a Voronoï partition of $\mathbb{R}$ induced by the N-quantizer Γ_N if, for every $i = 1, \dots, N$,

$$C_i(\Gamma_N) \subset \big\{\xi \in \mathbb{R}, |\xi - x_i^N| \leqslant \min_{j \neq i} |\xi - x_j^N|\big\}.$$

The Borel sets $C_i(\Gamma_N)$ are called Voronoï cells of the partition induced by Γ_N.

One can always consider that the quantizers are ordered: $x_1^N < x_2^N < \cdots < x_{N-1}^N < x_N^N$ and in that case the Voronoï cells are given by

$$C_k(\Gamma_N) = (x_{k-1/2}^N, x_{k+1/2}^N], \qquad k = 1, \dots, N-1, \qquad C_N(\Gamma_N) = (x_{N-1/2}^N, x_{N+1/2}^N)$$

where $\forall k = 2, \dots, N$, $x_{k-1/2}^N := \frac{x_{k-1}^N + x_k^N}{2}$ and $x_{1/2}^N := \inf\big(\mathrm{supp}(\mathbb{P}_X)\big)$ and $x_{N+1/2}^N := \sup\big(\mathrm{supp}(\mathbb{P}_X)\big)$.

Definition 1.2. Let $\Gamma_N = \{x_1^N, \dots, x_N^N\}$ be an N-quantizer. The nearest neighbour projection $\mathrm{Proj}_{\Gamma_N} : \mathbb{R} \to \{x_1^N, \dots, x_N^N\}$ induced by a Voronoï partition $\big(C_i(\Gamma_N)\big)_{i=1,\dots,N}$ is defined by

$$\forall \xi \in \mathbb{R}, \qquad \mathrm{Proj}_{\Gamma_N}(\xi) := \sum_{i=1}^{N} x_i^N \, \mathbb{1}_{\xi \in C_i(\Gamma_N)}.$$

We can now define the quantization of X by composing Proj_{Γ_N} and X

$$\widehat{X}^{\Gamma_N} = \mathrm{Proj}_{\Gamma_N}(X) = \sum_{i=1}^{N} x_i^N \, \mathbb{1}_{X \in C_i(\Gamma_N)}$$

and the point-wise error induced by the replacement of X by $\widehat{X}^{\Gamma_N}$ given by

$$|X - \widehat{X}^{\Gamma_N}| = \mathrm{dist}\big(X, \{x_1^N, \dots, x_N^N\}\big) = \min_{i=1,\dots,N} |X - x_i^N|.$$

In order to alleviate the notations, from now on we write $\widehat{X}^N$ in place of $\widehat{X}^{\Gamma_N}$.

Definition 1.3. The L^2-mean (or mean quadratic) quantization error induced by the replacement of X by the quantization of X using a N-quantizer $\Gamma_N \subset \mathbb{R}$ is defined as the quadratic norm of the point-wise error previously defined

$$\|X - \widehat{X}^N\|_2 := \left(\mathbb{E}\left[\min_{i=1,\dots,N} |X - x_i^N|^2 \right] \right)^{1/2} = \left(\int_{\mathbb{R}} \min_{i=1,\dots,N} |\xi - x_i^N|^2 \, \mathbb{P}_X(d\xi) \right)^{1/2}.$$

It is convenient to define the quadratic distortion function at level N as the squared mean quadratic quantization error on $(\mathbb{R})^N$:

$$\mathcal{Q}_{2,\mathrm{N}} : x = (x_1^N, \dots, x_N^N) \longmapsto \mathbb{E}\left[\min_{i=1,\dots,N} |X - x_i^N|^2 \right] = \|X - \widehat{X}^N\|_2^2.$$

Remark 1.4. All these definitions can be extended to the L^p case. For example the L^p-mean quantization error induced by a quantizer of size N is

$$\|X - \widehat{X}^N\|_p := \left(\mathbb{E}\left[\min_{i=1,\dots,N} |X - x_i^N|^p \right] \right)^{1/p} = \left(\int_{\mathbb{R}} \min_{i=1,\dots,N} |X - x_i^N|^p \, \mathbb{P}_X(d\xi) \right)^{1/p}.$$

We briefly recall some classical theoretical results, see [GL00, Pag18] for further details.

Theorem 1.5. *(Existence of optimal N-quantizers) Let $X \in L^2(\mathbb{P})$ and $N \in \mathbb{N}^*$.*

(a) The quadratic distortion function $\mathcal{Q}_{2,\mathrm{N}}$ at level N attains a minimum at an N-tuple $x^{(N)} = (x_1^N, \dots, x_N^N)$ and $\Gamma_N = \{x_i^N, i = 1, \dots, N\}$ is a quadratic optimal quantizer at level N.

(b) If the support of the distribution $\mathbb{P}_X$ of X has at least N elements, then $x^{(N)} = (x_1^N, \dots, x_N^N)$ has pairwise distinct components, $\mathbb{P}_X\left(C_i(x^{(N)})\right) > 0$, $i = 1, \dots, N$. Furthermore, the sequence $N \mapsto \inf_{x \in (\mathbb{R})^N} \mathcal{Q}_{2,\mathrm{N}}(x)$ converges to 0 and is decreasing as long as it is positive.

Following the existence of a minimum for $\mathcal{Q}_{2,\mathrm{N}}$ at $x^{(N)}$, we can define an optimal quadratic N-quantizer.

Definition 1.6. A grid associated to any N-tuple solution to the above distortion minimization problem is called an optimal quadratic N-quantizer.

A really interesting and useful property concerning quadratic optimal quantizers is the stationarity property.

Proposition 1.7. *(Stationarity) Assume that the support of $\mathbb{P}_X$ has at least N elements. Any L^2-optimal N-quantizer $\Gamma_N \in (\mathbb{R})^N$ is stationary in the following sense: for every Voronoï quantization $\widehat{X}^N$ of X,*

$$\mathbb{E}\left[X \mid \widehat{X}^N \right] = \widehat{X}^N.$$

Corollary 1.8. *If $\widehat{X}^N$ is a L^2-optimal quantization of X, hence has the above stationarity property, and $f(X) \in L^2(\mathbb{P})$ with $f : \mathbb{R} \to \mathbb{R}$ then*

$$\mathbb{E}\left[f(\widehat{X}^N)(X - \widehat{X}^N) \right] = 0.$$

Proof. The proof is straightforward, indeed

$$\begin{aligned} \mathbb{E}\left[f(\widehat{X}^N)(X - \widehat{X}^N) \right] &= \mathbb{E}\left[\mathbb{E}\left[f(\widehat{X}^N)(X - \widehat{X}^N) \mid \widehat{X}^N \right] \right] = \mathbb{E}\left[f(\widehat{X}^N)\, \mathbb{E}[X - \widehat{X}^N \mid \widehat{X}^N] \right] \\ &= \mathbb{E}\left[f(\widehat{X}^N)\big(\mathbb{E}\left[X \mid \widehat{X}^N \right] - \widehat{X}^N \big) \right] = 0. \end{aligned}$$

□

We now take a look at the asymptotic behaviour in N of the quadratic mean quantization error. We saw in Theorem 1.5 that the infimum of the quadratic distortion converges to 0 as N goes to infinity. The next Theorem, known as Zador's Theorem, analyzes the rate of convergence of the L^p-mean quantization error.

Theorem 1.9. *(Zador's Theorem) Let $p \in (0, +\infty)$.*

(a) SHARP RATE. *Let $X \in L^{p+\delta}(\mathbb{P})$ for some $\delta > 0$. Let $\mathbb{P}_X(d\xi) = \varphi(\xi) \cdot \lambda(d\xi) + \nu(d\xi)$, where $\nu \perp \lambda$ is the singular component of $\mathbb{P}_X$ with respect to the Lebesgue measure λ on $\mathbb{R}$. Then*

$$\lim_{N \to +\infty} N \min_{\Gamma_N \subset \mathbb{R}, |\Gamma_N| \leqslant N} \|X - \widehat{X}^N\|_p = \widetilde{J}_{p,1} \left[\int_{\mathbb{R}} \varphi^{\frac{1}{1+p}} d\lambda \right]^{1+\frac{1}{p}}$$

with $\widetilde{J}_{p,1} = \frac{1}{2^p(p+1)}$.

(b) NON ASYMPTOTIC UPPER-BOUND. *Let $\delta > 0$. There exists a real constant $C_{1,p,\delta} \in (0, +\infty)$ such that, for every $\mathbb{R}$-valued random variable X,*

$$\forall N \geqslant 1, \qquad \min_{\Gamma_N \subset \mathbb{R}, |\Gamma_N| \leqslant N} \|X - \widehat{X}^N\|_p \leqslant C_{1,p,\delta} \sigma_{\delta+p}(X) N^{-1}$$

where, for $r \in (0, +\infty)$, $\sigma_r(X) = \min_{a \in \mathbb{R}} \|X - a\|_r < +\infty$.

Now, we state some intuitive but remarkable results concerning the local behaviour of the optimal quantizers.

Lemma 1.10. *Let $\mathbb{P}_X$ be a distribution on the real line with connected support $I_{\mathbb{P}_X} := \mathrm{supp}(\mathbb{P}_X)$. Let $\Gamma_N = \{x_1^N, \ldots, x_N^N\}$ be a sequence of r-optimal quantizers, $r > 0$. Let $[a, b]$, be a closed interval then*

$$\bigcup_N \bigcup_{C_i(\Gamma_N) \cap [a,b] \neq \emptyset} C_i(\Gamma_N) \subset K_0$$

where K_0 is a compact set.

Proof. First, if $+\infty \notin \overline{I_{\mathbb{P}_X}}$ then the upper-bound of K_0 is the upper-bound of $\overline{I_{\mathbb{P}_X}}$ otherwise if $+\infty \in \overline{I_{\mathbb{P}_X}}$, let $b_0 \in I_{\mathbb{P}_X}$ such that $b_0 < b$, as $\mathbb{P}_X$ has a density, then $\mathbb{P}_X\big(\{b_0\}\big) = \mathbb{P}_X\big(\{b\}\big) = 0$. Considering the weighted empirical measure

$$\mathbb{P}_{\widehat{X}^N} := \sum_{i=1}^N \mathbb{P}_X\big(C_i(\Gamma_N)\big) \delta_{x_i^N} \xrightarrow{N \to +\infty} \mathbb{P}_X$$

then $\mathbb{P}_{\widehat{X}^N}\big([b_0, b]\big) \xrightarrow{N \to +\infty} \mathbb{P}_X\big([b_0, b]\big) < \mathbb{P}_X\big([b_0, +\infty)\big)$. Moreover, one notices that

$$\mathbb{P}_{\widehat{X}^N}\big([b_0, b]\big) = \mathbb{P}_X\left(\bigcup_{i \in \{i_{b_0}, \ldots, i_b\}} C_i(\Gamma_N) \right) = \mathbb{P}_{\widehat{X}^N}\left(\bigcup_{i \in \{i_{b_0}, \ldots, i_b\}} C_i(\Gamma_N) \right)$$

where $x_{i_u}^N$ is the centroid of the cell that contains u. Then, as $[b_0, x_{i_b+1/2}^N] \subset \bigcup_{i \in \{i_{b_0}, \ldots, i_b\}} C_i(\Gamma_N)$

$$\mathbb{P}_X\big([b_0, x_{i_b+1/2}^N]\big) \leqslant \mathbb{P}_{\widehat{X}^N}\big([b_0, b]\big) \xrightarrow{N \to +\infty} \mathbb{P}_X\big([b_0, b]\big) < \mathbb{P}_X\big([b_0, +\infty)\big)$$

hence, $\limsup_N x_{i_b+1/2}^N < +\infty$ and $\sup_N x_{i_b+1/2}^N < +\infty$, which gives us the upper-bound of K_0: $\sup_N x_{i_b+1/2}^N$.

Finally, if $-\infty \notin \overline{I_{\mathbb{P}_X}}$ then the lower-bound of K_0 is the lower-bound of $\overline{I_{\mathbb{P}_X}}$ otherwise if $-\infty \in \overline{I_{\mathbb{P}_X}}$, then following the same idea as above, we can apply the same deductions in order to show that $\inf_N x_{i_a-1/2}^N > -\infty$ which gives us the lower-bound of K_0: $\inf_N x_{i_a-1/2}^N$. In conclusion, $K_0 := \mathrm{supp}(\mathbb{P}_X) \bigcap [\inf_N x_{i_a-1/2}^N, \sup_N x_{i_b+1/2}^N]$. □

The next result, proved in [DFP04], deals with the local behaviour of optimal quantizer, more precisely it characterises the rate of convergence, in function of N, of the weights and the local distortions associated to an optimal quantizer. This is the key result of the first part of this paper. It allows us to extend the weak error bound of order two to less regular functions than those originally considered in [Pag98], namely differentiable functions with Lipschitz continuous derivative.

Theorem 1.11. *(Local behaviour of optimal quantizers) Let $\mathbb{P}_X$ be a distribution on the real line with connected support $\mathrm{supp}(\mathbb{P}_X)$. Assume that $\mathbb{P}_X$ has a probability density function φ which is positive and Lipschitz continuous on every compact set of the interior $(\underline{m}, \overline{m})$ of $\mathrm{supp}(\mathbb{P}_X)$. Let $\Gamma_N = \{x_1^N, \ldots, x_N^N\}$ be a sequence of stationary and L^r optimal quantizers, $r > 0$.*

(a) The sequence of functions $(\psi_N)_{N\geqslant 1}$ defined by

$$\psi_N(\xi) := N \sum_{i=1}^{N} \mathbb{1}_{C_i(\Gamma_N)}(\xi)\, \mathbb{P}_X\big(C_i(\Gamma_N)\big), \quad N \geqslant 1,$$

converges uniformly on compact sets of $(\underline{m}, \overline{m})$ towards $c_{\varphi,1/(r+1)}\varphi^{\frac{r}{r+1}}$, with $c_{\varphi,1/(r+1)} = \|\varphi\|_{1/(1+r)}^{-1/(1+r)}$ i.e., for every $[a,b] \subset (\underline{m}, \overline{m})$, $a < b$,

$$\sup_{\{i:x_i^N \in [a,b]\}} \left| N\, \mathbb{P}_X\big(C_i(\Gamma_N)\big) - c_{\varphi,1/(r+1)}\varphi^{\frac{r}{r+1}}(x_i^N) \right| \xrightarrow{N\to+\infty} 0. \tag{1.1}$$

The local distortion is asymptotically uniformly distributed i.e., for every $[a,b] \subset (\underline{m}, \overline{m})$,

$$\sup_{\{i:x_i^N \in [a,b]\}} \left| N^{r+1} \int_{C_i(\Gamma_N)} |x_i^N - \xi|^r\, \mathbb{P}_X(d\xi) - \frac{\|\varphi\|_{1/(r+1)}}{2^r(r+1)} \right| \xrightarrow{N\to+\infty} 0. \tag{1.2}$$

(b) Moreover, if $\mathbb{P}_X$ has a compact support $[\underline{m}, \overline{m}]$ and φ is bounded away from 0 on the whole interval $[m, M]$, then all the above convergences hold uniformly on $[\underline{m}, \overline{m}]$.

The next result is a weaker version of Theorem 1.11 but it is a really useful tool when dealing with weak error induced by quantization-based cubature formulas.

Corollary 1.12. *Under the same hypothesis as in Theorem 1.11 and if $1 \leqslant s \leqslant r$, we have the following result, for every $i \in \{1, \ldots, N\}$,*

$$\limsup_N N^{s+1} \int_{C_i(\Gamma_N)} |x_i^N - \xi|^s\, \mathbb{P}_X(d\xi) = \limsup_N N^{s+1}\, \mathbb{E}\left[|\widehat{X}^N - X|^s\, \mathbb{1}_{\{\widehat{X}^N = x_i^N\}}\right] < +\infty.$$

Proof. If $s = 1$, using Schwarz's inequality

$$\int_{C_i(\Gamma_N)} |x_i^N - \xi|\, \mathbb{P}_X(d\xi) \leqslant \left(\int_{C_i(\Gamma_N)} |x_i^N - \xi|^2\, \mathbb{P}_X(d\xi) \cdot \mathbb{P}_X\big(C_i(\Gamma_N)\big) \right)^{\frac{1}{2}}$$

$$\iff \qquad N^2 \int_{C_i(\Gamma_N)} |x_i^N - \xi|\, \mathbb{P}_X(d\xi) \leqslant \left(N^3 \int_{C_i(\Gamma_N)} |x_i^N - \xi|^2\, \mathbb{P}_X(d\xi) \cdot N\, \mathbb{P}_X\big(C_i(\Gamma_N)\big) \right)^{\frac{1}{2}}.$$

And applying Theorem 1.11 with $\mathbb{P}_X = \varphi \cdot \lambda$ and $r = 2$, one derives

$$\limsup_N N^2 \int_{C_i(\Gamma_N)} |x_i^N - \xi|\, \mathbb{P}_X(d\xi) \leqslant \frac{1}{2\sqrt{3}} \big(c_{\varphi,1/3} \|\varphi\|_{1/3} \|\varphi^{2/3}\|_\infty\big)^{\frac{1}{2}} < +\infty.$$

Otherwise, for $1 < s < r$, using Hölder's inequality with $p = \frac{1}{s}$ and $q = \frac{1}{1-s}$

$$\begin{aligned}
\int_{C_i(\Gamma_N)} |x_i^N - \xi|^s \, \mathbb{P}_X(d\xi) &\leqslant \left(\int_{C_i(\Gamma_N)} |x_i^N - \xi|^{ps} \, \mathbb{P}_X(d\xi) \right)^{1/p} \left(\int_{C_i(\Gamma_N)} \mathbb{P}_X(d\xi) \right)^{1/q} \\
&\leqslant \left(\int_{C_i(\Gamma_N)} |x_i^N - \xi| \, \mathbb{P}_X(d\xi) \right)^s \Big(\mathbb{P}_X\left(C_i(\Gamma_N)\right) \Big)^{1-s} \\
\iff N^{s+1} \int_{C_i(\Gamma_N)} |x_i^N - \xi|^s \, \mathbb{P}_X(d\xi) &\leqslant N^{s+1} \left(\int_{C_i(\Gamma_N)} |x_i^N - \xi| \, \mathbb{P}_X(d\xi) \right)^s \Big(\mathbb{P}_X\left(C_i(\Gamma_N)\right) \Big)^{1-s} \\
&\leqslant \left(N^2 \int_{C_i(\Gamma_N)} |x_i^N - \xi| \, \mathbb{P}_X(d\xi) \right)^s \Big(N \, \mathbb{P}_X\left(C_i(\Gamma_N)\right) \Big)^{1-s}.
\end{aligned}$$

And using the result proved above for $s = 1$ and (1.1), we obtain the desired result

$$\begin{aligned}
&\limsup_N N^{s+1} \int_{C_i(\Gamma_N)} |x_i^N - \xi|^s \, \mathbb{P}_X(d\xi) \\
&\qquad \leqslant \limsup_N \left(N^2 \int_{C_i(\Gamma_N)} |x_i^N - \xi| \, \mathbb{P}_X(d\xi) \right)^s \Big(N \, \mathbb{P}_X\left(C_i(\Gamma_N)\right) \Big)^{1-s} \\
&\qquad \leqslant \left(\frac{1}{12} \|\varphi\|_{1/3} \right)^{\frac{s}{2}} \left(c_{\varphi,1/3} \|\varphi^{2/3}\|_\infty \right)^{1-\frac{s}{2}} \\
&\qquad < +\infty.
\end{aligned}$$

□

The following result will be useful in the last part of the paper, which is the Theorem 6 in [DGLP04].

Theorem 1.13. *Let $(\Gamma_N)_{N \geqslant 1}$ a sequence of optimal quantizers for $\mathbb{P}_X$. Then*

$$\lim_{N \to +\infty} N^2 \, \mathbb{E}\left[g(\widehat{X}^N) |X - \widehat{X}^N|^2 \right] = \mathcal{Q}_2(\mathbb{P}_X) \int g(\xi) \, \mathbb{P}_X(d\xi)$$

for every function $g : \mathbb{R} \to \mathbb{R}$ such that $\mathbb{E}\left[g(X)\right] < +\infty$, with $\mathcal{Q}_2(\mathbb{P}_X)$ the Zador's constant.

The last result we state is an answer to the following question: what can we say about the rate of convergence of $\mathbb{E}\left[|X - \widehat{X}^N|^{2+\beta}\right]$ knowing that $\widehat{X}^N$ is a quadratic optimal quantization? This problem is known as the distortion mismatch problem and has been first addressed in [GLP08] and the results have been extended in Theorem 4.3 of [PS18].

Theorem 1.14. *[L^r-L^s-distortion mismatch] Let $X : (\Omega, \mathcal{A}, \mathbb{P}) \to \mathbb{R}$ be a random variable and let $r \in (0, +\infty)$. Assume that the distribution $\mathbb{P}_X$ of X has a non-zero absolutely continuous component with density φ, i.e. $\mathbb{P}_X(d\xi) = \varphi(\xi) \cdot \lambda(d\xi) + \nu(d\xi)$, where $\nu \perp \lambda$ is the singular component of $\mathbb{P}_X$ with respect to the Lebesgue measure λ on $\mathbb{R}$ and φ is non-identically null. Let $(\Gamma_N)_{N \geqslant 1}$ be a sequence of L^r-optimal grids. Let $s \in (r, r+1)$. If*

$$X \in L^{\frac{s}{1+r-s}+\delta}(\mathbb{P})$$

for some $\delta > 0$, then

$$\limsup_N N \|X - \widehat{X}^N\|_s < +\infty.$$

2 Weak Error bounds for Optimal Quantization ($d = 1$)

Let $X \in L^2(\mathbb{P})$ and $\widehat{X}^N$ a quadratic optimal quantizer of X which takes its values in the finite grid $\Gamma_N = \{x_1^N, \ldots, x_N^N\}$ of size N. We consider a function $f : \mathbb{R} \to \mathbb{R}$ with $f(X) \in L^2(\mathbb{P})$. One of the application of the framework developed above is the approximation of expectations of the form $\mathbb{E}\big[f(X)\big]$. Indeed, as $\widehat{X}^N$ is close to X in $L^2(\mathbb{P})$, a natural idea is to replace X by $\widehat{X}^N$ inside the expectation

$$\mathbb{E}\big[f(\widehat{X}^N)\big] = \sum_{i=1}^{N} f(x_i^N)\, \mathbb{P}_X\big(C_i(\Gamma_N)\big).$$

The above formula is referred as the quantization-based cubature formula to approximate $\mathbb{E}\big[f(X)\big]$. Now, we need to have an idea of the error we make when doing such an approximation and what is its rate of convergence as N tends to infinity? For that, we want to find the largest $\alpha \in \mathbb{R}$, such that the beyond limit is bounded

$$\lim_{N \to +\infty} N^{\alpha} \big|\, \mathbb{E}\big[f(X)\big] - \mathbb{E}\big[f(\widehat{X}^N)\big]\big| \leqslant C_{f,X} < +\infty. \tag{2.1}$$

The first class of function we consider is the class of Lipschitz continuous functions, more precisely piecewise affine functions and convex Lipschitz continuous functions. Then we deal with differentiable functions with piecewise-defined derivatives.

2.1 Piecewise affine functions

We improve the standard rate of convergence which is of order 1 for Lipschitz continuous functions by considering a subclass of the Lipschitz continuous functions, namely piecewise affine functions. This new result shows that the weak error induced is of order 2 ($\alpha = 2$ in (2.1)).

Lemma 2.1. *Assume that the distribution $\mathbb{P}_X = \varphi \cdot \lambda$ of X satisfies the conditions of Theorem 1.11. Let $f : \mathbb{R} \to \mathbb{R}$ be a Borel function.*

(a) If f is a continuous piecewise affine function with finitely many breaks of affinity, then there exists a real constant $C_{f,X} > 0$ such that

$$\limsup_N N^2 \big|\, \mathbb{E}\big[f(X)\big] - \mathbb{E}\big[f(\widehat{X}^N)\big]\big| \leqslant C_{f,X} < +\infty.$$

(b) However, if f is not supposed continuous but is still a piecewise affine function with finitely many breaks of affinity, then there exists a real constant $C_{f,X} > 0$ such that

$$\limsup_N N \big|\, \mathbb{E}\big[f(X)\big] - \mathbb{E}\big[f(\widehat{X}^N)\big]\big| \leqslant C_{f,X} < +\infty.$$

Proof. Let I be a compact interval containing all the affinity breaks of f denoted $a_1, \ldots, a_\ell$.

(a) Let f supposed to be continuous. Note that f is Lipschitz continuous (with coefficient denoted $[f]_{Lip} := \max_{i=1,\ldots,\ell} |a_i|$). Let $\Gamma_N = \{x_1^N, \ldots, x_N^N\}$ be an L^2- optimal quantizer at level $N \geqslant 1$.

$$\begin{aligned}
\mathbb{E}\big[f(X)\big] - \mathbb{E}\big[f(\widehat{X}^N)\big] &= \sum_{i=1}^{N} \int_{C_i(\Gamma_N)} \big(f(\xi) - f(x_i^N)\big)\, \mathbb{P}_X(d\xi) \\
&= \sum_{i \in J_f^N} \int_{C_i(\Gamma_N)} \big(f(\xi) - f(x_i^N)\big)\, \mathbb{P}_X(d\xi)
\end{aligned} \tag{2.2}$$

where $J_f^N = \{i : C_i(\Gamma_N) \text{ contains an affinity break}\}$ since all other terms are 0. Indeed, as $f(\xi) = \alpha_i\xi + \beta_i$ on $C_i(\Gamma_N)$ and using Corollary 1.8

$$\int_{C_i(\Gamma_N)} \big(f(\xi) - f(x_i^N)\big)\,\mathbb{P}_X(d\xi) = \alpha_i\,\mathbb{E}\big[(X - \widehat{X}^N)\,\mathbb{1}_{\{\widehat{X}^N = x_i^N\}}\big] = 0.$$

Now, taking the absolute value in (2.2), we have

$$\begin{aligned}\big|\,\mathbb{E}\big[f(X)\big] - \mathbb{E}\big[f(\widehat{X}^N)\big]\big| &\leqslant \mathrm{card}(J_f^N)\max_{i\in J_f^N}\int_{C_i(\Gamma_N)} |f(\xi) - f(x_i^N)|\,\mathbb{P}_X(d\xi)\\ &\leqslant \mathrm{card}(J_f^N)[f]_{Lip}\max_{i\in J_f^N}\int_{C_i(\Gamma_N)} |\xi - x_i^N|\,\mathbb{P}_X(d\xi) \qquad (2.3)\end{aligned}$$

and using Corollary 1.12 with $s = 1$, we have the desired result, with an explicit asymptotic upper bound,

$$\begin{aligned}\limsup_N N^2\big|\,\mathbb{E}\big[f(X)\big] - \mathbb{E}\big[f(\widehat{X}^N)\big]\big| &\leqslant [f]_{Lip}\lim_N \mathrm{card}(J_f^N)\max_{i\in J_f^N} N^2\int_{C_i(\Gamma_N)} |\xi - x_i^N|\,\mathbb{P}_X(d\xi)\\ &< [f]_{Lip}\frac{\ell}{2\sqrt{3}}\big(c_{\varphi,1/3}\|\varphi\|_{1/3}\|\varphi^{1/3}\|_\infty\big)^{\frac{1}{2}}\\ &< +\infty.\end{aligned}$$

(b) The sum in (2.2) in the discontinuous case is still true. However, the bound in (2.3) changes and becomes

$$\big|\,\mathbb{E}\big[f(X)\big] - \mathbb{E}\big[f(\widehat{X}^N)\big]\big| \leqslant 2\ell\|f\|_{\infty,K_0}\max_{i\in J_f^N}\mathbb{P}_X\big(C_i(\Gamma_N)\big)$$

where $\|f\|_{\infty,K_0}$ denotes the maximum of $|f|$ on K_0 and K_0 is defined as the compact appearing in Lemma 1.10 stating that the union over all N of all the cells where their intersection with the interval $[a_1, a_\ell]$ is non empty lies in a compact K_0, namely

$$\bigcup_N \bigcup_{C_i(\Gamma_N)\cap[a_1,a_\ell]\neq\varnothing} C_i(\Gamma_N) \subset K_0.$$

The desired limit is obtained using Theorem 1.11.

□

2.2 Lipschitz Convex functions

Thanks to the previous result on piecewise-affine functions, we can extend the rate of convergence of order 2 to a bigger class of functions: Lipschitz convex functions.

We recall that a real-valued function f defined on a non-trivial interval $I \subset \mathbb{R}$ is convex if

$$f\big(tx + (1-t)y\big) \leqslant tf(x) + (1-t)f(y),$$

for every $t \in [0,1]$ and $x, y \in I$. If $f : I \to \mathbb{R}$ is supposed to be a convex function, then its right and left derivatives exist, are non-decreasing on $\mathring{I}$ and $\forall x \in \mathring{I}$, $f'_-(x) \leqslant f'_+(x)$. Moreover, as f is supposed to be Lipschitz continuous, then f'_- and f'_+ are bounded on I by $[f]_{Lip}$.

Remark 2.2. One of the very interesting properties of convex functions when dealing with stationary quantizers follows from Jensen's inequality. Indeed, for every convex function $f : I \to \mathbb{R}$ such that $f(X) \in L^1(\mathbb{P})$,

$$\mathbb{E}\Big[f\big(\,\mathbb{E}\big[X \mid \widehat{X}^N\big]\big)\Big] \leqslant \mathbb{E}\Big[\,\mathbb{E}\big[f(X) \mid \widehat{X}^N\big]\Big]$$

so that,

$$\mathbb{E}\big[f(\widehat{X}^N)\big] \leqslant \mathbb{E}\big[f(X)\big].$$

This means that the quantization-based cubature formula used to approximate $\mathbb{E}\big[f(X)\big]$ is a lower-bound of the expectation.

We present, here, a more convenient and general form of the well known Carr-Madan formula representation (see [CM01]).

Proposition 2.3. *Let $f : I \to \mathbb{R}$ be a Lipschitz convex function and let I be any interval non trivial ($\neq \varnothing, \{a\}$) with endpoints $a, b \in \overline{\mathbb{R}}$. Then, there exists a unique finite non-negative Borel measure $\nu := \nu_f$ on I such that, for every $c \in I$,*

$$\forall x \in I, \quad f(x) = f(c) + (x-c)f'_+(c) + \int_{[a,c]\cap I} (u-x)_+ \nu(du) + \int_{(c,b]\cap I} (x-u)_+ \nu(du).$$

Proof. Let $f : I \to \mathbb{R}$ be a Lipschitz convex function. We can define the non-negative finite measure $\nu := \nu_f$ on I by setting

$$\forall x, y \in I, \quad x \leqslant y, \quad \nu\big((x,y]\big) = f'_+(y) - f'_+(x).$$

The finiteness of ν is induced by the Lipschitz continuity of f as the left and right derivatives are bounded by $[f]_{Lip} = \max(\|f'_+\|_\infty, \|f'_-\|_\infty)$. Let $c \in I$, for every $x \geqslant c$, we have the following representation of $f(x)$:

$$\begin{aligned}
f(x) &= f(c) + \int_c^x f'_+(u)du \\
&= f(c) + xf'_+(c) + \int_c^x \nu((c,u])du \\
&= f(c) + xf'_+(c) + \int\int \mathbb{1}_{(c,x]}(u)\,\mathbb{1}_{(c,u]}(v)\nu(dv)\,du \\
&= f(c) + xf'_+(c) + \int_{(c,x]} (x-v)du\;\nu(dv) \\
&= f(c) + xf'_+(c) + \int_{(c,b]\cap I} (x-v)_+\nu(dv)
\end{aligned}$$

using Fubini's Theorem and noting that $\mathbb{1}_{(c,x]}(u)\,\mathbb{1}_{(c,u]}(v) = \mathbb{1}_{(c,x]}(v)\,\mathbb{1}_{[v,x]}(u)$. Similarly for $x \leqslant c$

$$f(x) = f(c) + xf'_+(c) + \int_{[a,c]\cap I} (u-x)_+\nu(du).$$

Then,

$$\forall x \in \mathbb{R}, \quad f(x) = f(c) + xf'_+(c) + \int_{[a,c]\cap I} (u-x)_+\nu(du) + \int_{(c,b]\cap I} (x-u)_+\nu(du).$$

□

We can now use the representation of convex functions given above and extend the result concerning the weak error of order 2 ($\alpha = 2$ in (2.1)).

Proposition 2.4. *We assume that the distribution $\mathbb{P}_X = \varphi \cdot \lambda$ of X satisfies the conditions of Theorem 1.11. Let I be any non-trivial interval and let $f : I \to \mathbb{R}$ be a Lipschitz convex function with second derivative ν (see Proposition 2.3). If $I_{\mathbb{P}_X} \cap \mathrm{supp}(\nu)$ is compact, with $I_{\mathbb{P}_X} := \mathrm{supp}(\mathbb{P}_X)$, then there exists a real constant $C_{f,X} > 0$ such that*

$$\limsup_N N^2 \big|\,\mathbb{E}\big[f(X)\big] - \mathbb{E}\big[f(\widehat{X}^N)\big]\big| \leqslant C_{f,X} < +\infty.$$

Remark 2.5. Assuming that $\mathrm{supp}(\nu)$ is compact actually means that f is affine outside a compact set, namely that there exist $\alpha^{(\pm)}$ and $\beta^{(\pm)}$ such that $f(x) = \alpha^{(+)}x + \beta^{(+)}$, for x large enough ($x \geqslant K_+$) and $f(x) = \alpha^{(-)}x + \beta^{(-)}$, for x small enough ($x \leqslant K_-$). Therefore, this class of functions contains all classical vanilla financial payoffs: call, put, butterfly, saddle, straddle, spread, etc. Moreover, if $I_{\mathbb{P}_X}$ is compact, such as in the uniform distribution, then there is no need for the hypothesis on ν and we could consider any Lipschitz convex functions we want. The hypothesis on the intersection allows us to consider more cases.

Proof. First we decompose the expectations across the Voronoï cells as follows

$$\begin{aligned}\mathbb{E}\left[f(X) - f(\widehat{X}^N)\right] &= \sum_{i=1}^N \mathbb{E}\left[\left(f(X) - f(\widehat{X}^N)\right)\mathbb{1}_{\{X\in C_i(\Gamma_N)\}}\right] \\ &= \sum_{i=1}^N \mathbb{E}\left[\left(f(X) - f(x_i^N)\right)\mathbb{1}_{\{X\in(x_{i-1/2}^N, x_{i+1/2}^N]\}}\right].\end{aligned}$$

We use the integral representation of the convex function f, of the Proposition 2.3, with $x := X$ and $c := x_i$ and with the stationarity conditional property given by Corollary 1.8, the first term cancels out, for every i,

$$\mathbb{E}\left[(X - x_i^N)f'_+(x_i^N)\,\mathbb{1}_{\{X\in C_i(\Gamma_N)\}}\right] = 0.$$

Hence, we obtain

$$\begin{aligned}&\mathbb{E}\Big[\left(f(X) - f(x_i^N)\right)\mathbb{1}_{\{X\in(x_{i-1/2}^N, x_{i+1/2}^N]\}}\Big] \\ &= \mathbb{E}\left[\left(\int_{[a,x_i^N]\cap I}(u-X)_+\nu(du) + \int_{(x_i^N,b]\cap I}(X-u)_+\nu(du)\right)\mathbb{1}_{\{X\in(x_{i-1/2}^N, x_{i+1/2}^N]\}}\right] \\ &= \mathbb{E}\left[\int_{(x_{i-1/2}^N, x_i^N]}(u-X)_+\nu(du)\,\mathbb{1}_{\{X\in(x_{i-1/2}^N, x_i^N]\}}\right] \\ &\qquad + \mathbb{E}\left[\int_{(x_i^N, x_{i+1/2}^N)}(X-u)_+\nu(du)\,\mathbb{1}_{\{X\in[x_i^N, x_{i+1/2}^N]\}}\right].\end{aligned} \tag{2.4}$$

The interval $(x_{i-1/2}^N, x_i^N]$ in the integral is left-open because when $u = x_{i-1/2}^N$, as $X \in (x_{i-1/2}^N, x_i^N]$, $(u - X)_+ = 0$. The same remark can be made concerning the right open-bound of the interval $(x_i^N, x_{i+1/2}^N)$ in the integral. Now, using a crude upper-bound for (2.4), we get

$$\begin{aligned}\mathbb{E}\left[\left(f(X) - f(x_i^N)\right)\mathbb{1}_{\{X\in(x_{i-1/2}^N, x_{i+1/2}^N]\}}\right] &\leqslant \mathbb{E}\left[(x_i^N - X)\nu\big((x_{i-1/2}^N, x_i^N]\big)\,\mathbb{1}_{\{X\in(x_{i-1/2}^N, x_i^N]\}}\right] \\ &\quad + \mathbb{E}\left[(X - x_i^N)\nu\big((x_i^N, x_{i+1/2}^N)\big)\,\mathbb{1}_{\{X\in[x_i^N, x_{i+1/2}^N]\}}\right] \\ &\leqslant \mathbb{E}\left[|x_i^N - X|\,\mathbb{1}_{\{X\in C_i(\Gamma_N)\}}\right]\nu\big(C_i(\Gamma_N)\big)\end{aligned}$$

as $\nu\big((x_{i-1/2}^N, x_{i+1/2}^N)\big) \leqslant \nu\big(C_i(\Gamma_N)\big)$. Hence

$$\begin{aligned}0 \leqslant \mathbb{E}\left[f(X) - f(\widehat{X}^N)\right] &\leqslant \sum_{i=1}^N \mathbb{E}\left[|x_i^N - X|\,\mathbb{1}_{\{X\in C_i(\Gamma_N)\}}\right]\nu\big(C_i(\Gamma_N)\big) \\ &\leqslant \sum_{i=1}^N \mathbb{E}\left[|x_i^N - X|\,\mathbb{1}_{\{X\in C_i(\Gamma_N)\}}\right]\mathbb{1}_{\{x_i^N\in J_\nu\}}\,\nu\big(C_i(\Gamma_N)\big)\end{aligned}$$

with $J_\nu := [\inf_N x_{i_a-1/2}^N, \sup_N x_{i_b+1/2}^N]$ where $x_{i_a}^N$ and $x_{i_b}^N$ are the centroids of the optimal quantizer of size N that contains, respectively, the infimum and the supremum of the support of ν,

denoted by a and b, respectively. Hence, $x^N_{i_a-1/2}$ is the lower bound of the Voronoï cell $C_{i_a}(\Gamma_N)$ associated to the centroid $x^N_{i_a}$ and $x^N_{i_b+1/2}$ is the upper bound of the Voronoï cell $C_{i_b}(\Gamma_N)$ associated to the centroid $x^N_{i_b}$. If a is not contained in $I_{\mathbb{P}_X}$, then the lower bound of J_ν is set to a, and the same hold for b: if it is not contained in $I_{\mathbb{P}_X}$, the upper bound of J_ν is set to b. Then,

$$\begin{aligned} N^2\,\mathbb{E}\big[f(X)-f(\widehat{X}^N)\big] &\leqslant N^2\sum_{i=1}^N \mathbb{E}\big[|x_i^N-X|\,\mathbb{1}_{\{X\in C_i(\Gamma_N)\}}\big]\,\mathbb{1}_{\{x_i^N\in J_\nu\}}\,\nu(C_i(\Gamma_N)) \\ &\leqslant N^2 \sup_{i:x_i^N\in I_{\mathbb{P}_X}\cap J_\nu} \mathbb{E}\big[|\widehat{X}^N-X|\,\mathbb{1}_{\{X\in C_i(\Gamma_N)\}}\big]\sum_{i=1}^N \nu(C_i(\Gamma_N)) \\ &\leqslant \nu(I_{\mathbb{P}_X})N^2 \sup_{i:x_i^N\in I_{\mathbb{P}_X}\cap J_\nu} \mathbb{E}\big[|\widehat{X}^N-X|\,\mathbb{1}_{\{X\in C_i(\Gamma_N)\}}\big] \end{aligned}$$

yielding the desired result with Theorem 1.11 if $I_{\mathbb{P}_X}\cap J_\nu$ is compact.
Under the hypothesis $I_{\mathbb{P}_X}\cap\ \mathrm{supp}(\nu)$ compact, then by Lemma 1.10,

$$\bigcup_N \bigcup_{x_i^N\in I_{\mathbb{P}_X}\cap\ \mathrm{supp}(\nu)} C_i(\Gamma_N)\subset \bigcup_N \bigcup_{C_i(\Gamma_N)\cap I_{\mathbb{P}_X}\cap\ \mathrm{supp}(\nu)\neq\varnothing} C_i(\Gamma_N)\subset K_0,$$

with $K_0 := I_{\mathbb{P}_X}\cap J_\nu$ compact, which is what we were looking for. □

Proposition 2.6. *Assume that the distribution $\mathbb{P}_X=\varphi\cdot\lambda$ of X satisfies the conditions of Theorem 1.11 not only on compact sets but uniformly. Let I be any non-trivial interval then for every function $f: I\to\mathbb{R}$ Lipschitz convex with second derivative ν defined as in Proposition 2.3, there exists a real constant $C_{f,X}>0$ such that*

$$\limsup_N N^2\big|\,\mathbb{E}\big[f(X)\big]-\mathbb{E}\big[f(\widehat{X}^N)\big]\big|\leqslant C_{f,X}<+\infty.$$

Proof. This proof is exactly the same as above the Proposition. □

Remark 2.7. It has not be shown yet that Gaussian or Exponential random variables satisfy the conditions of Theorem 1.11 uniformly but empirical tests tend to confirm that they exhibit the error bound property for Lipschitz convex functions. More details are given in the numerical part.

2.3 Differentiable functions

In the following proposition, we deal with functions that are piecewise-defined and where their piecewise-defined derivatives are supposed to be locally-Lipschitz continuous or locally α-Hölder continuous on the non-bounded parts of the interval. We define below what we mean by locally-Lipschitz and locally α-Hölder.

Definition 2.8.

- A function $f: I\to\mathbb{R}$ is supposed to be locally-Lipschitz continuous, if

$$\forall x,y\in I\quad |f(x)-f(y)|\leqslant [f]_{Lip,loc}|x-y|\big(g(x)+g(y)\big)$$

  where $[f]_{Lip,loc}$ is a real constant and $g:\mathbb{R}\to\mathbb{R}_+$.

- A function $f: I\to\mathbb{R}$ is supposed to be locally α-Hölder continuous, if

$$\forall x,y\in I\quad |f(x)-f(y)|\leqslant [f]_{\alpha,loc}|x-y|^\alpha\big(g(x)+g(y)\big)$$

  where $[f]_{\alpha,loc}$ is a real constant and $g:\mathbb{R}\to\mathbb{R}_+$.

Proposition 2.9. *Assume that the distribution $\mathbb{P}_X$ of X satisfies the conditions of the L^r-L^s-distortion mismatch Theorem 1.14 and Theorem 1.11 concerning the local behaviours of optimal quantizers. If $f : \mathbb{R} \to \mathbb{R}$ is a piecewise-defined continuous function with finitely many breaks of affinity $\{a_1, \ldots, a_K\}$, where $-\infty = a_0 < a_1 < \cdots < a_K < a_{K+1} = +\infty$, such that the piecewise-defined derivatives denoted $(f'_k)_{k=0,\ldots,d}$ are either*

(a) locally-Lipschitz continuous on (a_k, a_{k+1}) where $\exists\, q_k > 3$ such that the q_k-th power of $g_k : (a_k, a_{k+1}) \to \mathbb{R}_+$ defined in Definition 2.8 are convex and $\big(\|g_k(X)\|_{q_k}\big)_{k=1,\ldots,K} < +\infty$. Then there exists a real constant $C_{f,X} > 0$ such that

$$\limsup_N N^2 \big|\, \mathbb{E}\big[f(X)\big] - \mathbb{E}\big[f(\widehat{X}^N)\big]\big| \leqslant C_{f,X} < +\infty.$$

(b) or locally α-Hölder continuous on (a_k, a_{k+1}), $\alpha \in (0,1)$, where $\exists\, q_k > \frac{3}{2-\alpha}$ such that the q_k-th power of $g_k : (a_k, a_{k+1}) \to \mathbb{R}_+$ defined in Definition 2.8 are convex and $\big(\|g_k(X)\|_{q_k}\big)_{k=1,\ldots,K} < +\infty$. Then there exists a real constant $C_{f,X} > 0$ such that

$$\limsup_N N^{1+\alpha} \big|\, \mathbb{E}\big[f(X)\big] - \mathbb{E}\big[f(\widehat{X}^N)\big]\big| \leqslant C_{f,X} < +\infty.$$

Proof. (a) Let $\Gamma_N = \{x_1^N, \ldots, x_N^N\}$ be a L^2- optimal quantizer at level $N \geqslant 1$. In the first place, we define the set of all the indexes of the Voronoï cells that contains a break of affinity

$$I_{reg}^N = \big\{i = 1, \ldots, N : C_i(\Gamma_N) \cap [a_1, a_K] \neq \varnothing\big\}.$$

Hence,

$$\mathbb{E}\big[f(\widehat{X}^N)\big] - \mathbb{E}\big[f(X)\big] = \underbrace{\sum_{i \in I_{reg}^N} \int_{C_i(\Gamma_N)} \big(f(x_i^N) - f(\xi)\big)\, \mathbb{P}_X(d\xi)}_{(A)} + \underbrace{\sum_{i \notin I_{reg}^N} \int_{C_i(\Gamma_N)} \big(f(x_i^N) - f(\xi)\big)\, \mathbb{P}_X(d\xi)}_{(B)}.$$

First, we deal with the (B) term. As, $i \notin I_{reg}^N$, f is differentiable in $C_i(\Gamma_N)$ and admits a first-order Taylor expansion at the point x_i^N, moreover by Corollary 1.8, $\int_{C_i(\Gamma_N)} f'(x_i^N)(\xi - x_i^N)\, \mathbb{P}_X(d\xi) = 0$, hence

$$\int_{C_i(\Gamma_N)} \big(f(x_i^N) - f(\xi)\big)\, \mathbb{P}_X(d\xi) = \int_{C_i(\Gamma_N)} \int_0^1 \big(f'(x_i^N) - f'(tx_i^N + (1-t)\xi)\big)(x_i^N - \xi) dt\, \mathbb{P}_X(d\xi).$$

Now, we take the absolute value and we use the locally Lipschitz property of the derivative, yielding

$$\begin{aligned}
&\left| \int_{C_i(\Gamma_N)} \big(f(x_i^N) - f(\xi)\big)\, \mathbb{P}_X(d\xi) \right| \\
&\qquad \leqslant \int_{C_i(\Gamma_N)} \int_0^1 |f'(x_i^N) - f'(tx_i^N + (1-t)\xi)||x_i^N - \xi| dt\, \mathbb{P}_X(d\xi) \\
&\qquad \leqslant [f']_{k,Lip,loc} \int_{C_i(\Gamma_N)} \int_0^1 (1-t)|x_i^N - \xi|^2 \big(g_{k_i}(x_i^N) + g_{k_i}(tx_i^N + (1-t)\xi)\big) dt\, \mathbb{P}_X(d\xi),
\end{aligned} \tag{2.5}$$

with $k_i := \{k = 0, \dots, d : x_i \in (a_k, a_{k+1})\}$. Under the convex hypothesis of $g_{k_i}^{q_{k_i}}$, we have that

$$g_{k_i}(tx_i^N + (1-t)\xi) \leqslant \max\big(g_{k_i}(x_i^N), g_{k_i}(\xi)\big) \leqslant g_{k_i}(x_i^N) + g_{k_i}(\xi),$$

thus

$$\int_{C_i(\Gamma_N)} \int_0^1 (1-t)|x_i^N - \xi|^2 \big(g_k(x_i^N) + g_k(tx_i^N + (1-t)\xi)\big) dt\, \mathbb{P}_X(d\xi) \\ \leqslant \frac{1}{2} \int_{C_i(\Gamma_N)} |x_i^N - \xi|^2 \big(2g_k(x_i^N) + g_k(\xi)\big)\, \mathbb{P}_X(d\xi).$$

Now, taking the sum over all $i \notin I_{reg}^N$ and denoting $[f']_{Lip,loc} := \max_k [f']_{k,Lip,loc}$

$$\begin{aligned} \big|(B)\big| &\leqslant \frac{1}{2}[f']_{Lip,loc} \sum_{i \notin I_{reg}^N} \int_{C_i(\Gamma_N)} |x_i^N - \xi|^2 \big(2g_{k_i}(x_i^N) + g_{k_i}(\xi)\big)\, \mathbb{P}_X(d\xi) \\ &\leqslant \frac{K}{2}[f']_{Lip,loc} \max_k \mathbb{E}\Big[|\widehat{X}^N - X|^2 \big(2g_k(\widehat{X}^N) + g_k(X)\big)\Big] \\ &\leqslant \frac{K}{2}[f']_{Lip,loc} \max_k \|\widehat{X}^N - X\|_{2p_k}^2 \big(2\|g_k(\widehat{X}^N)\|_{q_k} + \|g_k(X)\|_{q_k}\big) \\ &\leqslant \frac{K}{2}[f']_{Lip,loc} \|\widehat{X}^N - X\|_{2p}^2 \max_k \big(2\|g_k(\widehat{X}^N)\|_{q_k} + \|g_k(X)\|_{q_k}\big) \\ &\leqslant \frac{3K}{2}[f']_{Lip,loc} \|\widehat{X}^N - X\|_{2p}^2 \max_k \|g_k(X)\|_{q_k} \end{aligned} \tag{2.6}$$

using Hölder inequality, such that $\frac{1}{p_k} + \frac{1}{q_k} < 1$ and the convexity of g^{q_k}. Under the hypothesis $q_k > 3$, p_k has to be in contained in the interval $(1, 3/2)$, hence p is defined as $p := \max_k p_k$ and using the non-decreasing property of the L^p norm, we obtain the fourth inequality in (2.6). Now, if we use the L^r-L^s-distortion mismatch Theorem 1.14 with $r = 2$ and $s = 2p < 3$ under the condition $X \in L^{\frac{2p}{3-2p}+\delta}(\mathbb{P})$, we have

$$\begin{aligned} N^2\big|(B)\big| &\leqslant N^2 \frac{3K}{2}[f']_{Lip,loc} \|\widehat{X}^N - X\|_{2p}^2 \max_k \|g_k(X)\|_{q_k} \\ &\xrightarrow{N \to +\infty} C_2 < +\infty. \end{aligned} \tag{2.7}$$

Secondly, we take care of the (A) term. Using Lemma 1.10 stating that the union over all N of all the cells where their intersection with the interval $[a_1, a_K]$ is non empty lies in a compact K_0, namely

$$\bigcup_N \bigcup_{C_i(\Gamma_N) \cap [a_1, a_K] \neq \varnothing} C_i(\Gamma_N) \subset K_0$$

and using that f' is bounded on K_0 by $[f']_{Lip,K_0}$, we can use the following integral representation of f

$$f(x) = \int_0^x f'(u)du + f(0)$$

and the stationarity property of the optimal quantizer on $C_i(\Gamma_N)$, yielding

$$\begin{aligned} \Big| \int_{C_i(\Gamma_N)} \big(f(x_i^N) - f(\xi)\big)\, \mathbb{P}_X(d\xi) \Big| &= \Big| \int_{C_i(\Gamma_N)} \int_\xi^{x_i^N} f'(u)du\, \mathbb{P}_X(d\xi) \Big| \\ &\leqslant [f']_{Lip,K_0} \int_{C_i(\Gamma_N)} |\xi - x_i^N|\, \mathbb{P}_X(d\xi). \end{aligned}$$

Now, we sum among all $i \in I_{reg}^N$

$$\big|(A)\big| \leqslant [f']_{Lip,K_0} \sum_{i\in I_{reg}^N} \int_{C_i(\Gamma_N)} |\xi - x_i^N| \, \mathbb{P}_X(d\xi).$$

Hence, using the result concerning the local behaviour of optimal quantizers Corollary 1.12 as $[a_1, a_K]$ is compact, we have

$$\begin{aligned} N^2\big|(A)\big| &\leqslant N^2[f']_{Lip,K_0} \sum_{i\in I_{reg}^N} \int_{C_i(\Gamma_N)} |\xi - x_i^N| \, \mathbb{P}_X(d\xi) \\ &\leqslant N^2 K[f']_{Lip,K_0} \sup_{i:x_i^N\in K_0} \int_{C_i(\Gamma_N)} |\xi - x_i^N| \, \mathbb{P}_X(d\xi) \\ &\xrightarrow{N\to+\infty} C_1 < +\infty. \end{aligned} \tag{2.8}$$

Finally, using (2.8) and (2.7), we have the desired result

$$N^2\big|\,\mathbb{E}\,\big[f(X)\big] - \mathbb{E}\,\big[f(\widehat{X}^N)\big]\big| \leqslant N^2\Big(\big|(A)\big| + \big|(B)\big|\Big) \xrightarrow{N\to+\infty} C_1 + C_2 < +\infty.$$

(b) When the piecewise-defined derivatives are locally α-Hölder continuous on $(-\infty, a_1]$ and $[a_K, +\infty)$, $\alpha \in (0,1)$, the proof is very close to the locally Lipschitz case. Indeed, the first difference is in (2.5), where the $|x_i^N - \xi|^2$ is replaced by $|x_i^N - \xi|^{1+\alpha}$ and the constant is the one of the locally α-Hölder hypothesis. This implies that (2.6) is replaced by

$$|(B)| \leqslant \frac{3K[f']_{Hol,loc}}{2} \|\widehat{X}^N - X\|_{(1+\alpha)p}^{1+\alpha} \max_k \|g_k(X)\|_{q_k}.$$

Finally, using the L^r-L^s-distortion mismatch Theorem 1.14 with $r = 2$ and $s = (1+\alpha)p < 3$ under the condition $X \in L^{\frac{(1+\alpha)p}{3-(1+\alpha)p}+\delta}(\mathbb{P})$, we have

$$\begin{aligned} N^{1+\alpha}|(B)| &\leqslant N^{1+\alpha} \frac{3K[f']_{Hol,loc}}{2} \|\widehat{X}^N - X\|_{(1+\alpha)p}^{1+\alpha} \max_k \|g_k(X)\|_{q_k} \\ &\xrightarrow{N\to+\infty} C_3 < +\infty. \end{aligned}$$

The other parts of the proof are identical, yielding the desired result.

□

Remark 2.10. If one strengthens the hypothesis concerning the piecewise locally Lipschitz continuous derivative and considers in place that the derivative is piecewise Lipschitz continuous, then the hypothesis that X should satisfy the conditions of Theorem 1.14 can be relaxed. Indeed, the term $\frac{3K}{2}[f']_{Lip,loc}\|\widehat{X}^N - X\|_{2p}^2 \max_k \|g_k(X)\|_{q_k}$ in (2.6) would become $\frac{1}{2}[f']_{Lip}\|\widehat{X}^N - X\|_2^2$ and we would conclude using Zador's Theorem 1.9.

3 Weak Error and Richardson-Romberg Extrapolation

One can improve the previous speeds of convergence using Richardson-Romberg extrapolation method. The Richardson extrapolation is a method that was originally introduced in numerical analysis by Richardson in 1911 (see [RG10]) and developed later by Romberg in 1955 (see [Rom55]) whose aim was to speed-up the rate of convergence of a sequence, to accelerate the research of a solution of an ODE's or to approximate more precisely integrals.

[TT90] and [Pag07, Pag18] used this concept for the computation of the expectation $\mathbb{E}\big[f(X_T)\big]$ of a diffusion $(X_t)_{t\in[0,T]}$ that cannot be simulated exactly at a given time T but can be approximated by a simulable process $\widetilde{X}_T^{(h)}$ using a Euler scheme with time step $h = T/n$ and n the number of time step. The main idea is to use the weak error expansion of the approximation in order to highlight the term we would *kill*. For example, using the following weak time discretization error of order 1

$$\mathbb{E}\big[f(X_T)\big] = \mathbb{E}\big[f(\widetilde{X}_T^{(h)})\big] + \frac{c_1}{n} + O(n^{-2}),$$

one reduces the error of the approximation using a linear combination of the approximating process $\widetilde{X}_T^{(h)}$ and a refiner process $\widetilde{X}_T^{(h/2)}$, namely

$$\mathbb{E}\big[f(X_T)\big] = \mathbb{E}\big[2f(\widetilde{X}_T^{(h/2)}) - f(\widetilde{X}_T^{(h)})\big] - \frac{1}{2}\frac{c_2}{n^2} + O(n^{-2}).$$

Our goal within the optimal quantization framework is to improve the speed of convergence of the cubature formula using the same ideas. Let us consider a random variable $X : (\Omega, \mathcal{A}, \mathbb{P}) \to \mathbb{R}$ and a quadratic-optimal quantizer $\widehat{X}^N$ of X. In our case we show that, if we are in dimension one there exists, for some functions f, a *weak error expansion* of the form:

$$\mathbb{E}\big[f(X)\big] = \mathbb{E}\big[f(\widehat{X}^N)\big] + \frac{c_2}{N^2} + O(N^{-(2+\beta)})$$

with $\beta \in (0, 1)$. We present in Section 3.2 a similar result in higher dimension.

3.1 In dimension one

This first result is focused on function $f : \mathbb{R} \to \mathbb{R}$ with Lipschitz continuous second derivative. In that case, we have a *weak error quantization* of order two. The first term of the expansion is equal to zero, thanks to the stationarity of the quadratic optimal quantizer.

Proposition 3.1. *Let $f : \mathbb{R} \to \mathbb{R}$ be a twice differentiable function with Lipschitz continuous second derivative. Let $X : (\Omega, \mathcal{A}, \mathbb{P}) \to \mathbb{R}$ be a random variable and the distribution of $\mathbb{P}_X$ of X has a non-zero absolutely continuous density φ and, for every $N \geqslant 1$, let Γ_N be an optimal quantizer at level $N \geqslant 1$ for X. Then, $\forall\, \beta \in (0, 1)$, we have the following expansion*

$$\mathbb{E}\big[f(X)\big] = \mathbb{E}\big[f(\widehat{X}^N)\big] + \frac{c_2}{N^2} + O(N^{-(2+\beta)}).$$

Moreover, if $\varphi : [a, b] \to \mathbb{R}_+$ is a Lipschitz continuous probability density function, bounded away from 0 on $[a, b]$ then we can choose $\beta = 1$, yielding

$$\mathbb{E}\big[f(X)\big] = \mathbb{E}\big[f(\widehat{X}^N)\big] + \frac{c_2}{N^2} + O(N^{-3}).$$

Proof. If f is twice differentiable with Lipschitz continuous second derivatives, we have the following expansion

$$f(x) = f(y) + f'(y)(x-y) + \frac{1}{2}f''(y)(x-y)^2 + \int_0^1 (1-t)\big(f''(tx + (1-t)y) - f''(y)\big)(x-y)^2 dt$$

hence replacing x and y by X and $\widehat{X}^N$ respectively and taking the expectation yields

$$\mathbb{E}\big[f(X)\big] = \mathbb{E}\big[f(\widehat{X}^N)\big] + \frac{1}{2}\,\mathbb{E}\big[f''(\widehat{X}^N)|X - \widehat{X}^N|^2\big] + R(X, \widehat{X}^N)$$

where $R(X,\widehat{X}) = \int_0^1 (1-t)\,\mathbb{E}\big[\big(f''(tX+(1-t)\widehat{X}) - f''(\widehat{X})\big)|X-\widehat{X}|^2\big]dt$.

First, using Theorem 1.13 with f'', we have the following limit

$$\lim_{N\to+\infty} N^2\,\mathbb{E}\big[f''(\widehat{X}^N)|X-\widehat{X}^N|^2\big] = \mathcal{Q}_2(\mathbb{P}_X)\int f''(\xi)\,\mathbb{P}_X(d\xi),$$

hence

$$\mathbb{E}\big[f(X)\big] = \mathbb{E}\big[f(\widehat{X}^N)\big] + \frac{c_2}{N^2} + R(X,\widehat{X}^N).$$

Now, we look closely at asymptotic behaviour of $R(X,\widehat{X}^N)$. One notices that, if we consider a Lipschitz continuous function $g:\mathbb{R}\to\mathbb{R}$, for any fixed $\alpha\in(0,1)$,

$$\forall\, x,\, y\in\mathbb{R},\quad |g(x)-g(y)| \leqslant 2\|g\|_\infty^\alpha [g]_{Lip}^{1-\alpha}|x-y|^{1-\alpha}.$$

In our case, taking $g\equiv f''$, we have

$$\begin{aligned}
&\mathbb{E}\Big[\big(f''(tX+(1-t)\widehat{X}^N) - f''(\widehat{X}^N)\big)|X-\widehat{X}^N|^2\Big] \\
&\qquad\qquad \leqslant \mathbb{E}\Big[2\|f''\|_\infty^\alpha [f'']_{Lip}^{1-\alpha} t^{1-\alpha}|X-\widehat{X}^N|^{1-\alpha}|X-\widehat{X}^N|^2\Big] \\
&\qquad\qquad \leqslant C_{\beta,f''}t^\beta\,\mathbb{E}\big[|X-\widehat{X}^N|^{2+\beta}\big]
\end{aligned}$$

with $0<\beta<1$ where $\beta = 1-\alpha$, hence

$$R(X,\widehat{X}^N) \leqslant \widetilde{C}_{\beta,f''}\,\mathbb{E}\big[|X-\widehat{X}^N|^{2+\beta}\big],$$

with $\widetilde{C}_{\beta,f''} = C_{\beta,f''}\frac{1}{(2+\beta)(1+\beta)}$. Using now Theorem 1.14 with $r=2$ and $s=2+\beta$, we have the desired result: $\mathbb{E}\big[|X-\widehat{X}^N|^{2+\beta}\big] = O(N^{-(2+\beta)})$ and finally

$$\mathbb{E}\big[f(X)\big] = \mathbb{E}\big[f(\widehat{X}^N)\big] + \frac{c_2}{N^2} + O(N^{-(2+\beta)}),$$

for every $\beta\in(0,1)$. If moreover, the density φ of X is Lipschitz continuous, bounded away from 0 on $[a,b]$ then we can take $\beta=1$.

□

Now, following the Richardson-Romberg idea, we could combine approximations with optimal quantizers $\widehat{X}^N$ of size N and $\widehat{X}^{\widetilde{N}}$ of size $\widetilde{N}$, with $\widetilde{N}>N$ in order to *kill* the residual term, leading

$$\mathbb{E}\big[f(X)\big] = \mathbb{E}\left[\frac{\widetilde{N}^2 f(\widehat{X}^{\widetilde{N}}) - N^2 f(\widehat{X}^N)}{\widetilde{N}^2 - N^2}\right] + O(N^{-(2+\beta)}). \tag{3.1}$$

Remark 3.2. For the choice of $\widetilde{N}$, we consider $\widetilde{N} := k\times N$. A natural choice for k could be $k=2$ or $k=\sqrt{2}$ but note that the complexity is proportional to $(k+1)N$. In practice it is therefore preferable to take a small k that does not increase complexity too much. For the numerical example, we choose $\widetilde{N} := k\times N$ with $k=1.2$, this is arbitrary and probably not optimal, however even with this k, we attain a weak error of order 3.

3.2 A first extension in higher dimension

In this part, we give a first result on higher dimension concerning the weak error expansion of $\mathbb{E}\big[f(X)\big]$ when approximated by $\mathbb{E}\big[f(\widehat{X}^N)\big]$. In the next part, we use the following matrix norm: let $M\in\mathbb{R}^{d\times d}$, then $|||M||| := \sup_{u:|u|=1}|u^TMu|$.

Proposition 3.3. *Let $f : \mathbb{R}^d \to \mathbb{R}$ be a twice differentiable function with a bounded and Lipschtiz Hessian H, namely $\forall x, y \in \mathbb{R}^d$, $|||H(x) - H(y)||| \leqslant [H]_{Lip}|x-y|$. Let $X : (\Omega, \mathcal{A}, \mathbb{P}) \to \mathbb{R}^d$ be a random vector with independent components $(X_k)_{k=1,\dots,d}$. For every $(N_k)_{k=1,\dots,d} \geqslant 1$, let $(\widehat{X}_d^{N_d})_{k=1,\dots,d}$ be quadratic optimal quantizers of $(X_k)_{k=1,\dots,d}$ taking values in the grids $(\Gamma_{N_k})_{k=1,\dots,d}$ respectively and we define $\widehat{X}^N$ as the product quantizer X taking values in the finite grid $\Gamma_N := \bigotimes_{k=1,\dots,d} \Gamma_{N_d}$ of size $N := N_1 \times \cdots \times N_d$. Then, we have the following expansion*

$$\mathbb{E}\big[f(X)\big] = \mathbb{E}\big[f(\widehat{X}^N)\big] + \sum_{k=1}^{d} \frac{c_k}{N_k^2} + O\bigg(\Big(\min_{k=1:d} N_k\Big)^{-(2+\beta)}\bigg).$$

Proof. If f is twice differentiable, hence we have the following Taylor's expansion

$$\begin{aligned} f(x) = f(a) + \nabla f(a)(x-a) + \frac{1}{2} H(a) \cdot (x-a)^{\otimes 2} \\ + \int_0^1 (1-t)\big(H(tx + (1-t)a) - H(a)\big) \cdot (x-a)^{\otimes 2} dt \end{aligned}$$

where the notation $f(x,a) \cdot (x-a)^{\otimes 2}$ stands for $(x-a)^T f(x,a)(x-a)$. Replacing x and a by X and $\widehat{X}^N$ respectively and taking the expectation

$$\begin{aligned} \mathbb{E}\big[f(X)\big] = \mathbb{E}\big[f(\widehat{X}^N)\big] + \mathbb{E}\big[\nabla f(\widehat{X}^N)(X - \widehat{X}^N)\big] + \frac{1}{2}\mathbb{E}\big[H(\widehat{X}^N) \cdot (X - \widehat{X}^N)^{\otimes 2}\big] \\ + \int_0^1 (1-t)\,\mathbb{E}\Big[\big(H(tX + (1-t)\widehat{X}^N) - H(\widehat{X}^N)\big) \cdot (X - \widehat{X}^N)^{\otimes 2}\Big] dt. \end{aligned}$$

Noticing that, by Corollary 1.8,

$$\begin{aligned} \mathbb{E}\big[\nabla f(\widehat{X}^N)(X - \widehat{X}^N)\big] &= \sum_{k=1}^{d} \mathbb{E}\left[\frac{\partial f}{\partial x_k}(\widehat{X}^N)(X_k - \widehat{X}_k^{N_k})\right] \\ &= \sum_{k=1}^{d} \mathbb{E}\left[\mathbb{E}\left[\frac{\partial f}{\partial x_k}(\widehat{X}^N)(X_k - \widehat{X}_k^{N_k}) \mid \widehat{X}_{-k}\right]\right] \\ &= 0. \end{aligned}$$

where $\widehat{X}_{-k}$ denotes $(\widehat{X}_1^{N_1}, \dots, \widehat{X}_{k-1}^{N_{k-1}}, \widehat{X}_{k+1}^{N_{k+1}}, \dots, \widehat{X}_d^{N_d})$. Hence

$$\begin{aligned} \mathbb{E}\big[f(X)\big] = \mathbb{E}\big[f(\widehat{X}^N)\big] + \frac{1}{2}\mathbb{E}\big[H(\widehat{X}^N) \cdot (X - \widehat{X}^N)^{\otimes 2}\big] \\ + \int_0^1 (1-t)\,\mathbb{E}\Big[\big(H(tX + (1-t)\widehat{X}^N) - H(\widehat{X}^N)\big) \cdot (X - \widehat{X}^N)^{\otimes 2}\Big] dt \end{aligned} \tag{3.2}$$

and looking at the second term in (3.2)

$$
\begin{aligned}
&\mathbb{E}\big[H(\widehat{X}^N)\cdot(X-\widehat{X}^N)^{\otimes 2}\big]\\
&\quad=\sum_{k=1}^{d}\mathbb{E}\left[\frac{\partial^2 f}{\partial x_k^2}(\widehat{X}^N)|X_k-\widehat{X}_k^{N_k}|^2\right]+2\sum_{k\neq l}\mathbb{E}\left[\frac{\partial^2 f}{\partial x_k\partial x_l}(\widehat{X}^N)(X_k-\widehat{X}_k^{N_k})(X_l-\widehat{X}_l^{N_l})\right]\\
&\quad=\sum_{k=1}^{d}\mathbb{E}\left[\mathbb{E}\left[\frac{\partial^2 f}{\partial x_k^2}(\widehat{X}^N)|X_k-\widehat{X}_k^{N_k}|^2\mid\widehat{X}_{-k}\right]\right]\\
&\qquad\qquad+2\sum_{k\neq l}\mathbb{E}\left[\underbrace{\mathbb{E}\left[\frac{\partial^2 f}{\partial x_k\partial x_l}(\widehat{X}^N)(X_k-\widehat{X}_k^{N_k})\mid X_l\right]}_{=0}(X_l-\widehat{X}_l^{N_l})\right]\\
&\quad=\sum_{k=1}^{d}\mathbb{E}\left[\mathbb{E}\left[\frac{\partial^2 f}{\partial x_k^2}(\widehat{X}^N)|X_k-\widehat{X}_k^{N_k}|^2\mid\widehat{X}_{-k}\right]\right]\\
&\quad=\sum_{k=1}^{d}\mathbb{E}\left[\mathbb{E}\left[\frac{\partial^2 f}{\partial x_k^2}(x_1,\dots,x_{k-1},\widehat{X}_k^{N_k},x_{k+1},\dots,x_d)|X_k-\widehat{X}_k^{N_k}|^2\right]\Big|_{\widehat{X}_{-k}=x_{-k}}\right]\\
&\quad=\sum_{k=1}^{d}\mathbb{E}\Big[\mathbb{E}\big[g_{k,x_{-k}}(\widehat{X}_k^{N_k})|X_k-\widehat{X}_k^{N_k}|^2\big]\big|_{\widehat{X}_{-k}=x_{-k}}\Big].
\end{aligned}
$$

Now, using Theorem 1.13, we have the following limits, for each k

$$
\lim_{N_k\to+\infty}N_k^2\,\mathbb{E}\big[g_{k,x_{-k}}(\widehat{X}_k^{N_k})|X_k-\widehat{X}_k^{N_k}|^2\big]=\mathcal{Q}_2(\mathbb{P}_{X_k})\int g_{k,x_{-k}}(\xi)\,\mathbb{P}_X(d\xi).
$$

Giving us the first part of the desired result

$$
\mathbb{E}\big[f(X)\big]=\mathbb{E}\big[f(\widehat{X}^N)\big]+\sum_{k=1}^{d}\frac{c_k}{N_k^2}+\int_0^1(1-t)\,\mathbb{E}\Big[\big(H(tX+(1-t)\widehat{X}^N)-H(\widehat{X}^N)\big)\cdot(X-\widehat{X}^N)^{\otimes 2}\Big]dt
$$

with $c_k:=\frac{1}{2}\mathcal{Q}_2(\mathbb{P}_{X_k})\iint g_{k,x_{-k}}(x)\,\mathbb{P}_{X_k}(dx)\,\mathbb{P}_{X_{-k}}(dy)$. Now, we take care of the integral part, we proceed using the same methodology as in the one dimensional case, using the hypothesis on the Hessian

$$
\mathbb{E}\Big[\big|\big(H(tX+(1-t)\widehat{X}^N)-H(\widehat{X}^N)\big)\cdot(X-\widehat{X}^N)^{\otimes 2}\big|\Big]\leqslant 2t^{\beta}[H]_{Lip}^{\beta}|||H|||_{\infty}^{1-\beta}\,\mathbb{E}\big[|X-\widehat{X}^N|^{2+\beta}\big]
$$

with $\beta\in(0,1)$ and $|||H|||_{\infty}:=\sup_{x\in\mathbb{R}^d}|||H(x)|||$. Hence

$$
\begin{aligned}
&\int_0^1(1-t)\,\mathbb{E}\Big[\big(H(tX+(1-t)\widehat{X}^N)-H(\widehat{X}^N)\big)\cdot(X-\widehat{X}^N)^{\otimes 2}\Big]dt\\
&\qquad\leqslant\frac{1}{(2+\beta)(1+\beta)}C_{H,X}\,\mathbb{E}\big[|X-\widehat{X}^N|^{2+\beta}\big].
\end{aligned}
$$

Using now Theorem 1.14, let $s=2+\beta$, we have the desired result: $\mathbb{E}\big[|X_k-\widehat{X}_k^{N_k}|^{2+\beta}\big]=O(N_k^{-(2+\beta)})$ and finally

$$
\mathbb{E}\big[f(X)\big]=\mathbb{E}\big[f(\widehat{X}^N)\big]+\sum_{k=1}^{d}\frac{c_k}{N_k^2}+O\Big(\Big(\min_{k=1:d}N_k\Big)^{-(2+\beta)}\Big),
$$

for every $\beta\in(0,1)$. If moreover, the densities φ_k of X_k, for all $k=1,\dots,k$, are Lipschitz continuous, bounded away from 0 on $[a,b]$ then we can take $\beta=1$.

□

Remark 3.4. Even-though, we could be interested by considering non-independent components $(X_k)_{k=1,\dots,d}$, the independence hypothesis on the components is necessary in the proof because we proceed component by component. For example the first order term of the expansion would not be null by stationarity if the components are not independent.

4 Applications

4.1 Quantized Control Variates in Monte Carlo simulations

Let $Z \in L^2(\mathbb{P})$ be a random vector with components $(Z_k)_{k=1,\dots,d}$, we assume that we have a closed-form for $\mathbb{E}[Z_k]$, $k = 1, \dots, d$, and $f : \mathbb{R}^d \to \mathbb{R}$ our function of interest. We are interested in the quantity

$$I := \mathbb{E}\big[f(Z)\big]. \tag{4.1}$$

The standard method for approximating (4.1) if we are able to simulate independent copies of Z is to devise a Monte Carlo estimator. In this part, we present a reduction variance method based on quantized control variates. Let Ξ_N our d dimensional control variate

$$\Xi^N := (\Xi_k^N)_{k=1,\dots,d}$$

where each component Ξ_k^N is defined by

$$\Xi_k^N := f_k(Z_k) - \mathbb{E}\big[f_k(\widehat{Z}_k^N)\big],$$

with $f_k(z) := f(\mathbb{E}[Z_1], \dots, \mathbb{E}[Z_{k-1}], z, \mathbb{E}[Z_{k+1}], \dots, \mathbb{E}[Z_d])$ and $\widehat{Z}_k^N$ is an optimal quantizer of cardinality N of the component Z_k. One notices that the complexity for the evaluation of f_k is the same as the one of f. Now, defining $X^\lambda := f(Z) - \langle \lambda, \Xi^N \rangle$ where $\lambda \in \mathbb{R}^d$, we can introduce $I^{\lambda,N}$ as an approximation for (4.1)

$$\begin{aligned} I^{\lambda,N} &:= \mathbb{E}\big[X^\lambda\big] \\ &= \mathbb{E}\big[f(Z) - \langle \lambda, \Xi^N \rangle\big] \\ &= \mathbb{E}\left[f(Z) - \sum_{k=1}^d \lambda_k f_k(Z_k)\right] + \sum_{k=1}^d \lambda_k \, \mathbb{E}\big[f_k(\widehat{Z}_k^N)\big]. \end{aligned} \tag{4.2}$$

The terms $\mathbb{E}\big[f_k(\widehat{Z}_k^N)\big]$ in (4.2) can be computed easily using the quantization-based cubature formula if we known the grids of the quantizers $(\widehat{Z}_k^N)_{k=1,\dots,d}$ and their associated weights.

Remark 4.1. We look for the $\lambda_{\min}$ minimizing the variance of X^λ

$$\mathbb{V}\mathrm{ar}(X^{\lambda_{\min}}) = \min\big\{ \mathbb{V}\mathrm{ar}\big(f(Z) - \langle \lambda, \Xi^N \rangle\big), \lambda \in \mathbb{R}^d \big\}.$$

The solution of the above optimization problem is the solution of following system

$$D(Z) \cdot \lambda = B$$

where $D(Z)$, the covariance-variance matrix of $\big(f_k(Z_k)\big)_{k=1,\dots,d}$, and B are given by

$$D(Z) = \begin{pmatrix} \mathbb{V}\mathrm{ar}\big(f_1(Z_1)\big) & \cdots & \mathbb{C}\mathrm{ov}\big(f_1(Z_1), f_d(Z_d)\big) \\ \vdots & \ddots & \vdots \\ \mathbb{C}\mathrm{ov}\big(f_d(Z_d), f_1(Z_1)\big) & \cdots & \mathbb{V}\mathrm{ar}\big(f_d(Z_d)\big) \end{pmatrix}, \quad B = \begin{pmatrix} \mathbb{C}\mathrm{ov}\big(f(Z), f_1(Z_1)\big) \\ \vdots \\ \mathbb{C}\mathrm{ov}\big(f(Z), f_d(Z_d)\big) \end{pmatrix}.$$

The solution to this optimization problem can easily be solved numerically using any library of linear algebra able to solve linear systems thanks to QR or LU decompositions.

Remark 4.2. If the Z_k's are independent hence λ can be determined easily. Indeed, in that case the matrix $D(Z)$ is diagonal. Then, the λ_k's are given by

$$\lambda_k = \frac{\mathbb{C}\mathrm{ov}\big(f_k(Z_k), f(Z)\big)}{\mathbb{V}\mathrm{ar}\big(f_k(Z_k)\big)}.$$

Now, we can define $\widehat{I}_M^{\lambda,N}$ the associated Monte Carlo estimator of $I^{\lambda,N}$

$$\widehat{I}_M^{\lambda,N} = \frac{1}{M}\sum_{m=1}^{M}\left(f(Z^m) - \sum_{k=1}^{d}\lambda_k f_k(Z_k^m)\right) + \sum_{k=1}^{d}\lambda_k\,\mathbb{E}\big[f_k(\widehat{Z}_k^N)\big].$$

One notices that $\mathbb{E}\big[I - I^{\lambda,N}\big] \neq 0$, with bias equal to $\sum_{k=1}^{d}\lambda_k\big(\mathbb{E}\big[f_k(\widehat{Z}_k^N)\big] - \mathbb{E}\big[f_k(Z_k)\big]\big)$. However the quantity we are really interested by is not the bias but the *MSE* (Mean Squared Error), yielding a *bias-variance decomposition*

$$\mathrm{MSE}(\widehat{I}_M^{\lambda,N}) = \underbrace{\left(\sum_{k=1}^{d}\lambda_k\Big(\mathbb{E}\big[f_k(\widehat{Z}_k^N)\big] - \mathbb{E}\big[f_k(Z_k)\big]\Big)\right)^2}_{bias^2} + \frac{1}{M}\underbrace{\mathbb{V}\mathrm{ar}\left(f(Z) - \sum_{k=1}^{d}\lambda_k f_k(Z_k)\right)}_{Monte\ Carlo\ variance}.$$

Our aim is to minimize the cost of the Monte Carlo simulation for a given *MSE* or upper-bound of the *MSE*. Consequently, for a given Monte Carlo estimator $\widehat{I}_M^{\lambda,N}$ our minimization problem reads

$$\inf_{MSE(\widehat{I}_M^{\lambda,N})\leqslant\epsilon^2}\mathrm{Cost}(\widehat{I}_M^{\lambda,N}). \tag{4.3}$$

Let $\kappa = \mathrm{Cost}(f(z))$ for a given $z \in \mathbb{R}^d$, the cost of a standard Monte Carlo estimator $\widehat{I}_M$ of size M is $\mathrm{Cost}(\widehat{I}_M) = \kappa M$. In our controlled case, if we neglect the cost for building an optimal quantizer, the global complexity associated to the Monte-Carlo estimator $\widehat{I}_M^{\lambda,N}$ is given by

$$\mathrm{Cost}(\widehat{I}_M^{\lambda,N}) = \kappa\big((d+1)M + dN\big)$$

where the cost of the computation of $f(z) - \lambda\sum_{k=1}^{d} f_k(z)$ is upper-bounded by $(d+1)\kappa$ whereas κdN is the cost of the quantized part. Indeed, there is d expectations of functions of N-quantizers to compute, inducing a cost of order κdN. Some optimizations can be implemented when computing $f_k(z)$, in that case $\mathrm{Cost}(f_k(z)) < \kappa$. So, (4.3) becomes

$$\inf_{MSE(\widehat{I}_M^{\lambda,N})\leqslant\epsilon^2}\kappa\big((d+1)M + dN\big).$$

Moreover, using the results in the first part of the paper concerning the weak error, we could define an upper-bound for the $MSE(\widehat{I}_M^{\lambda,N})$, indeed if each f_k is in a class of function where the weak error of order two is attained when using a quantization-based cubature formula then

$$MSE(\widehat{I}_M^{\lambda,N}) = \left(\sum_{k=1}^{d}\lambda_k\Big(\mathbb{E}\big[f_k(\widehat{Z}_k^N)\big] - \mathbb{E}\big[f_k(Z_k)\big]\Big)\right)^2 + \frac{\sigma_\lambda^2}{M} \leqslant \frac{C}{N^4} + \frac{\sigma_\lambda^2}{M}$$

with $\sigma_\lambda^2 := \mathbb{V}\mathrm{ar}\big(f(Z) - \sum_{k=1}^{d}\lambda_k f_k(Z_k)\big)$. Now, our minimization problem becomes

$$\inf_{\frac{C}{N^4}+\frac{\sigma_\lambda^2}{M}\leqslant\epsilon^2}\kappa\big((d+1)M + dN\big).$$

$\frac{C}{N^4}$ corresponds to the squared empirical bias and $\frac{\sigma_\lambda^2}{M}$ to the empirical variance, hence a standard approach when dealing with this kind of problem, is to equally divide ϵ^2 between the bias and the variance: $\frac{C}{N^4} = \frac{\epsilon^2}{2}$ and $\frac{\sigma_\lambda^2}{M} = \frac{\epsilon^2}{2}$ yielding

$$N = O(\epsilon^{-\frac{1}{2}}) \quad \text{and} \quad M = O(\epsilon^{-2}),$$

hence the cost would be of order $O(\epsilon^{-2})$. However, as the cost is additive and in the case where σ_λ^2 is close to $\mathbb{V}\mathrm{ar}\big(f(Z)\big)$, meaning that the control variate does not really reduce the variance, we want to reduce the bias as much as we can. So another idea could be to choose both terms M and N of order $O(\epsilon^{-2})$, because the impact on the cost of the Monte Carlo is at least of this order. Then, we search $\theta \in (0,1)$ defined by

$$\theta\epsilon^2 = \frac{C}{N^4} \quad \text{and } (1-\theta)\epsilon^2 = \frac{\sigma_\lambda^2}{M},$$

such that the impact on the cost of the Monte Carlo part and the quantization part are of same order: $O(\epsilon^{-2})$. In that case, θ is given by

$$\begin{cases} \theta\epsilon^2 = \frac{C}{N^4} \\ \kappa d N = O(\epsilon^{-2}) \end{cases} \qquad \Longrightarrow \qquad \theta = O(\epsilon^6).$$

In practice, we do take not that high value for N. Indeed, the bias converges to 0 as N^{-4}, so taking optimal quantizers of size 200 or 500 is enough for considering that the bias is negligible compared to the residual variance of the Monte Carlo estimator.

Remark 4.3. Now, if we consider that we have no closed-form for $\mathbb{E}[Z_k]$, $k = 1, \dots, d$, then we need to approximate them by m_k (this would impact the total cost of the method, as one would need to use a numerical method for computing the m_k's but this can be done once and for all before estimating $\widehat{I}_M^{\lambda,N}$). These approximations yield different control variates: the functions $\widetilde{f}_k(z) := f(m_1, \dots, m_{k-1}, z, m_{k+1}, \dots, m_d)$, inducing a different *MSE*

$$MSE(\widehat{I}_M^{\widetilde{\lambda},N}) = \left(\sum_{k=1}^{d} \widetilde{\lambda}_k \Big(\mathbb{E}\left[\widetilde{f}_k(\widehat{Z}_k^N)\right] - \mathbb{E}\left[\widetilde{f}_k(Z_k)\right] \Big) \right)^2 + \frac{\widetilde{\sigma}_{\widetilde{\lambda}}^2}{M}$$

with $\widetilde{\sigma}_{\widetilde{\lambda}}^2 := \mathbb{V}\mathrm{ar}\left(f(Z) - \sum_{k=1}^d \widetilde{\lambda}_k \widetilde{f}_k(Z_k)\right)$ and $\widetilde{\lambda}_k$, $k = 1, \dots, d$. Finally, we can conclude in the same way as before if the $\widetilde{f}_k$'s are in a class of function where the weak error of order two is attained when using a quantization-based cubature formula.

4.2 Numerical results

Let $(S_t)_{t\in[0,T]}$ be a geometric Brownian motion representing the dynamic of a *Black-Scholes* asset between time $t = 0$ and time $t = T$ defined by

$$S_t = S_0\, \mathrm{e}^{(r-\sigma^2/2)t + \sigma W_t}$$

with $(W_t)_{t\in[0,T]}$ a standard Brownian motion defined on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$, r the interest rate and σ the volatility. When considering to use optimal quantization with a Black-Scholes asset, we have two possibilities: either we take an optimal quantizer of a normal distribution as $W_T \sim \mathcal{N}(0,T)$ or we build an optimal quantizer of a log-normal distribution as $\log(\mathrm{e}^{(r-\sigma^2/2)T+\sigma W_T}) \sim \mathcal{N}\left((r-\sigma^2/2)T, \sigma^2 T\right)$. In this part we consider both approaches since each one has its benefits and drawbacks.

Optimal Quantizers of log-normal random variables need to be computed each time we consider different parameters for the Black-Scholes asset. Indeed, the only operations preserving the optimality of the quantizers are translations and scaling. However, this transformations are not enough if one wishes to build an optimal quantizer of a Log-Normal random variables with parameters μ and σ from an optimal quantizer of a standardized Log-Normal random variable. However, if one looses time by computing for each set of parameters an optimal quantizer for the log-normal random variable, it gains in precision.

Now, if we consider the case of optimal quantizers of normal random variables, we loose in precision because we do not quantize directly our asset but the optimal quantizers of normal random variables can be computed once and for all and stored on a file. Indeed, we can build every normal random variable from a standard normal random variable using translations and scaling. Moreover, high precision grids of the $\mathcal{N}(0,1)$-distribution are in free access for download at the website: `www.quantize.maths-fi.com`.

Substantial details concerning the optimization problem and the numerical methods for building quadratic optimal quantizers can be found in [Pag18, PP03, PPP04, MRKP18]. In our case, we chose to build all the optimal quantizers with the Newton-Raphson algorithm (see [PP03] for more details on the gradient and Hessian formulas for the $\mathcal{N}(0,1)$-distribution and [MRKP18] for other distributions) modified with the Levenberg-Marquardt procedure which improves the robustness of the method.

4.2.1 Vanilla Call

The payoff of a Call expiring at time T is

$$(S_T - K)_+$$

with K the strike and T the maturity of the option. Its price, in the special case of *Black-Scholes* model, is given by the following closed formula

$$I_0 := \mathbb{E}\left[\mathrm{e}^{-rT}(S_T - K)_+\right] = Call_{BS}(S_0, K, r, \sigma, T) = S_0\,\mathcal{N}(d_1) - K\,\mathrm{e}^{-rT}\,\mathcal{N}(d_2) \tag{4.4}$$

where $\mathcal{N}(x)$ is the cumulative distribution function of the standard normal distribution, $d_1 := \frac{\log(S_0/K)+(r+\sigma^2/2)T}{\sigma\sqrt{T}}$ and $d_2 := d_1 - \sigma\sqrt{T}$. Although the price of a Call in the Black-Scholes model can be expressed in a closed form, it is a good exercise to test new numerical methods against this benchmark. We compare the use of optimal quantizers of normal distribution, when one quantizes the law of the Brownian motion at time T and log-normal distribution when one quantizes directly the law of the asset S_T at time T.

In the first case, we can rewrite I_0 as a function of a random variable Z with a $\mathcal{N}(0,1)$-distribution, namely a normal distributed random variable,

$$\mathbb{E}\left[\mathrm{e}^{-rT}(S_T - K)_+\right] = \mathbb{E}\left[f(Z)\right]$$

where $f(x) := \mathrm{e}^{-rT}(s_0\,\mathrm{e}^{(r-\sigma^2/2)T+\sigma\sqrt{T}x} - K)_+$ is continuous with a piecewise-defined locally-Lipschitz derivative, with respect to the function $g(x) = \mathrm{e}^{\sigma\sqrt{T}|x|}$.

In the second case, we have

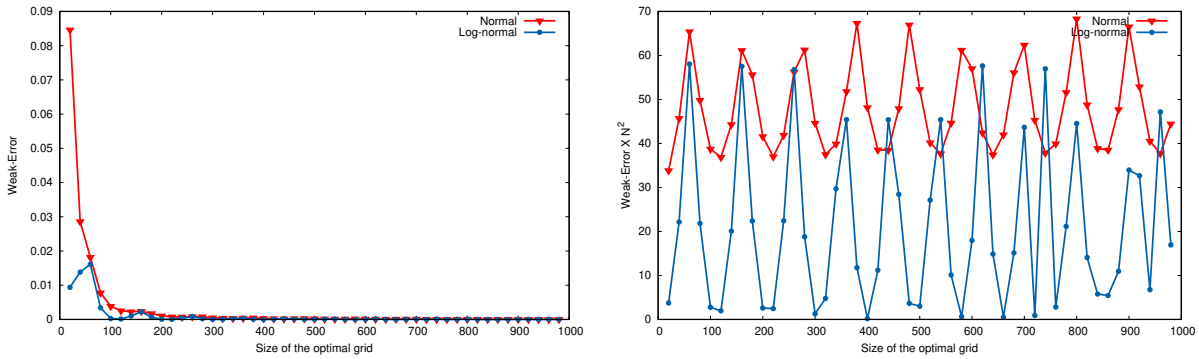
$$\mathbb{E}\left[\mathrm{e}^{-rT}(S_T - K)_+\right] = \mathbb{E}\left[\varphi(S_T)\right]$$

where $\varphi(x) := \mathrm{e}^{-rT}(x - K)_+$ is piecewise affine with one break of affinity.

The Black-Scholes parameters considered are

$$s_0 = 100, \qquad r = 0.1, \qquad \sigma = 0.5,$$

whereas those of the Call option are $T = 1$ and $K = 80$. The reference value is 34.15007. The first graphic in the Figure 1 represents the weak error between the benchmark and the quantization-based approximations in function of the size of the grid: $N \longmapsto \left|I_0 - \mathbb{E}\left[f(\widehat{Z}^N)\right]\right|$ and $N \longmapsto \left|I_0 - \mathbb{E}\left[\varphi(\widehat{X}^N)\right]\right|$, the second represents the weak error multiplied by N^2 in function of N: $N \longmapsto N^2 \times \left|I_0 - \mathbb{E}\left[f(\widehat{Z}^N)\right]\right|$ and $N \longmapsto N^2 \times \left|I_0 - \mathbb{E}\left[\varphi(\widehat{X}^N)\right]\right|$.



(a) $N \longmapsto \left|I_0 - \mathbb{E}\left[f(\widehat{Z}^N)\right]\right|$ *(▼) and* $N \longmapsto \left|I_0 - \mathbb{E}\left[\varphi(\widehat{X}^N)\right]\right|$ *(●)*

(b) $N \longmapsto N^2 \times \left|I_0 - \mathbb{E}\left[f(\widehat{Z}^N)\right]\right|$ *(▼) and* $N \longmapsto N^2 \times \left|I_0 - \mathbb{E}\left[\varphi(\widehat{X}^N)\right]\right|$ *(●)*

Figure 1: *Call option in a Black-Scholes model.*

First, we notice that both methods yield a weak-error of order 2, as desired. Second, if we look closely at the results the log-normal grids give a more precise price. However we need to build a specific grid each time we have a new set of parameters for the asset, whereas such is not the case when we choose to quantize the normal random variable, we can directly read precomputed grids with their associated weights in files.

4.2.2 Compound Option

The second product we consider is a Compound Option: a Put-on-Call. The payoff of a Put-on-Call expiring at time T_1 is the following

$$\left(K_1 - \mathbb{E}\left[\,\mathrm{e}^{-r(T_2-T_1)}(S_{T_2} - K_2)_+ \mid S_{T_1}\right]\right)_+$$

with price

$$I_0 := \mathbb{E}\left[\,\mathrm{e}^{-rT_1}\left(K_1 - \mathbb{E}\left[\,\mathrm{e}^{-r(T_2-T_1)}(S_{T_2} - K_2)_+ \mid S_{T_1}\right]\right)_+\right]. \tag{4.5}$$

The inner expectation can be computed, using the fact that S_{T_2} is a *Black-Scholes* asset and we know the conditional law of S_{T_2} given S_{T_1}. Using (4.4), the value of the inner expectation is

$$\mathbb{E}\left[\,\mathrm{e}^{-r(T_2-T_1)}(S_{T_2} - K_2)_+ \mid S_{T_1}\right] = Call_{BS}(S_{T_1}, K_2, r, \sigma, T_2 - T_1).$$

Hence, the price of the Put-On-Call option in (4.5) can be rewritten as

$$I_0 = \mathbb{E}\left[\,\mathrm{e}^{-rT_1}\left(K_1 - Call_{BS}(S_{T_1}, K_2, r, \sigma, T_2 - T_1)\right)_+\right].$$

The Black-Scholes parameters considered are

$$s_0 = 100, \qquad r = 0.03, \qquad \sigma = 0.2,$$

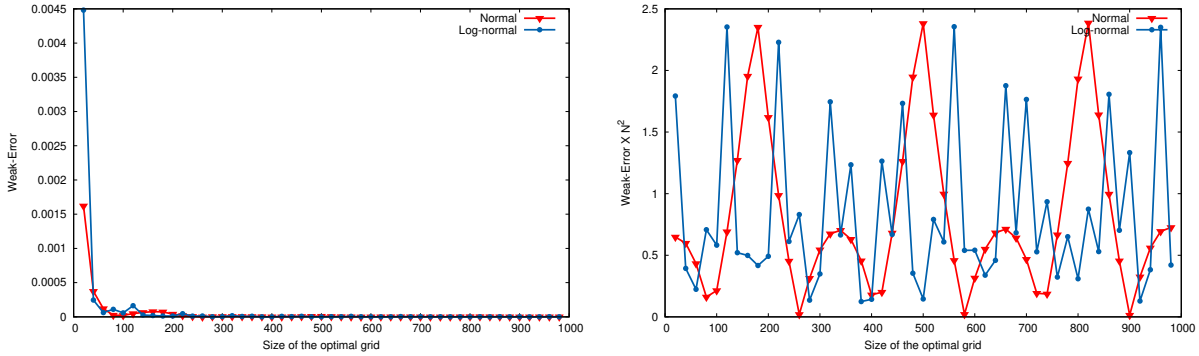
whereas those of the Put-On-Call option are $T_1 = 1/12$, $T_2 = 1/2$, $K_1 = 6.5$ and $K_2 = 100$. The reference value, obtained using an optimal quantizer of size 10000 of the $\mathcal{N}(0,1)$-distribution, is 1.3945704. As in the vanilla case, we compare the use of optimal quantizers of normal distribution and log-normal distribution. In the first case, we have

$$I_0 = \mathbb{E}\big[f(Z)\big]$$

where $Z \sim \mathcal{N}(0,1)$ and $f(z) = \mathrm{e}^{-rT_1}\left(K_1 - Call_{BS}(s_0\,\mathrm{e}^{(r-\sigma^2/2)T_1+\sigma\sqrt{T_1}z}, K_2, r, \sigma, T_2 - T_1)\right)_+$, and in the second case

$$I_0 = \mathbb{E}\big[\varphi(X)\big]$$

where $\log(X) \sim \mathcal{N}((r-\sigma^2/2)T, \sigma\sqrt{T})$ and $\varphi(x) = \mathrm{e}^{-rT_1}\left(K_1 - Call_{BS}(s_0 x, K_2, r, \sigma, T_2 - T_1)\right)_+$. The first graphic in Figure 2 represents the weak error between the benchmark and the quantization-based approximations in function of the size of the grid: $N \longmapsto \big|I_0 - \mathbb{E}\big[f(\widehat{Z}^N)\big]\big|$ and $N \longmapsto \big|I_0 - \mathbb{E}\big[\varphi(\widehat{X}^N)\big]\big|$, the second allows us to observe if the rate of convergence is indeed of order 2.



(a) $N \longmapsto \big|I_0 - \mathbb{E}\big[f(\widehat{Z}^N)\big]\big|$ (▼) *and* $N \longmapsto \big|I_0 - \mathbb{E}\big[\varphi(\widehat{X}^N)\big]\big|$ (●)

(b) $N \longmapsto N^2 \times \big|I_0 - \mathbb{E}\big[f(\widehat{Z}^N)\big]\big|$ (▼) *and* $N \longmapsto N^2 \times \big|I_0 - \mathbb{E}\big[\varphi(\widehat{X}^N)\big]\big|$ (●)

Figure 2: *option in a Black-Scholes model.*

We notice that both methods yield a weak-error of order 2 as desired, however it is not clear that one should use the log-normal representation of (4.5) in place of the Gaussian representation. Indeed, both constants in the rate of convergence are of the desired order and getting Gaussian optimal quantizers is much cheaper than building optimal quantizers of log-normal random variables. Hence, one should choose the Gausian representation as it is as precise as the log-normal one and is much cheaper.

4.2.3 Exchange spread Option

In this part, we consider a higher dimensional problem. Let two Black-Scholes assets $(S_T^i)_{i=1,2}$ at time T related to two Brownian motions $(W_T^i)_{i=1,2}$, with correlation $\rho \in [-1,1]$. We are interested by an exchange spread option with strike K with payoff

$$(S_T^1 - S_T^2 - K)_+$$

whose price is

$$I_0 := \mathbb{E}\big[\,\mathrm{e}^{-rT}(S_T^1 - S_T^2 - K)_+\big]. \tag{4.6}$$

Decomposing the two Brownian motions into two independent parts, we have $(W_T^1, W_T^2) = \sqrt{T}(\sqrt{1-\rho^2}Z_1 + \rho Z_2, Z_2)$, where Z_1 and Z_2 are two independent $\mathcal{N}(0,1)$-distributed Gaussian

random variables. Now, pre-conditioning on Z_2 in (4.6) and using (4.4), we have

$$I_0 = \mathbb{E}\left[\varphi(Z_2)\right]$$

where

$$\varphi(z) = Call_{BS}(s_0^1\, \mathrm{e}^{-\rho^2\sigma_1^2 T/2+\sigma_1\rho\sqrt{T}z}, s_0^2\, \mathrm{e}^{(r-\sigma_2^2/2)T+\sigma_2\sqrt{T}z} + K, r, \sigma_1\sqrt{1-\rho^2}, T).$$

The numerical specifications of the function φ are as follows:

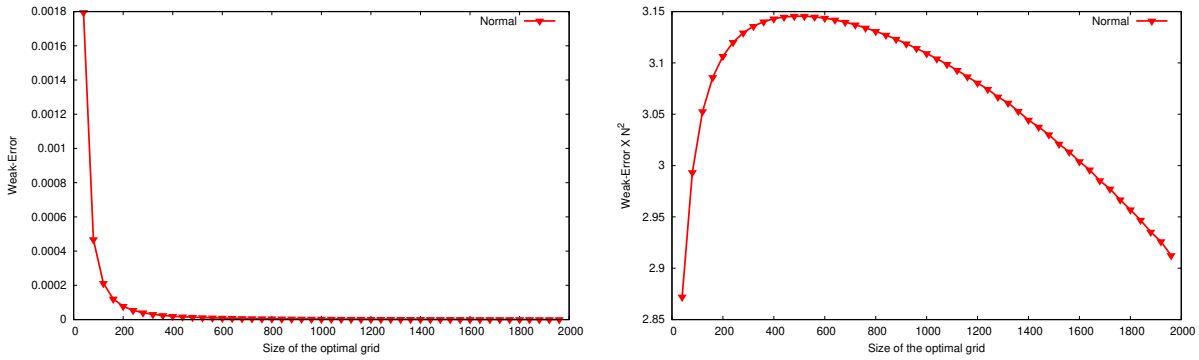
$$s_0^i = 100, \quad r = 0.02, \quad \sigma_i = 0.5, \quad \rho = 0.5, \quad T = 10, \quad K = 10.$$

In that case, the reference value is 53.552678.

First, we look at the weak error induced by the quantization-based cubature formula when approximating (4.6). We use optimal quantizers of the normal random variable Z_2. The quantization-based approximation is denoted $\widehat{I}_N$,

$$\widehat{I}_N := \mathbb{E}\left[\varphi(\widehat{Z}^N)\right].$$

The first graphic in Figure 3 represents the weak error between the benchmark and the quantization-based approximation in function of the size of the grid: $N \longmapsto \left|I_0 - \mathbb{E}\left[\varphi(\widehat{Z}^N)\right]\right|$, the second plots $N \longmapsto N^2 \times \left|I_0 - \mathbb{E}\left[\varphi(\widehat{Z}^N)\right]\right|$ and allows us to observe that the rate of convergence is indeed of order 2.

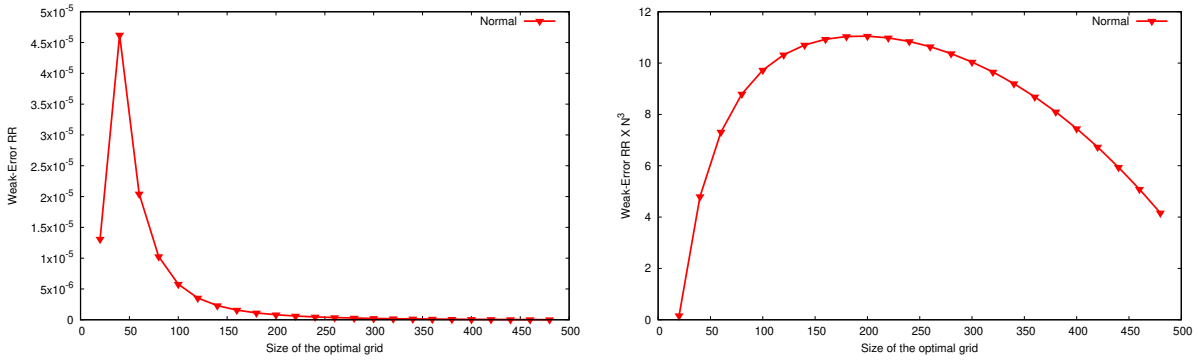


(a) $N \longmapsto \left|I_0 - \mathbb{E}\left[\varphi(\widehat{Z}^N)\right]\right|$ (▼)

(b) $N \longmapsto N^2 \times \left|I_0 - \mathbb{E}\left[\varphi(\widehat{Z}^N)\right]\right|$ (▼)

Figure 3: *Exchange spread option pricing in a Black-Scholes model.*

Now, noticing that φ is a twice differentiable function with a bounded second derivative, we show that we can attain a weak error of order 3 when using a Richardson-Romberg extrapolation denoted $\widehat{I}^{RR}_{\widetilde{N},N}$ and defined in (3.1).



(a) $N \longmapsto |I_0 - \widehat{I}^{RR}_{\widetilde{N},N}|$ (▼) (b) $N \longmapsto N^3 \times |I_0 - \widehat{I}^{RR}_{\widetilde{N},N}|$ (▼)

Figure 4: *Richardson-Romberg extrapolation, with* $\widetilde{N} = 1.2 \times N$*, for Exchange spread option pricing in a Black-Scholes model.*

4.2.4 Basket Option

A typical financial product that allows to diversify the market risk and to invest in options is a basket option. The simplest one is an option on a weighted average of stocks. For example, if we consider an option on the FTSE index, this is a basket option where the assets are the companies defined in the description of the index and the weights are the market capitalization of each company at the time we built the index normalized by the sum on all market capitalizations.

In this part, we consider d correlated assets $(S_T^k)_{k=1,\dots,d}$ following a Black-Scholes model and the payoff we consider is

$$f(S_t^1, \dots, S_T^d) := \left(\sum_{k=1}^{d} \alpha_k S_T^k - K \right)_+ \tag{4.7}$$

whose price is

$$I_0 := \mathrm{e}^{-rT}\, \mathbb{E}\left[\left(\sum_{k=1}^{d} \alpha_k S_T^k - K \right)_+ \right].$$

I_0 cannot be computed directly, hence we use a Monte Carlo estimator in order to approximate the expectation. The standard estimator, denoted $\widehat{I}_M$, is the crude Monte Carlo estimator and is given by

$$\widehat{I}_M := \mathrm{e}^{-rT}\, \frac{1}{M} \sum_{m=1}^{M} \left(\sum_{k=1}^{d} \alpha_k S_T^{k,(m)} - K \right)_+$$

where $(S_T^{k,(m)})_{m=1,\dots,M}$ are i.i.d. copies of S_T^k. We compare the crude estimator to our novel approach based on a d-dimensional quantized control variates Ξ^N. In that case, I_0 is approximated by I^N defined by

$$I^N := \mathrm{e}^{-rT}\, \mathbb{E}\left[\left(\sum_{k=1}^{d} \alpha_k S_T^k - K \right)_+ - \langle \lambda, \Xi^N \rangle \right]$$

where Ξ^N is defined later, yielding the following Monte Carlo estimator

$$\widehat{I}_M^{\lambda,N} := \mathrm{e}^{-rT}\, \frac{1}{M} \sum_{m=1}^{M} \left(\sum_{k=1}^{d} \alpha_k S_T^{k,(m)} - K \right)_+ - \langle \lambda, \Xi^{N,(m)} \rangle.$$

We propose two different control variates Ξ^N based on optimal quantizers either of log-normal random variables or of Gaussian random variables.

1. The control variate, denoted $\overline{\Xi}^N$, is defined by, $\forall k = 1, \dots, d$

$$\overline{\Xi}_k^N := f(\mathbb{E}[S_T^1], \dots, S_T^k, \dots, \mathbb{E}[S_T^d]) - \mathbb{E}\left[f(\mathbb{E}[S_T^1], \dots, \widehat{S}_T^{k,N}, \dots, \mathbb{E}[S_T^d])\right]$$

where $(\widehat{S}_T^{k,N})_{k=1,\dots,d}$ are optimal quantizers of cardinality N of S_T^k. In that case, the Monte Carlo estimator is denoted $\widehat{\overline{I}}_M^{\lambda,N}$.

2. The control variate, denoted $\widetilde{\Xi}^N$, is using another representation of the payoff (4.7), using d Gaussian random variables i.i.d in place of the assets S_T^k because the d underlying correlated Brownian Motions can be expressed from d rescaled independent Gaussian random variables, thus we define φ our new representation for the payoff as

$$\varphi(Z^1, \dots, Z^d) := f(S_T^1, \dots, S_T^d)$$

where $(Z^k)_{k=1,\dots,d}$ are i.i.d Gaussian random variables. Now, defining our control variates with the function φ, $\forall k = 1, \dots, d$

$$\widetilde{\Xi}_k^N := \varphi(0, \dots, Z^k, \dots, 0) - \mathbb{E}\left[\varphi(0, \dots, \widehat{Z}^N, \dots, 0)\right]$$

where $(\widehat{Z}^N)_{k=1,\dots,d}$ is an optimal quantizer of $Z \sim \mathcal{N}(0,1)$. In that case, the Monte Carlo estimator is denoted $\widehat{\widetilde{I}}_M^{\lambda,N}$.

The Black-Scholes parameters considered are

$$s_0^i = 100, \qquad r = 2\%, \qquad \sigma_i = \frac{i}{d+1}, \quad \rho = 0.5,$$

and the specifications of the product are

$$K = 100, \quad \alpha_i = \frac{2i}{d(d+1)}, \quad T = 1$$

such that $\sum \alpha_i = 1$. The benchmarks used for the computation of the *MSE* has been computed using a Monte Carlo estimator with control variate without quantization where the term $\sum_{k=1}^d \mathbb{E}[X_k]$ is computed using Black-Scholes Call pricing closed formulas. The *Mean Squared Error* of an estimator I is computed using the formula

$$MSE(I) = \frac{1}{n}\sum_{i=1}^{n}(I^{(i)} - I_0)^2$$

where $(I^{(i)})_{i=1,\dots,n}$ are n independent copies of I.

Table 1 compares three different types of Monte Carlo estimators: the standard (Crude) Monte Carlo estimator $\widehat{I}_M$, our novel Monte Carlo estimator with control variate based on optimal quantizers of Gaussian random variables $\widehat{\widetilde{I}}_M^{\lambda,N}$ and another one with optimal quantizers of log-normal random variables $\widehat{\overline{I}}_M^{\lambda,N}$. The notation n corresponds to the number of Monte Carlo used for computing the *MSE*, M is the size of each Monte Carlo and N is the size of the optimal quantizers. The prices of reference for each d are

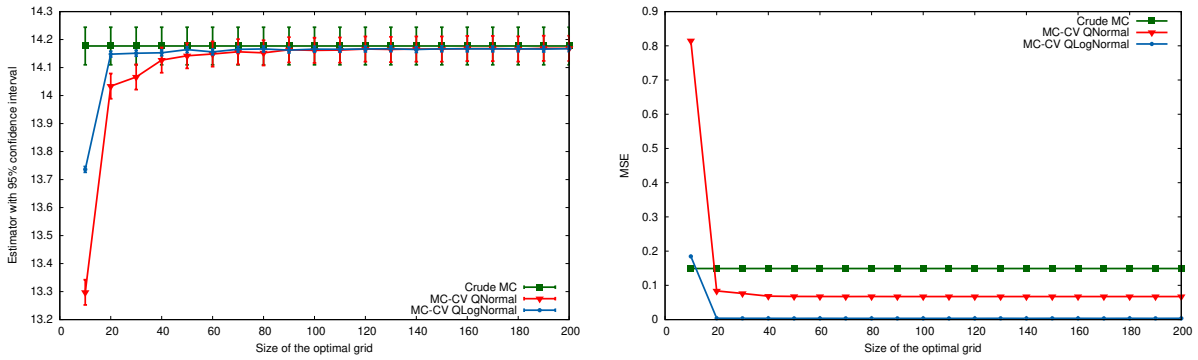
- for $d = 2$: 14.2589 (± 0.0010),
- for $d = 3$: 14.1618 (± 0.0015),
- for $d = 5$: 13.9005 (± 0.0022),

		$N = 20$		$N = 200$	
d	MC Estimator	Mean ($\pm 1.96 \times$std)	MSE	Mean ($\pm 1.96 \times$std)	MSE
$d = 2$	Crude	14.2695 ($\pm$0.0662)	0.1450	14.2695 ($\pm$0.0662)	0.1450
	CV Gaussian	14.1017 ($\pm$0.0399)	0.0774	14.2773 ($\pm$0.0399)	0.0530
	CV Log-Normal	14.2351 ($\pm$0.0078)	0.0026	14.2614 ($\pm$0.0078)	0.0020
$d = 3$	Crude MC	14.1770 ($\pm$0.0671)	0.1492	14.1770 ($\pm$0.0671)	0.1492
	CV Gaussian	14.0336 ($\pm$0.0451)	0.0837	14.1685 ($\pm$0.0451)	0.0673
	CV Log-Normal	14.1479 ($\pm$0.0104)	0.0038	14.1674 ($\pm$0.0104)	0.0036
$d = 5$	Crude MC	13.8803 ($\pm$0.0720)	0.1717	13.8803 ($\pm$0.0720)	0.1717
	CV Gaussian	13.6686 ($\pm$0.0562)	0.1580	13.8883 ($\pm$0.0562)	0.1044
	CV Log-Normal	13.8797 ($\pm$0.0151)	0.0080	13.9008 ($\pm$0.0151)	0.0076
$d = 10$	Crude MC	13.5046 ($\pm$0.0599)	0.1186	13.5046 ($\pm$0.0599)	0.1186
	CV Gaussian	13.2429 ($\pm$0.0515)	0.1527	13.5113 ($\pm$0.0515)	0.0878
	CV Log-Normal	13.4221 ($\pm$0.0194)	0.0181	13.4983 ($\pm$0.0194)	0.0124

Table 1: $n = 128$, $M = 1e4$

- for $d = 10$: 13.4979 ($\pm$0.0034).

One remarks in Table 1 the efficiency of the optimal quantization-based variance reduction method. The variance, in the best cases, can be divided by almost 100 when using the optimal quantizers of Log-Normal random variables. Figure 5 shows the effect of N (for $d = 3$), the size the optimal quantizers, on the bias. The same seeds are used for all the Monte Carlo estimator, the only thing varying is N.



(a) $N \longmapsto |I_0 - \widehat{\widehat{I}}_M^{\lambda,N}|$ (▼), $N \longmapsto |I_0 - \widehat{\widehat{I}}_M^{\lambda,N}|$ (●) *and the Crude Monte Carlo estimator* (■) *with their associated confidence interval at* 95%.

(b) $N \longmapsto MSE(\widehat{I}_M)$ (■), $N \longmapsto MSE(\widehat{\widehat{I}}_M^{\lambda,N})$ (▼) *and* $N \longmapsto MSE(\widehat{\widehat{I}}_M^{\lambda,N})$ (●).

Figure 5: $n = 128$, $M = 1e4$, $d = 3$.

Acknowledgment

The authors wish to thank Pauline Corblet and Eric Tea for their useful feedback. The PhD thesis of Thibaut Montes is funded by a CIFRE grand from The Independent Calculation Agent (The ICA) and French ANRT.

References

[BP03] Vlad Bally and Gilles Pagès. A quantization algorithm for solving multidimensional discrete-time optimal stopping problems. *Bernoulli*, 9(6):1003–1049, 2003.

[BPP01] Vlad Bally, Gilles Pagès, and Jacques Printems. A stochastic quantization method for nonlinear problems. *Monte Carlo Methods and Applications*, 7:21–34, 2001.

[BPP05] Vlad Bally, Gilles Pagès, and Jacques Printems. A quantization tree method for pricing and hedging multi-dimensional american options. *Mathematical Finance*, 15(1):119–168, 2005.

[CM01] Peter Carr and Dilip Madan. Optimal positioning in derivative securities. *Quantitative Finance*, 1(1):19–37, 2001.

[DFP04] Sylvain Delattre, Jean-Claude Fort, and Gilles Pagès. Local distortion and μ-mass of the cells of one dimensional asymptotically optimal quantizers. *Communications in Statistics - Theory and Methods*, 33(5):1087–1117, 2004.

[DGLP04] Sylvain Delattre, Siegfried Graf, Harald Luschgy, and Gilles Pagès. Quantization of probability distributions under norm-based distortion measures. *Statistics & Decisions*, 22(4):261–282, 2004.

[GG82] Allen Gersho and Robert M Gray. Special issue on quantization. *IEEE Transactions on Information Theory*, 29, 1982.

[GL00] Siegfried Graf and Harald Luschgy. *Foundations of Quantization for Probability Distributions*. Springer-Verlag, Berlin, Heidelberg, 2000.

[Gla13] Paul Glasserman. *Monte Carlo methods in financial engineering*, volume 53. Springer Science & Business Media, 2013.

[GLP08] Siegfried Graf, Harald Luschgy, and Gilles Pagès. Distortion mismatch in the quantization of probability measures. *ESAIM: Probability and Statistics*, 12:127–153, 2008.

[MRKP18] Thomas A McWalter, Ralph Rudd, Jörg Kienitz, and Eckhard Platen. Recursive marginal quantization of higher-order schemes. *Quantitative Finance*, 18(4):693–706, 2018.

[Pag98] Gilles Pagès. A space quantization method for numerical integration. *Journal of computational and applied mathematics*, 89(1):1–38, 1998.

[Pag07] Gilles Pagès. Multi-step richardson-romberg extrapolation: remarks on variance control and complexity. *Monte Carlo Methods and Applications*, 13(1):37–70, 2007.

[Pag15] Gilles Pagès. Introduction to vector quantization and its applications for numerics. *ESAIM: proceedings and surveys*, 48:29–79, 2015.

[Pag18] Gilles Pagès. *Numerical Probability: An Introduction with Applications to Finance*. Springer, 2018.

[PP03] Gilles Pagès and Jacques Printems. Optimal quadratic quantization for numerics: the gaussian case. *Monte Carlo Methods and Applications*, 9(2):135–165, 2003.

[PPP04] Gilles Pagès, Huyên Pham, and Jacques Printems. *Optimal Quantization Methods and Applications to Numerical Problems in Finance*, pages 253–297. Birkhäuser Boston, 2004.

[PS12] Gilles Pagès and Abass Sagna. Asymptotics of the maximal radius of an l^r-optimal sequence of quantizers. *Bernoulli*, 18(1):360–389, 2012.

[PS18] Gilles Pagès and Abass Sagna. Improved error bounds for quantization based numerical schemes for bsde and nonlinear filtering. *Stochastic Processes and their Applications*, 128(3):847–883, 2018.

[RG10] Lewis Fry Richardson and Richard Tetley Glazebrook. On the approximate arithmetical solution by finite differences of physical problems involving differential equations, with an application to the stresses in a masonry dam. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, 83(563):335–336, 1910.

[Rom55] Werner Romberg. Vereinfachte numerische integration. *Norske Vid. Selsk. Forh.*, 28:30–36, 1955.

[She97] William Fleetwood Sheppard. On the calculation of the most probable values of frequency-constants, for data arranged according to equidistant division of a scale. *Proceedings of the London Mathematical Society*, 1(1):353–380, 1897.

[TT90] Denis Talay and Luciano Tubaro. Romberg extrapolations for numerical schemes solving stochastic differential equations. *Structural Safety*, 8(1-4):143–150, 1990.