Marginal quantization of an Euler diffusion process and its application

ABASS SAGNA *

Abstract

We propose a new approach to quantize the marginals of the discrete Euler process resulting from the discretization of a brownian diffusion process using the Euler scheme. The method is built recursively using the distribution of the marginals of the discrete Euler process. The quantization error associated to the marginals is shown to goes toward 0 at the optimal rate associated to the quantization of an $\mathbb{R}^d$-valued random vector. In the one dimensional setting we illustrate how to perform the optimal grids using the Newton algorithm and show how to estimate the associated weights from a recursive formula. Numerical tests are carried out for the pricing of European options in a local volatility model and a comparison with the Monte Carlo simulations shows that the proposed method is more efficient than the Monte Carlo method.

1 Introduction

Optimal quantization method appears first in [20] where the author studies in particular the optimal quantization problem for the uniform distribution. It has become an important field of information theory since the early 1940's. A common use of quantization is the conversion of a continuous signal into a discrete signal that assumes only a finite number of values.

Since then, optimal quantization is applied in many fields as in Physics, in Computer Sciences, ..., and recently in Numerical Probability from the seminal work [12]. Its application to Numerical Probability relies on the possibility to resume a random vector taking values in a set of infinite cardinality by a discrete random vector valued in a set of finite cardinality with the associated weights. This allows to approximate either expectations or more significantly some conditional expectations. This is the key property used to solve some problems arising in Quantitative Finance as optimal stopping problems (see [1, 2]), the pricing of swing options (see [3]), stochastic control problems (see [6, 14]), nonlinear filtering problems (see e.g. [13, 17, 5, 18]), the pricing of barrier options (see [19]).

In Quantitative Finance, several problems of interest are the estimation of quantities like (for a given function $f : \mathbb{R}^d \mapsto \mathbb{R}$)

$$\mathbb{E}\big[f(X_T)\big], \quad T > 0, \tag{1}$$

or involving terms like

$$\mathbb{E}\big[f(X_t)|X_s = x\big], \quad 0 < s < t, \tag{2}$$

where $(X_t)_{t\in[0,T]}$ is a stochastic process which evolves following the stochastic differential equation

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dW_t, \quad X_0 = x_0 \in \mathbb{R}^d, \tag{3}$$

where W is a standard d-dimensional Brownian motion starting at 0 and where the functions $b : [0, T] \times \mathbb{R}^d \mapsto \mathbb{R}^d$ and the matrix diffusion coefficient function $\sigma : [0, T] \times \mathbb{R}^d \mapsto \mathbb{R}^d \times \mathbb{R}^d$ are measurable and satisfy

*Fédération de Mathématiques d'Evry, Laboratoire Analyse et Probabilités, 23 Boulevard de France, 91037 Evry, & ENSIIE, e-mail: abass.sagna@ensiie.fr. This research benefited from the support of the " Chaire Marchés en Mutation", Fédération Bancaire Française.

some appropriate conditions which ensure the existence of a strong solution of the stochastic differential equation. Since in general the solution of (3) is not explicit we have first to approximate the continuous paths of the process $(X_t)_{t\in[0,T]}$ by discrete paths using some discretization scheme like the Euler scheme. Given the (regular) time discretization steps $t_k = k\Delta$, $k = 0, \dots, n$, $\Delta = T/n$, the "discrete Euler process" $(\tilde{X}_{t_k})_k$, $k = 0, \dots, n$, associated to the previous diffusion process $(X_t)_{t\in[0,T]}$ is defined recursively as

$$\tilde{X}_{t_{k+1}} = \tilde{X}_{t_k} + b(t_k, \tilde{X}_{t_k})\Delta + \sigma(t_k, \tilde{X}_{t_k})(W_{t_{k+1}} - W_{t_k}), \quad \tilde{X}_0 = X_0.$$

Then, once we have access to the discrete paths of the stochastic process $(X_t)_{t\in[0,T]}$, the quantities (1) and (2) are estimated by

$$\mathbb{E}\big[f(\tilde{X}_{t_n})\big] \tag{4}$$

and

$$\mathbb{E}\big[f(\tilde{X}_{t_{k+1}})|\tilde{X}_{t_k} = x\big], \quad t_{k+1} = t, t_k = s. \tag{5}$$

Remark 1.1. a) Under some assumptions on f (for e.g. if f is four times continuously differentiable with polynomial growth) the estimation of $\mathbb{E}(f(X_T)$ by $\mathbb{E}(f(\tilde{X}_T))$ induces the following weak error convergence:

$$|\mathbb{E}f(X_T) - \mathbb{E}f(\tilde{X}_T)| \leq \frac{C}{n}$$

with $C > 0$ and where n is the number of time discretization steps.
b) Also remark that for every $p \geq 1$,

$$\mathbb{E}\Big(\sup_{k=0,\dots,n} |\tilde{X}_{t_k}|^p\Big) < +\infty. \tag{6}$$

The estimation of quantities like (4) or (5) can be performed using Monte Carlo simulations. Nevertheless, an alternative to the Monte Carlo method may be the optimal quantization method, specially in small dimension ($d \leq 4$ in the theory but in practice it may stay competitive with respect to the Monte Carlo method up to dimension $d = 10$, see [16]).

In fact, suppose that we have access to the optimal quantization or to some "good" (in a sense to be specified later) quantizations $(\tilde{X}_{t_k}^{x^{N_k}})_k$ of the process $(\tilde{X}_{t_k})_k$ on the grids $x^{N_k} = \{x_1^{N_k}, \dots, x_{N_k}^{N_k}\}$ of size N_k, for $k = 0, \dots, n$. Suppose also that we can get the associated weights $\mathbb{P}(\hat{X}_{t_k}^{x^{N_k}} = x_i^{N_k})$, $i = 1, \dots, N_k$, $k = 0, \dots, n$, and the transition probabilities $\hat{p}_k^j(x) = \mathbb{P}(\hat{X}_{t_{k+1}}^{x^{N_{k+1}}} = x_j^{N_{k+1}}|\hat{X}_{t_k}^{x^{N_k}} = x)$ for every $k = 0, \dots, n-1$. Then using optimal quantization method, the expressions (4) and (5) are estimated respectively by

$$\mathbb{E}\big[f(\hat{X}_{t_n}^{x^{N_n}})\big] = \sum_{i=1}^{N_n} f(x_i^{N_n})\, \mathbb{P}\big(\hat{X}_{t_n}^{x^{N_n}} = x_i^{N_n}\big)$$

and

$$\mathbb{E}\big[f(\hat{X}_{t_{k+1}}^{x^{N_{k+1}}})|\hat{X}_{t_k}^{x^{N_k}} = x\big] = \sum_{j=1}^{N_{k+1}} f(x_j^{N_{k+1}})\, \hat{p}_k^j(x).$$

The question is then to know how to get the optimal grids x^{N_k}, for $k = 0, \dots, n$, the associated weights and transition probabilities. In a more general framework, as soon as the stochastic process $(\tilde{X}_{t_k})_k$ (or the underlying diffusion process $(X_t)_{t\geq 0}$) can be simulated one may use stochastic (or Lloyd) algorithms to estimate the (optimal) grids and the associated weights or transition probabilities. In the special case of the one dimensional setting we can use the Newton algorithm in some situations. This deterministic algorithm leads to more precise estimations and is more fast (at far in some cases) than stochastic algorithms.

Consider the first quantity (4), given that the second one is of interest in a working paper. Suppose for example that we aim to estimate the price of a Put option with a maturity T, a strike K, an interest rate r and

a given present value X_0 in the Black-Scholes model where the dynamics of the stock price process under the risk neutral probability is given by:

$$dX_t = rX_t dt + \sigma X_t dW_t, \quad X_0 = x_0 \in \mathbb{R},$$

which solution reads

$$X_t = X_0\, e^{(r-\sigma^2/2)t+\sigma W_t} \overset{\mathcal{L}}{=} X_0\, e^{(r-\sigma^2/2)t+\sigma\sqrt{t}Z}, \quad Z \sim \mathcal{N}(0;1).$$

Then we want to estimate the quantity (keep in mind that this is a toy example given that it can be written in a semi-closed formula involving the cumulative distribution function of the standard Gaussian distribution: the famous Black-Scholes formula)

$$e^{-rT}\mathbb{E}(f(X_T)),$$

where the payoff function $f(x) = \max(K - x, 0)$. To estimate such an expression by optimal quantization, one may quantize the random variable Z and compute the associated weights using the Newton algorithm. Then, setting $\hat{Z}^{x^N}$ its optimal quantization on the grid $x^N = \{x_1^N, \dots, x_N^N\}$ we estimate the quantity of interest by

$$e^{-rT}\sum_{i=1}^{N} f(g(x_i^N))\mathbb{P}(\hat{Z}^{x^N} = x_i^N)$$

where $g : x \mapsto X_0\, e^{(r-\sigma^2/2)T+\sigma\sqrt{T}x}$ (see e.g. [12]). Now, suppose that the model is a local volatility model where the dynamics of the stock price process evolves following the stochastic differential equation (called Pseudo-CEV in [10]):

$$dX_t = rX_t dt + \vartheta \frac{X_t^{\delta+1}}{\sqrt{1+X_t^2}} dW_t, \quad X_0 = x_0 \tag{7}$$

for some $\delta \in (0,1)$ and $\vartheta \in (0, \underline{\vartheta}]$ with $\underline{\vartheta} > 0$, where r is the interest rate. In this situation the distribution of X_T is not known and if we want to estimate the quantity of interest: $e^{-rT}\mathbb{E}(f(X_T)$, where $f(x) := \max(K - x, 0)$ is the payoff function, we have to discretize the process $(X_t)_{t\geq 0}$ as $(\tilde{X}_{t_k})_{k=0,\dots,n}$, with $t_n = T$, using for e.g. the Euler scheme, then, estimate

$$e^{-rT}\mathbb{E}(f(\tilde{X}_T))$$

by optimal quantization. Up to now, the only way to get the optimal grids and the associated weights in this situation is carrying out stochastic algorithms (see e.g. [16]) or Lloyd's algorithms (see e.g. [8]), even in the one dimensional framework. As pointed out previously this method may be very time consuming.

In this paper we propose a new approach to quantize the process $(\tilde{X}_{t_k})_k$, based on a recursive method involving the conditional distribution of the marginals: means, those of $\tilde{X}_{t_{k+1}}|\tilde{X}_{t_k}$. This approach raises some interesting problems among with the computation of the rate of convergence of the quantization error associated to $\tilde{X}_{t_k}$ (for $k = 0, \dots, n$), the question of knowing if the defined quantization $\hat{X}_{t_k}^{x^{N_k}}$ of $\tilde{X}_{t_k}$ on the grid x^{N_k} satisfy the stationary property, ...

We observe that by construction these marginal quantizations are stationary. This property is very useful when estimating expressions like (4) and (5) by optimal quantization. We also observe that the rate of convergence of the quantization error associated to the marginal $\tilde{X}_{t_k}$ is of $\mathcal{O}(N_k^{-1})$, where N_k is the level of the quantization. As a consequence of this result, if f is a Lipschitz function then estimating $\mathbb{E}f(\tilde{X}_T)$ by $\mathbb{E}f(\hat{X}_T^{N_n})$ where $\hat{X}_T^{N_n}$ is the optimal quantization of $\tilde{X}_T$ of size N_n will induce the following error bound:

$$|\mathbb{E}f(\tilde{X}_T) - \mathbb{E}f(\hat{X}_T^{N_n})| = \mathcal{O}(N_n^{-1}).$$

On the other hand, if $k \in \{0, \dots, n-1\}$, we show that one may rely the L^2-quantization error $\mathbb{E}|\tilde{X}_{t_{k+1}} - \hat{X}_{t_{k+1}}^{x^{N_{k+1}}}|^2$ associated to $\tilde{X}_{t_{k+1}}$ with the distribution of $\tilde{X}_{t_k}$ as

$$\mathbb{E}|\tilde{X}_{t_{k+1}} - \hat{X}_{t_{k+1}}^{x^{N_{k+1}}}|^2 = \int_{\mathbb{R}^d} \mathbb{E}\Big(\min_{j=1,\dots,N_{k+1}} \big|Y_k(x) - x_j^{N_{k+1}}\big|^r\Big)\mathbb{P}(\tilde{X}_{t_k} \in dx) \tag{8}$$

where $Y_k(x)$ is the d-dimensional gaussian distribution with known mean and variance. Thus, in practice, to quantize the random variable $\tilde{X}_{t_{k+1}}$ we have to compute an expectation with respect to $\tilde{X}_{t_k}$. The distribution of $\tilde{X}_{t_k}$ is not explicit but since the random variable $\tilde{X}_{t_k}$ is supposed to be already quantized we will estimate (8) by

$$\sum_{i=1}^{N_k} \mathbb{E}\Big(\min_{j=1,\dots,N_{k+1}} \big|Y_k(x_i^{N_k}) - x_j^{N_{k+1}}\big|^2 \Big) \mathbb{P}(\hat{X}_{t_k}^{x^{N_k}} = x_i^{N_k}) \tag{9}$$

where $\hat{X}_{t_k}^{x^{N_k}}$ is the quantization of $\tilde{X}_{t_k}$ on the grid x^{N_k}. We show that the error induced by such an estimation is bounded by the quantization error $\|X_{t_k} - \hat{X}_{t_k}^{x^{N_k}}\|_2$.

From the numerical viewpoint our method allows us to estimate efficiently, as well from the precision of the estimates as from the computational time, expressions like (4). In fact, in the specific case of the one dimensional setting, this approach allows us to use the Newton algorithm as an alternative to stochastic algorithms. This is a very important fact because it is more precise and more fast than stochastic algorithms. Notice however that when dealing with high dimension ($d \geq 2$) there is at the moment no way to avoid the use of stochastic (or Lloyd) like algorithms.

The paper is organized as follows. We recall first some basic facts about optimal quantization. The marginal quantization method is described in Section 3. We give in this section the induced quantization error and compute the error bound derived from the estimation of (8) by (9). We also illustrate how to get the optimal grids using Newton's algorithm and show how to estimate the associated weights. The last section, Section 4, is devoted to the application of the proposed method to the pricing of an European Put option in a local volatility model (as well as in the Black-Scholes model) and to the comparison with the Monte Carlo method.

2 Overview on optimal quantization methods

Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and let $X : (\Omega, \mathcal{A}, \mathbb{P}) \longrightarrow \mathbb{R}^d$ be a random variable with distribution $\mathbb{P}_X$. The $L^r(\mathbb{P}_X)$-optimal quantization problem of size N for the random vector X (or for the distribution $\mathbb{P}_X$) consists in finding the best approximation of X by a Borel function of X taking at most N values. Assuming that $X \in L^r(\mathbb{P})$, we associated to every Borel function $q(X)$ taking at most N values the L^r-error $\|X - q(X)\|_r$ measuring the distance between the two random vectors X and $q(X)$ w.r.t. the L^r-norm, where $\|X\|_r := (\mathbb{E}|X|^r)^{1/r}$ and where $|\cdot|$ denotes an arbitrary norm on $\mathbb{R}^d$. Then finding the best approximation of X by a Borel function of X taking at most N values turns out to find the solution of the following minimization problem:

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

Now, let $x^N = \{x_1^N, \dots, x_N^N\} \subset \mathbb{R}^d$ be a subset (a codebook) of size N (it is also called an N-quantizer or a grid of size N) and define a Voronoi partition $C_i(x^N)_{i=1,\dots,N}$ of $\mathbb{R}^d$, which is a Borel partition of $\mathbb{R}^d$ satisfying for every $i \in \{1, \dots, N\}$,

$$C_i(x^N) \subset \{x \in \mathbb{R}^d : |x - x_i^N| = \min_{j=1,\dots,N} |x - x_j^N|\}.$$

Consider the Voronoi quantization of X (or simply called quantization of X) on the N-quantizer x^N defined by

$$\hat{X}^{x^N} = \sum_{i=1}^{N} x_i^N \mathbf{1}_{\{X \in C_i(x^N)\}}.$$

Then for any Borel function $q : \mathbb{R}^d \to x^N = \{x_1^N, \dots, x_N^N\}$ we have

$$|X - q(X)| \geq \min_{i=1,\dots,N} d(X, x_i^N) = d(X, x^N) = |X - \hat{X}^{x^N}| \quad \mathbb{P} \text{ a.s}$$

so that the optimal quantization error $e_{N,r}(X)$ reads

$$\begin{aligned} e_{N,r}(X) &= \inf\{\|X - \hat{X}^{x^N}\|_r, x^N \subset \mathbb{R}^d, \text{ card}(x^N) \leq N\} \\ &= \inf_{\substack{x^N \subset \mathbb{R}^d \\ \text{card}(x^N) \leq N}} \left(\int_{\mathbb{R}^d} d(z, x^N)^r d\mathbb{P}_X(z) \right)^{1/r}. \end{aligned} \tag{10}$$

Note that for every $N \geq 1$, the infimum in (10) is reached at one codebook at least. Any N-quantizer realizing this infimum is called an L^r-optimal N-quantizer. Morever, if $\text{card}(\text{supp}(\mathbb{P}_X)) \geq N$ then the optimal N-quantizer is of size exactly N (see [9] or [12]). On the other hand, the quantization error, $e_{N,r}(X)$, decreases to zero as the grid size N goes to infinity and its rate of convergence is ruled by the so-called Zador Theorem given below.

Theorem 2.1. *(**Zador Theorem**, see [9]) : Let X be an $\mathbb{R}^d$-valued random vector such that $\mathbb{E}|X|^{r+\eta} < +\infty$ for some $\eta > 0$ and let $\mathbb{P}_X = P_a + P_s$ be the Lebesgue decomposition of $\mathbb{P}_X$ with respect to the Lebesgue measure λ_d, where P_a denotes the absolutely continuous part and P_s the singular part. Then*

$$\lim_{N \to +\infty} N^{r/d} (e_{N,r}(P))^r = Q_r(\mathbb{P}_X) \tag{11}$$

with

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

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

where $U([0,1]^d)$ denotes the uniform distribution on the set $[0,1]^d$ and $f = \frac{dP_a}{d\lambda_d}$.

We will call $Q_r(\mathbb{P}_X)$ the Zador's constant associated to X. From the Numerical Probability point of view, finding an optimal N-quantizer x^N may be a challenging task. In practice (we will only consider the quadratic case, i.e. when $r = 2$ for numerical implementations) we are sometimes led to find some "good" quantizations $\hat{X}^{x^N}$ which are close to X in distribution, so that for every Borel function $f : \mathbb{R}^d \mapsto \mathbb{R}$, we can approximate $\mathbb{E}[f(X)]$ by

$$\mathbb{E} f\big(\hat{X}^{x^N}\big) = \sum_{i=1}^{N} f(x_i^N)\, p_i, \tag{12}$$

where $p_i = \mathbb{P}(\hat{X}^{x^N} = x_i^N)$. Among "good" quantizations of X we have stationary quantizers defined as follows.

Definition 2.1. An N-quantizer $x^N = \{x_1^N, \ldots, x_N^N\}$ inducing the quantization $\hat{X}^{x^N}$ of X is said stationary if

$$\forall\, i \neq j, \qquad x_i \neq x_j \;\text{ and }\; \mathbb{P}\left(X \in \cup_i \partial C_i(x^N)\right) = 0$$

and

$$\mathbb{E}\left[X | \hat{X}^{x^N}\right] = \hat{X}^{x^N}. \tag{13}$$

Defining the distortion function by

$$D_{N,2}(x^N) = \int_{\mathbb{R}^d} d(z, x^N)^2 d\mathbb{P}_X(z) = \sum_{i=1}^{N} \int_{C_i(x^N)} |z - x_i^N|^2 d\mathbb{P}_X(z), \tag{14}$$

a stationary quantizer $x^N = \{x_1^N, \ldots, x_N^N\}$ is in fact an N-quantizer satisfying the stationary equality: $\nabla D_{N,2}(x^N) = 0$. The following result justify the introversion of the differentiation and the integral leading to (13) when differentiating (14), see [9].

Proposition 2.2. *The function $D_{N,2}$ is differentiable at any N-tuple $x^N \in (\mathbb{R}^d)^N$ having pairwise distinct components and satisfying*

$$\mathbb{P}\left(X \in \cup_i \partial C_i(x^N)\right) = 0$$

(which condition holds in particular when X is a continuous random vector) and we have

$$\nabla D_{N,2}(x^N) = \Big(2\int_{C_i(x^N)} (x_i^N - z) d\mathbb{P}_X(z)\Big)_{i=1,\dots,N}. \tag{15}$$

For numerical implementations, the stationary quantizers search is based on zero search recursive procedures like Newton's algorithm in the one dimensional framework and some algorithms as Lloyd's I algorithms (see e.g. [8]), the Competitive Learning Vector Quantization (CLVQ) algorithm (see [8]) or stochastic algorithms (see [15]) in the multidimensional framework. One may download optimal quantizers associated to some random variables as the multivariate Gaussian random vector in the website `www.quantize.math-fi.com`.

When estimating the expectation $\mathbb{E}f(X)$ by $\mathbb{E}f(\hat{X}^{x^N})$ where x^N is an N-quantizer and $\hat{X}^{x^N}$ the quantization of X on x^N, we make an error which may be bounded by the quantization error $\mathbb{E}|X - \hat{X}^{x^N}|^2$, depending to the regularity of the function f. We next recall some error bounds induced from the approximation of $\mathbb{E}f(X)$ by (12), we refer to [16]) for further details.

a) Let x^N be a stationary quantizer and f be a Borel function on $\mathbb{R}^d$. If f is a convex function then

$$\mathbb{E}f(\hat{X}^{x^N}) \leq \mathbb{E}f(X). \tag{16}$$

b) Lipschitz functions:

- If f is Lipschitz continuous then for any N-quantizer x^N we have

$$\left|\mathbb{E}f(X) - \mathbb{E}f(\hat{X}^{x^N})\right| \leq [f]_{\mathrm{Lip}} \|X - \hat{X}^{x^N}\|_2,$$

where

$$[f]_{\mathrm{Lip}} := \sup_{x \neq y} \frac{|f(x) - f(y)|}{|x - y|}.$$

- Let $\theta : \mathbb{R}^d \to \mathbb{R}_+$ be a nonnegative convex function such that $\theta(X) \in L^2(\mathbb{P})$. If f is locally Lipschitz with at most θ-growth, i.e. $|f(x) - f(y)| \leq [f]_{\mathrm{Lip}}|x - y|(\theta(x) + \theta(y))$ then $f(X) \in L^1(\mathbb{P})$ and

$$\left|\mathbb{E}f(X) - \mathbb{E}f(\hat{X}^{x^N})\right| \leq 2[f]_{\mathrm{Lip}} \|X - \hat{X}^{x^N}\|_2 \|\theta(X)\|_2.$$

c) Differentiable functionals: if f is differentiable on $\mathbb{R}^d$ with an α-Hölder differential $\mathrm{D}f$ ($\alpha \in [0,1]$), then for any stationary N-quantizer x^N,

$$\left|\mathbb{E}f(X) - \mathbb{E}f(\hat{X}^{x^N})\right| \leq [Df]_\alpha \|X - \hat{X}^{x^N}\|_2^{1+\alpha}.$$

3 Marginal quantization of a discrete stochastic diffusion process

Let $(X_t)_{t \geq 0}$ be a stochastic process taking values in d-dimensional Euclidean space $\mathbb{R}^d$ and evolving following the stochastic differential equation:

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dW_t, \quad X_0 = x_0 \in \mathbb{R}^d, \tag{17}$$

where W is a standard d-dimensional Brownian motion starting at 0 and where the functions $b : [0,T] \times \mathbb{R}^d \mapsto \mathbb{R}^d$ and the matrix diffusion coefficient function $\sigma : [0,T] \times \mathbb{R}^d \mapsto \mathbb{R}^d \times \mathbb{R}^d$ are measurable and satisfy the global Lipschitz and linear growth conditions:

$$\begin{aligned} |b(t,x) - b(t,y)| + |\sigma(t,x) - \sigma(t,y)| &\leq K|x-y| \quad (18) \\ |b(t,x)| + |\sigma(t,x)| &\leq K(1+|x|). \quad (19) \end{aligned}$$

This guaranties the existence of a strong solution of (17). We also suppose that the matrix σ is positive definite. In the rest of the paper we will suppose that $\mathbb{R}^d$ is equipped with the Euclidean norm.

Let $t_k = k\Delta, k = 0, 1, \ldots, n$ be time discretization steps, with $\Delta > 0$. Consider the Euler scheme of the process $(X_t)_{t\geq 0}$ starting from $X_0 = x_0$:

$$\tilde{X}_{t_{k+1}} = \tilde{X}_{t_k} + b(t_k, \tilde{X}_{t_k})\Delta + \sigma(t_k, \tilde{X}_{t_k})(W_{t_{k+1}} - W_{t_k}).$$

Define the functions m_k and Σ_k by

$$m_k(x) = x + \Delta b(t_k, x) \qquad \text{and} \qquad \Sigma_k(x) = \Delta\, \sigma(t_k, x)\sigma(t_k, x)^T.$$

It is clear that

$$\tilde{X}_{t_{k+1}} | \tilde{X}_{t_k} = x \sim \mathcal{N}\big(m_k(x); \Sigma_k(x)\big)$$

and will denote by $\Phi_{m_k(x),\Sigma_k(x)}(\cdot)$ its density and by $\Gamma_{m_k(x),\Sigma_k(x)}(\cdot)$ its cumulative distribution function:

$$\Phi_{m_k(x),\Sigma_k(x)}(x_{k+1}) = \frac{1}{(2\pi)^{d/2}|\Sigma_k(x)|^{1/2}} \exp\Big(-\frac{1}{2}\big(x_{k+1} - m_k(x)\big)^T \Sigma_k^{-1}\big(x_{k+1} - m_k(x)\big)\Big)$$

where $|\Sigma_k(x)|$ denote the determinant of $\Sigma_k(x)$ and for $x_{k+1} \in \mathbb{R}^d$,

$$\Gamma_{m_k(x),\Sigma_k(x)}(x_{k+1}) = \int_{]-\infty, x_{k+1}]} \Phi_{m_k(x),\Sigma_k(x)}(y)dy.$$

For every $n \geq 0$, we will denote by $\mathbb{P}_x(\tilde{X}_{t_n} \in dx_n)$ the density of X_{t_n} when starting the process at $X_0 = x$.

3.1 Distortion function of the marginals

To compute the distortion function of the marginals of the Euler diffusion process we need to know the probability distribution of the marginals. We recall how to compute the distribution of the marginal $\tilde{X}_{t_{k+1}}$ with respect to the distribution of $\tilde{X}_{t_k}$.

Using Fubini's theorem we have

$$\begin{aligned} \mathbb{P}(\tilde{X}_{t_{k+1}} \leq x_{k+1}) &= \int_{]-\infty, x_{k+1}]} \int_{\mathbb{R}^d} \mathbb{P}(\tilde{X}_{t_{k+1}} \in dy | \tilde{X}_{t_k} = x_k)\mathbb{P}(\tilde{X}_{t_k} \in dx_k) \\ &= \int_{\mathbb{R}^d} \Gamma_{m_k(x_k),\Sigma_k(x_k)}(x_{k+1})\mathbb{P}(\tilde{X}_{t_k} \in dx_k). \end{aligned}$$

Then

$$\begin{aligned} \mathbb{P}(\tilde{X}_{t_{k+1}} \in dx_{k+1}) &= dx_{k+1}\, \mathbb{E}\Big[\Phi_{m_k(\tilde{X}_{t_k}),\Sigma_k(\tilde{X}_{t_k})}(x_{k+1})\Big] \quad (20) \\ &= dx_{k+1} \int_{\mathbb{R}^d} \Phi_{m_k(x_k),\Sigma_k(x_k)}(x_{k+1})\mathbb{P}(\tilde{X}_{t_k} \in dx_k) \quad (21) \\ &= dx_{k+1} \int_{\mathbb{R}^d} \cdots \int_{\mathbb{R}^d} \prod_{i=0}^{k} \Phi_{m_i(x_i),\Sigma_i(x_i)}(x_{i+1})\mu(dx_0)dx_1 \ldots dx_k \quad (22) \end{aligned}$$

It follows from Equation (21) that the probability density function of the Euler diffusion process may be computed recursively if the distribution of X_0 (supposing that X_0 is random) is known.

On the other hand this result allows us to compute the (quadratic) quantization error of the marginal distribution $\tilde{X}_{t_{k+1}}$ given the distribution of $\tilde{X}_{t_k}$ and then to deduce from a recursive procedure the optimal quantization of $\tilde{X}_{t_{k+1}}$ given the distribution of $\tilde{X}_{t_k}$ (or, from the numerical point of view, given the optimal quantization of $\tilde{X}_{t_k}$ and its associated weights). In fact, let $\hat{X}_0^{x^{N_0}}$ be the quadratic optimal quantization of X_0 over the optimal grid $x^{N_0} = \{x_1^{N_0}, \ldots, x_{N_0}^{N_0}\}$ (if $X_0 = x$ is not random we set $N_0 = 1$ and $\hat{X}_0^{x^{N_0}} = x$). We will deduce the quantization of the marginal random variable $\tilde{X}_{t_1}$ using Equation (21). Once this quantization is perform we will get from the same way the quantization of $\tilde{X}_{t_2}$ given the distribution of $\tilde{X}_{t_1}$, and so on. In the general setting suppose that the quadratic optimal quantizations $\hat{X}_{t_l}^{x^{N_l}}$, $l = 0, \ldots, k$, of the $\tilde{X}_{t_l}$'s over the optimal N_l-quantizers $x^{N_l} = \{x_1^{N_l}, \ldots, x_{N_l}^{N_l}\}$ are given. We propose in the following, a way to deduce the quadratic optimal quantization of $\tilde{X}_{t_{k+1}}$.

Let $k \in \{0, \ldots, n-1\}$ and let $D_{k+1}(x^{N_{k+1}})$ be the distortion function associated to the N_{k+1}-quantizer $x^{N_{k+1}} = \{x_1^{N_{k+1}}, \ldots, x_{N_{k+1}}^{N_{k+1}}\}$. Then it follows from (21) that

$$\begin{aligned} D_{k+1}(x^{N_{k+1}}) &= \sum_{j=1}^{N_{k+1}} \int_{C_j(x^{N_{k+1}})} \big(x_j^{N_{k+1}} - z\big)^2 \mathbb{P}(\tilde{X}_{t_{k+1}} \in dz) \\ &= \int_{\mathbb{R}^d} \Big(\sum_{j=1}^{N_{k+1}} \int_{C_j(x^{N_{k+1}})} \big(x_j^{N_{k+1}} - z\big)^2 \Phi_{m_k(x),\Sigma_k(x)}(z) dz \Big) \mathbb{P}(\tilde{X}_{t_k} \in dx), \end{aligned}$$

so that we have the following result.

Proposition 3.1. *We have*

$$D_{k+1}(x^{N_{k+1}}) = \int_{\mathbb{R}^d} \mathbb{E}\Big(\min_{j=1,\ldots,N_{k+1}} \big|Y_k(x) - x_j^{N_{k+1}}\big|^2 \Big) \mathbb{P}(\tilde{X}_{t_k} \in dx), \tag{23}$$

where $Y_k(x)$ is the d-dimensional gaussian distribution with mean $m_k(x)$ and variance $\Sigma_k(x)$.

Keep in mind that the L^r-quantization error for the marginal random variable $\tilde{X}_{t_{k+1}}$ is defined as

$$e_{N_{k+1},r}(\tilde{X}_{t_{k+1}}) = \inf \big\{ \big\| \tilde{X}_{t_{k+1}} - \hat{X}_{t_{k+1}}^{x^{N_{k+1}}} \big\|_r,\ x^{N_{k+1}} \subset \mathbb{R}^d,\ \text{card}(x^{N_{k+1}}) \leq N_{k+1} \big\}.$$

A natural question arising after our constructive approach is to compute the rate of convergence of the quantization error associated to the marginals.

3.2 Error analysis

We observe that for every time step t_k, the quantization error goes to 0 at the optimal rate $N_k^{-1/d}$ where N_k is the grid size of the optimal N_k-quantizer $x^{N_k} = \{x_1^{N_k}, \ldots, x_{N_k}^{N_k}\}$. In fact, it follows from Equation (22) that $\tilde{X}_{t_k}$ admits a density function for every $k =, 1, \ldots, n$. Furthermore, we deduce from (6) that for every $r \geq 1$ and for every $k = 1, \ldots, n$, $\mathbb{E}|\tilde{X}_{t_k}|^r < +\infty$. Then Zador's theorem applies and we have for every $k = 1, \ldots, n$, for any L^r-optimal quantizer x^{N_k} for X_{t_k},

$$\lim_{N_k \to +\infty} N_k^{r/d} e_{N_k,r}^r(\tilde{X}_{t_k}) = Q_r\big(\mathbb{P}_{\tilde{X}_{t_k}}\big) \tag{24}$$

where $e_{N_k,r}^r(\tilde{X}_{t_k}) := \mathbb{E}\big(|\tilde{X}_{t_k} - \hat{X}_{t_k}^{x^{N_k}}|^r\big)$ is the quantization error associated to the L^r-optimal quantizer x^{N_k} and $Q_r\big(\mathbb{P}_{\tilde{X}_{t_k}}\big)$ is the Zador constant associated to the distribution of $\tilde{X}_{t_k}$.

As mentioning before, the computation of the quadratic optimal quantizers for $\tilde{X}_{t_{k+1}}$ suppose that we have access to the distribution of $\tilde{X}_{t_k}$. In practice, since we suppose that $\tilde{X}_{t_k}$ has already be quantized and that we have access to its associated weights, we will estimate the distortion function D_{k+1} associated to $\tilde{X}_{t_{k+1}}$ and given in (23) by

$$\hat{D}_{k+1}(x^{N_{k+1}}) = \sum_{i=1}^{N_k} \mathbb{E}\Big(\min_{j=1,\dots,N_{k+1}} \big|Y_k(x_i^{N_k}) - x_j^{N_{k+1}}\big|^2\Big)\mathbb{P}(\hat{X}_{t_k}^{x^{N_k}} = x_i^{N_k}). \tag{25}$$

We give in the following an error bound induced by such an approximation.

Theorem 3.2. *Suppose that there is is a non null polynomial function $\mathcal{P}_\alpha(a,b)$ of degree $\alpha \in \mathbb{N}$ of the form*

$$\mathcal{P}_\alpha(a,b) = \sum_{i,j\le\alpha} c_{ij} a^i b^j, \quad c_{i,j} \ge 0,\ i,j \in \mathbb{N} \tag{26}$$

such that for every $x, z \in \mathbb{R}^d$,

$$\sup_{t\in[0,T]} |\sigma\sigma^T(t,x) - \sigma\sigma^T(t,z)| \le |x-z|\mathcal{P}_\alpha(|x|,|z|). \tag{27}$$

Let $n \ge 1$. Then, for any sequence of stationary N_k-quantizer x^{N_k} for $\tilde{X}_{t_k}$ s.t. for every $k = 0,\dots,n$,

$$\big\|\tilde{X}_{t_k} - \widehat{X}_{t_k}^{x^{N_k}}\big\| \longrightarrow 0, \text{ as } N_k \to +\infty$$

we have

$$\big|D_{k+1}(x^{N_{k+1}}) - \hat{D}_{k+1}(x^{N_{k+1}})\big| \le K\,\big\|\tilde{X}_{t_k} - \hat{X}_{t_k}^{x^{N_k}}\big\|_2, \quad \text{for } k = 1,\dots,n-1. \tag{28}$$

for some real constant $K > 0$.

Before dealing with the proof of the theorem let us make some remarks on Assumption (26).

Remark 3.1. Note that Assumption (26) is satisfied when $\sigma\sigma^T$ is uniformly Lipschitz in space, uniformly in time: there is $L > 0$ s.t.

$$\sup_{t\in[0,T],(x,z)\in\mathbb{R}^{2d},x\neq z} \frac{|\sigma\sigma^T(t,x) - \sigma\sigma^T(t,z)|}{|x-z|} \le L.$$

We also note that if σ is symmetric then Assumption (26) holds with $\mathcal{P}_1(a,b) = K^2(2+a+b)$, where K is the constant appearing in (18) and (19). In fact, if σ is symmetric then

$$\begin{aligned} |\sigma\sigma^T(t,x) - \sigma\sigma^T(t,z)| &= |\sigma^2(t,x) - \sigma^2(t,z)| \\ &\le |\sigma(t,x) - \sigma(t,z)|\big(|\sigma(t,x)| + |\sigma(t,z)|\big) \\ &\le K^2|x-z|(2+|x|+|z|), \end{aligned}$$

where the first inequality follows from the Hölder and Minkowski inequalities and the last inequality from (18) and (19). As a consequence, Assumption (26) is satisfied by a large class of model in the one dimensional framework, including the Black-Scholes model and the Pseudo-CEV model.

Let us prove the theorem now.

Proof. We have

$$|D_{k+1}(x^{N_{k+1}}) - \hat{D}_{k+1}(x^{N_{k+1}})| = \big|\mathbb{E}f(\tilde{X}_{t_k}) - \mathbb{E}f(\hat{X}_{t_k}^{x^{N_k}})\big| \le \mathbb{E}\big|f(\tilde{X}_{t_k}) - f(\hat{X}_{t_k}^{x^{N_k}})\big|$$

where

$$f(x) = \mathbb{E}\Big(\min_{j=1,\dots,N_{k+1}} \big|Y_k(x) - x_j^{N_{k+1}}\big|^2 \Big) = \mathbb{E}\Big(\min_{j=1,\dots,N_{k+1}} f_j\big(Y_k(x)\big) \Big),$$

with $f_j\big(Y_k(x)\big) = \big|Y_k(x) - x_j^{N_{k+1}}\big|^2$. It follows that for every $z \in \mathbb{R}^d$,

$$f(x) \le \mathbb{E}\Big(\min_{j=1,\dots,N_{k+1}} \big|f_j(Y_k(x)) - f_j(Y_k(z))\big| \Big) + \mathbb{E}\Big(\min_{j=1,\dots,N_{k+1}} f_j\big(Y_k(z)\big) \Big).$$

or, in other words

$$f(x) - f(z) \le \mathbb{E}\Big(\min_{j=1,\dots,N_{k+1}} \big|f_j(Y_k(x)) - f_j(Y_k(z))\big| \Big).$$

Using a symmetric reasoning gives, for every $x, z \in \mathbb{R}^d$,

$$|f(x) - f(z)| \le \mathbb{E}\Big(\min_{j=1,\dots,N_{k+1}} \big|f_j(Y_k(x)) - f_j(Y_k(z))\big| \Big).$$

Since for every $a, b \in \mathbb{R}^d$, it holds the inequality $|a|^2 - |b|^2 \le |a-b|(|a|+|b|)$, we have for every $j = 1, \dots, N_{k+1}$,

$$\big|f_j(Y_k(x)) - f_j(Y_k(z))\big| \le |Y_k(x) - Y_k(z)|\big(|Y_k(x)| + |Y_k(z)| + 2|x_j^{N_{k+1}}|\big).$$

Using Hölder's inequality and the inequality $(|a|+|b|)^2 \le 2(|a|^2+|b|^2)$ yields, for every $x, z \in \mathbb{R}^d$

$$\big|f(x) - f(z)\big| \le \Big(\mathbb{E}|Y_k(x) - Y_k(z)|^2\Big)^{1/2} \Big(\mathbb{E}|Y_k(x)|^2 + \mathbb{E}|Y_k(z)|^2 + 4 \min_{j=1,\dots,N_{k+1}} |x_j^{N_{k+1}}|^2\Big)^{1/2}.$$

Now, using the uniform Lipschitz assumption on $b(t,x)$ and Assumption (27) leads to

$$\begin{aligned}
\mathbb{E}|Y_k(x) - Y_k(z)|^2 &= \mathbb{E}\big|m_k(x) - m_k(z) + \big(\Sigma_k(x) - \Sigma_k(z)\big)Z\big|^2 \\
&\le K_1\Big(|x-z|^2 + |b(t,x) - b(t,z)|^2 + \mathbb{E}|\big(\Sigma_k(x) - \Sigma_k(z)\big)Z|^2\Big) \\
&\le K_1\Big(|x-z|^2 + |\Sigma_k(x) - \Sigma_k(z)|^2\mathbb{E}|Z|^2\Big) \\
&\le K_1|x-z|^2\mathcal{P}_{2\alpha}(|x|,|z|)
\end{aligned}$$

for some positive and generic real constant K_1 which may vary from line to line and where $\mathcal{P}_{2\alpha}$ is a polynomial function of degree 2α of the form (26). Moreover, using the linear growth assumption (19) on the coefficient b of the diffusion and assumption (27) we get, for every $x \in \mathbb{R}^d$,

$$\begin{aligned}
\mathbb{E}|Y_k(x)|^2 = \mathbb{E}|m_k(x) + \Sigma_k(x)Z|^2 &\le 2(|m_k(x)|^2 + \mathbb{E}|\Sigma_k(x)Z|^2) \\
&\le 2\big(|m_k(x)|^2 + |\Sigma_k(x)|^2\mathbb{E}|Z|^2\big) \\
&\le K\big(|m_k(x)|^2 + |\sigma(t_k,x)|^2\big) \\
&\le K(1 + |x| + |x|^2)
\end{aligned}$$

where K is a generic positive constant depending in particular on Δ.

On the other hand, since $\mathbb{E}\big(d(\tilde{X}_{t_{k+1}}, x^{N_{k+1}})^2\big) \to 0$ as $N_{k+1} \to +\infty$, the following asymptotic density property of $(x^{N_{k+1}})$ in the support of the distribution of $\tilde{X}_{t_{k+1}}$ holds :

$$\forall \varepsilon > 0,\ \forall\, x \in \mathrm{supp}(\mathbb{P}_{\tilde{X}_{t_{k+1}}}),\ \exists\, n_{\varepsilon,x,k} \in \mathbb{N},\ \forall\, N_{k+1} \ge n_{\varepsilon,x,k}, \quad B(x,\varepsilon) \cap x^{N_{k+1}} \neq \emptyset. \tag{29}$$

Otherwise, there exists $x \in \mathrm{supp}(\mathbb{P}_{\tilde{X}_{t_{k+1}}}), \varepsilon > 0$ and a subsequence $(x^{N_{k+1}^{(p)}})_{p\ge1}$ so that $\forall p \ge 1, B(x,\varepsilon) \cap x^{N_{k+1}^{(p)}} = \emptyset$. Then, for every $p \ge 1$,

$$\big\|d(\tilde{X}_{t_{k+1}}, x^{N_{k+1}^{(p)}})\big\|_2 \ge \big\|d(\tilde{X}_{t_{k+1}}, x^{N_{k+1}^{(p)}})\mathbf{1}_{\tilde{X}_{t_{k+1}} \in B(x,\varepsilon/2)}\big\|_2 \ge \frac{\varepsilon}{2}\mathbb{P}_{\tilde{X}_{t_{k+1}}}(B(x,\varepsilon/2))^{1/2} > 0$$

which contradicts the fact that $\big\|d(\tilde{X}_{t_{k+1}}, x^{N_{k+1}})\big\|_2 \to 0$ as $N_{k+1} \to +\infty$.

Then, it follows from (29) that there is $C > 0$ such that

$$\min_{j=1,\dots,N_{k+1}} |x_j^{N_{k+1}}| \leq C.$$

Combining the previous results gives for every $x, z \in \mathbb{R}^d$

$$|f(x) - f(z)| \leq K|x-z|\Big((1+|x|+|z|+|x|^2+|z|^2)\mathcal{P}_{2\alpha}(|x|,|z|)\Big)^{1/2} = K|x-z|\mathcal{P}_\beta(|x|,|z|)^{1/2},$$

for a positive real constant K and for $\mathcal{P}_\beta$, a polynomial function of degree β of the form (26). So, using once again Hölder inequality yields

$$\mathbb{E}\big|f(\tilde{X}_{t_k}) - f(\hat{X}_{t_k}^{x^{N_k}})\big| \;\leq\; K\big\|\tilde{X}_{t_k} - \hat{X}_{t_k}^{x^{N_k}}\big\|_2\Big(\mathbb{E}\big[\mathcal{P}_\beta(|\tilde{X}_{t_k}|, |\hat{X}_{t_k}^{x^{N_k}}|)\big]\Big)^{1/2}.$$

Now, we know that for every $r \in \mathbb{N}$, $\mathbb{E}|\tilde{X}_{t_k}|^r < +\infty$. Moreover, since $\hat{X}_{t_k}^{x^{N_k}}$ is stationary, it follows from the previous statement and from inequality (16) that for every $r \in \mathbb{N}$,

$$\mathbb{E}|\hat{X}_{t_k}^{x^{N_k}}|^r \leq \mathbb{E}|\tilde{X}_{t_k}|^r < +\infty.$$

Hence

$$\mathbb{E}\big[\mathcal{P}_\beta(|\tilde{X}_{t_k}|, |\hat{X}_{t_k}^{x^{N_k}}|)\big] \leq \sum_{i,j \leq 2\alpha} c_{ij}\big(\mathbb{E}|\tilde{X}_{t_k}|^{2i}\big)^{1/2}\big(\mathbb{E}|\hat{X}_{t_k}^{x^{N_k}}|^{2j}\big)^{1/2} < +\infty.$$

Finally we have shown that

$$\mathbb{E}\big|f(\tilde{X}_{t_k}) - f(\hat{X}_{t_k}^{x^{N_k}})\big| \;\leq\; K\big\|\tilde{X}_{t_k} - \hat{X}_{t_k}^{x^{N_k}}\big\|_2,$$

for some positive real constant K. □

We focus now on the numerical computation of the quadratic optimal quantizers of the marginal random variable $\tilde{X}_{t_{k+1}}$ given the probability distribution function of $\tilde{X}_{t_k}$. Such a task requires the use of some algorithms like stochastic algorithms, Lloyd's algorithms (both requiring the computation of the gradient of the distortion function) or Newton's algorithm (specially for the one-dimensional setting) which involves the gradient and the hessian matrix of the distortion (we refer to [16] for more details).

3.3 How to get optimal quantizers of the marginals?

Owing to Proposition 2.2 and Equation (22)), the distortion $D_{k+1}(x^{N_{k+1}})$ is continuously differentiable as a function of the N_{k+1}-quantizer $x^{N_{k+1}}$ (having pairwise distinct components) and its gradient is given by

$$\begin{aligned}\nabla D_{k+1}(x^{N_{k+1}}) &= \Big(2\int_{C_j(x^{N_{k+1}})}(x_j^{N_{k+1}} - y)\mathbb{P}(\tilde{X}_{t_{k+1}} \in dy)\Big)_{j=1,\dots,N_{k+1}} \\ &= \Big(2\int_{\mathbb{R}^d}\int_{C_j(x^{N_{k+1}})}(x_j^{N_{k+1}} - y)\mathbb{P}(Y_k(x) \in dy)\mathbb{P}(\tilde{X}_{t_k} \in dx)\Big)_{j=1,\dots,N_{k+1}}\end{aligned}$$

where $Y_k(x)$ is the d-dimensional gaussian distribution with mean $m_k(x)$ and variance $\Sigma_k(x)$. Keep in mind that the quadratic optimal quantizer at level N_{k+1} (that we still denote by $x^{N_{k+1}}$) of $X_{t_{k+1}}$ is defined as

$$x^{N_{k+1}} \in \arg\min\{D_{k+1}(z), z \subset \mathbb{R}^d, \mathrm{card}(z) \leq N_{k+1}\}. \tag{30}$$

Remark 3.2. *a)* If $x^{N_{k+1}}$ is a quadratic optimal N_{k+1}-quantizer of $\tilde{X}_{t_{k+1}}$ and if $\hat{X}_{t_{k+1}}^{x^{N_{k+1}}}$ denotes the quantization of $\tilde{X}_{t_{k+1}}$ over the grid $x^{N_{k+1}}$. Then $x^{N_{k+1}}$ is a stationary quantizer: it is such that

$$\nabla D_{k+1}(x^{N_{k+1}}) = 0.$$

Equivalently, we have for every $j = 1, \ldots, N_{k+1}$,

$$\begin{aligned} x_j^{N_{k+1}} &= \frac{\int_{\mathbb{R}^d} \mathbb{E}\big(Y_{t_k}(x)\mathbf{1}_{\{Y_{t_k}(x)\in C_j(x^{N_{k+1}})\}}\big)\mathbb{P}(\tilde{X}_{t_k}\in dx)}{\int_{\mathbb{R}^d}\mathbb{P}\big(Y_{t_k}(x)\in C_j(x^{N_{k+1}})\big)\mathbb{P}(\tilde{X}_{t_k}\in dx)} \\ &= \frac{\mathbb{E}\big(\tilde{X}_{t_{k+1}}\mathbf{1}_{\{\tilde{X}_{t_{k+1}}\in C_j(x^{N_{k+1}})\}}\big)}{\mathbb{P}\big(\tilde{X}_{t_{k+1}}\in C_j(x^{N_{k+1}})\big)}, \end{aligned} \tag{31}$$

where as previously $Y_k(x)$ is the d-dimensional gaussian distribution with mean $m_k(x)$ and variance $\Sigma_k(x)$. This also means that

$$\mathbb{E}\Big(\tilde{X}_{t_{k+1}}\big|\hat{X}_{t_{k+1}}^{x^{N_{k+1}}}\Big) = \hat{X}_{t_{k+1}}^{x^{N_{k+1}}}.$$

b) In the d-dimensional ($d \geq 2$) setting, Equation (31) allows us to compute stationary quantizers for $\tilde{X}_{t_{k+1}}$ given the distribution of $\tilde{X}_{t_k}$ using Lloyd's type algorithms.

As mentioned in [16], the usual stochastic or Lloyd's companion algorithms become quickly intractable when the dimension d of the random vector $\tilde{X}_{t_k}$ is greater or equal to 2 due to the fact that we have to compute d-dimension integrals on Voronoi cells. Moreover, in our setting the complexity of the algorithms will increase in hight dimension since we have to compute additional d-dimensional integrals. For these reasons, we will restrict our analysis to the one-dimensional setting where we will use the Newton algorithm to perform recursively quadratic optimal quantizers of the marginals $\tilde{X}_{t_k}$ given the distribution of X_0.

3.3.1 The Newton algorithm

Let $\nabla D_{k+1}(x^{N_{k+1}})$ and $\nabla^2 D_{k+1}(x^{N_{k+1}})$ denote respectively the gradient vector and the hessian matrix of the distortion function. Using the Newton algorithm, a zero of the gradient may be computed via the following recursive procedure starting from a given $x^{N_{k+1},0} \in \mathbb{R}^{N_{k+1}}$:

$$x^{N_{k+1},n+1} = x^{N_{k+1},n} - \big(\nabla^2 D_{k+1}(x^{N_{k+1},n})\big)^{-1}\nabla D_{k+1}(x^{N_{k+1},n}). \tag{32}$$

So we have to compute the gradient and the inverse of the hessian of the distortion function. To simplify notations set

$$x_{j-1/2}^{N_{k+1}} = \frac{x_j^{N_{k+1}} + x_{j-1}^{N_{k+1}}}{2},\ x_{j+1/2}^{N_{k+1}} = \frac{x_j^{N_{k+1}} + x_{j+1}^{N_{k+1}}}{2},\ \text{ with } x_{1/2}^{N_{k+1}} = -\infty, x_{N_{k+1}+1/2}^{N_{k+1}} = +\infty,$$

and let

$$x_{k+1,j-}(x) := \frac{x_{j-1/2}^{N_{k+1}} - m_k(x)}{v_k(x)} \quad \text{and} \quad x_{k+1,j+}(x) := \frac{x_{j+1/2}^{N_{k+1}} - m_k(x)}{v_k(x)},$$

where $v_k(x) = \Delta\sigma^2(t_k, x)$. The components of the gradient can be computed after some elementary computations and we have for every $j = 1, \ldots, N_{k+1}$,

$$\begin{aligned} \frac{\partial D_{k+1}(x^{N_{k+1}})}{\partial x_j^{N_{k+1}}} &= \int_{\mathbb{R}} \Big\{\big(x_j^{N_{k+1}} - m_k(x)\big)\Big(\Gamma_{0,1}(x_{k+1,j+}(x)) - \Gamma_{0,1}(x_{k+1,j-}(x))\Big) \\ &\qquad + v_k(x)\big(\Phi_{0,1}(x_{k+1,j+}(x)) - \Phi_{0,1}(x_{k+1,j-}(x))\big)\Big\}\mathbb{P}(\tilde{X}_{t_k}\in dx). \end{aligned}$$

The diagonal terms of the hessian matrix are given by:

$$\frac{\partial^2 D_{k+1}(x^{N_{k+1}})}{\partial^2 x_j^{N_{k+1}}} = \int_{\mathbb{R}} \Big[\Gamma_{0,1}(x_{k+1,j+}(x)) - \Gamma_{0,1}(x_{k+1,j-}(x)) - \frac{1}{4v_k(x)}\Phi_{0,1}(x_{k+1,j+}(x))(x_{j+1}^{N_{k+1}} - x_j^{N_{k+1}}) - \frac{1}{4v_k(x)}\Phi_{0,1}(x_{k+1,j-}(x))(x_j^{N_{k+1}} - x_{j-1}^{N_{k+1}})\Big]\mathbb{P}(\tilde{X}_{t_k} \in dx).$$

The sub-diagonal terms are

$$\frac{\partial^2 D_{k+1}(x^{N_{k+1}})}{\partial x_j^{N_{k+1}} \partial x_{j-1}^{N_{k+1}}} = -\frac{1}{4}\int_{\mathbb{R}} \frac{1}{v_k(x)}(x_j^{N_{k+1}} - x_{j-1}^{N_{k+1}})\Phi_{0,1}(x_{k+1,j-}(x))\mathbb{P}(\tilde{X}_{t_k} \in dx)$$

and the super-diagonals are

$$\frac{\partial^2 D_{k+1}(x^{N_{k+1}})}{\partial x_j^{N_{k+1}} \partial x_{j+1}^{N_{k+1}}} = -\frac{1}{4}\int_{\mathbb{R}} \frac{1}{v_k(x)}(x_{j+1}^{N_{k+1}} - x_j^{N_{k+1}})\Phi_{0,1}(x_{k+1,j+}(x))\mathbb{P}(\tilde{X}_{t_k} \in dx).$$

Note that for practical implementations we have to estimate the integral terms with respect to the distribution of $\tilde{X}_{t_k}$ appearing in the gradient and the hessian of the distortion function. Since our procedure is recursive and we have supposed that $\tilde{X}_{t_k}$ has already been quantized and that the associated weights: $\mathbb{P}(\tilde{X}_{t_k} \in C_i(x^{N_k})), i = 1, \ldots, N_k$ are accessible, we may estimate them by optimal quantization. The Newton procedure (32) will then be modified as

$$\hat{x}^{N_{k+1},n+1} = \hat{x}^{N_{k+1},n} - \big(\nabla^2 \hat{D}_{k+1}(\hat{x}^{N_{k+1},n})\big)^{-1}\nabla \hat{D}_{k+1}(\hat{x}^{N_{k+1},n}) \tag{33}$$

where the components of the modified gradient $\hat{D}_{k+1}(\hat{x}^{N_{k+1}})$ are given for every $j = 1, \ldots, N_{k+1}$ by

$$\frac{\partial \hat{D}_{k+1}(\hat{x}^{N_{k+1}})}{\partial \hat{x}_j^{N_{k+1}}} = \sum_{i=1}^{N_k} \Big\{ \big(\hat{x}_j^{N_{k+1}} - m_k(\hat{x}_i^{N_k})\big)\Big(\Gamma_{0,1}(\hat{x}_{k+1,j+}(\hat{x}_i^{N_k})) - \Gamma_{0,1}(\hat{x}_{k+1,j-}(\hat{x}_i^{N_k}))\Big) + v_k(\hat{x}_i^{N_k})\big(\Phi_{0,1}(\hat{x}_{k+1,j+}(\hat{x}_i^{N_k})) - \Phi_{0,1}(\hat{x}_{k+1,j-}(\hat{x}_i^{N_k}))\big)\Big\}\mathbb{P}(\tilde{X}_{t_k} \in C_i(\hat{x}^{N_k})).$$

The diagonal terms of the hessian matrix $\nabla^2 \hat{D}_{k+1}(\hat{x}^{N_{k+1},n})$ are given by:

$$\frac{\partial^2 \hat{D}_{k+1}(\hat{x}^{N_{k+1}})}{\partial^2 \hat{x}_j^{N_{k+1}}} = \sum_{i=1}^{N_k} \Big[\Gamma_{0,1}(\hat{x}_{k+1,j+}(\hat{x}_i^{N_k})) - \Gamma_{0,1}(\hat{x}_{k+1,j-}(\hat{x}_i^{N_k})) - \frac{1}{4v_k(\hat{x}_i^{N_k})}\Phi_{0,1}(\hat{x}_{k+1,j+}(\hat{x}_i^{N_k}))(\hat{x}_{j+1}^{N_{k+1}} - \hat{x}_j^{N_{k+1}}) - \frac{1}{4v_k(\hat{x}_i^{N_k})}\Phi_{0,1}(\hat{x}_{k+1,j-}(\hat{x}_i^{N_k}))(\hat{x}_j^{N_{k+1}} - \hat{x}_{j-1}^{N_{k+1}})\Big]\mathbb{P}(\tilde{X}_{t_k} \in C_i(\hat{x}^{N_k}))$$

and its sub-diagonal terms are

$$\frac{\partial^2 \hat{D}_{k+1}(\hat{x}^{N_{k+1}})}{\partial \hat{x}_j^{N_{k+1}} \partial \hat{x}_{j-1}^{N_{k+1}}} = -\frac{1}{4}\sum_{i=1}^{N_k} \frac{1}{v_k(\hat{x}_i^{N_k})}(\hat{x}_j^{N_{k+1}} - \hat{x}_{j-1}^{N_{k+1}})\Phi_{0,1}(\hat{x}_{k+1,j-}(\hat{x}_i^{N_k}))\mathbb{P}(\tilde{X}_{t_k} \in C_i(\hat{x}^{N_k})).$$

The super-diagonals terms are

$$\frac{\partial^2 \hat{D}_{k+1}(\hat{x}^{N_{k+1}})}{\partial \hat{x}_j^{N_{k+1}} \partial \hat{x}_{j+1}^{N_{k+1}}} = -\frac{1}{4}\sum_{i=1}^{N_k} \frac{1}{v_k(\hat{x}_i^{N_k})}(\hat{x}_{j+1}^{N_{k+1}} - \hat{x}_j^{N_{k+1}})\Phi_{0,1}(\hat{x}_{k+1,j+}(\hat{x}_i^{N_k}))\mathbb{P}(\tilde{X}_{t_k} \in C_i(\hat{x}^{N_k})).$$

The $\hat{x}^{N_{k+1}}$'s (and its companions $\hat{x}^{N_{k+1}}_{j-1}$, $\hat{x}^{N_{k+1}}_{j+1}$, ...) are the modified N_{k+1}-quantizers induced by the approximation of the distribution of the $\tilde{X}_{t_k}$'s by the distribution of the $\hat{X}^{x^{N_k}}_{t_k}$'s in the computation of the gradient and the hessian matrix of the Newton algorithm steps.

A similar idea combining (vector or functional) optimal quantization with Newton-Raphson zero search procedure is used in [7] in variance reduction context as an alternative and robust method to simulation based recursive importance sampling procedure to estimate the optimal change of measure. Furthermore, the convergence of the modified Newton algorithm to the optimal quantizer is shown in the framework of [7] to be bounded by the quantization error. However, the tools used to show it do not apply directly to our context and the investigation of the convergence of our modified Newton algorithm to an optimal quantizer is a open question. To perform it, we have to show, among others, that the optimal quantizers of marginals are unique. This seems to be not obvious to show even if we may reasonably think that the unicity holds.

Once we have access to the quadratic optimal quantizers x^{N_k} of the marginals $\tilde{X}_{t_k}$, for $k = 0, \ldots, n$ (which are estimated using the Newton algorithm described previously) we have to compute the associated weights $\mathbb{P}(\tilde{X}_{t_k} \in C_j(x^{N_k}))$, $j = 1 \ldots, N_k$, for $k = 0, \ldots, n$. We show in the next result how to estimate them.

Proposition 3.3. *Let $x^{N_{k+1}}$ be the quadratic optimal quantizer of the marginal random variable $\tilde{X}_{t_{k+1}}$. Given the quadratic optimal quantizer x^{N_k} of $\tilde{X}_{t_k}$ and the associated weights $\mathbb{P}(\tilde{X}_{t_k} \in C_i(x^{N_k}))$, $i = 1, \ldots, N_k$, the probability $\mathbb{P}(\tilde{X}_{t_{k+1}} \in C_j(x^{N_{k+1}}))$ is approximated for every $j = 1, \ldots, N_{k+1}$ by*

$$\mathbb{P}\big(\tilde{X}_{t_{k+1}} \in C_j(x^{N_{k+1}})\big) \approx \sum_{i=1}^{N_k} \Big(\Gamma_{0,1}(x_{k+1,j+}(x_i^{N_k})) - \Gamma_{0,1}(x_{k+1,j-}(x_i^{N_k}))\Big)\mathbb{P}\big(\tilde{X}_{t_k} \in C_i(x^{N_k})\big). \tag{34}$$

Proof. For every $k \in \{1, \ldots, n-1\}$ and for every $j = 1, \ldots, N_{k+1}$, we have

$$\mathbb{P}\big(\tilde{X}_{t_{k+1}} \in C_j(x^{N_{k+1}})\big) = \sum_{i=1}^{N_k} \mathbb{P}\big(\tilde{X}_{t_{k+1}} \in C_j(x^{N_{k+1}}); \tilde{X}_{t_k} \in C_i(x^{N_k})\big) \tag{35}$$

and

$$\begin{aligned}\mathbb{P}\big(\tilde{X}_{t_{k+1}} \in C_j(x^{N_{k+1}}); \tilde{X}_{t_k} \in C_i(x^{N_k})\big) =& \mathbb{P}\big(\tilde{X}_{t_k} \le x^{N_{k+1}}_{j+1/2}; \tilde{X}_{t_k} \in C_i(x^{N_k})\big) \\ &- \mathbb{P}\big(\tilde{X}_{t_k+1} \le x^{N_{k+1}}_{j-1/2}; \tilde{X}_{t_k} \in C_i(x^{N_k})\big).\end{aligned} \tag{36}$$

Now we have for every $z \in \mathbb{R}$,

$$\begin{aligned}\mathbb{P}\big(\tilde{X}_{t_{k+1}} \le z; \tilde{X}_{t_k} \in C_i(x^{N_k})\big) &= \int_{-\infty}^{z} \Big(\int_{C_i(x^{N_k})} \mathbb{P}(\tilde{X}_{t_{k+1}} \in dx | \tilde{X}_{t_k} = y)\mathbb{P}(\tilde{X}_{t_k} \in dy)\Big) \\ &= \int_{C_i(x^{N_k})} \Gamma_{m_k(y), v_k(y)}(z)\mathbb{P}(\tilde{X}_{t_k} \in dy) \\ &\approx \Gamma_{m_k(x_i^{N_k}), v_k(x_i^{N_k})}(z)\mathbb{P}\big(\tilde{X}_{t_k} \in C_i(x^{N_k})\big),\end{aligned}$$

which gives the announced result replacing the quantity $\mathbb{P}\big(\tilde{X}_{t_{k+1}} \in C_j(x^{N_{k+1}}); \tilde{X}_{t_k} \in C_i(x^{N_k})\big)$ in (35) by its approximation in (36). □

In the next section we propose an application of the proposed method to the pricing of European options in a local volatility model.

4 Pricing of European options in a local volatility model

4.1 The model

We consider a pseudo-CEV model (see e.g. [10]) where the dynamics of the stock price process is ruled by the following SDE (under the risk neutral probability)

$$dX_t = rX_t dt + \vartheta \frac{X_t^{\delta+1}}{\sqrt{1+X_t^2}} dW_t, \quad X_0 = x_0 \tag{37}$$

for some $\delta \in (0,1)$ and $\vartheta \in (0, \underline{\vartheta}]$ with $\underline{\vartheta} > 0$. The parameter r stands for the interest rate and $\sigma(x) := \vartheta \frac{x^\delta}{\sqrt{1+x^2}}$ corresponds to the local volatility function. This model becomes very close to the CEV model, specially when the initial value of the stock process X_0 is large enough. In this case the local volatility $\sigma(x) \approx \vartheta x^{\delta-1}$.

	ϑ	MC (10^5)	MC (10^6)	MQ
	0.5	0.0022	0.0018	0.0017
CI		$[\mathbf{0.0017}; \mathbf{0.0028}]$	$[\mathbf{0.0017}; \mathbf{0.0019}]$	
	0.6	0.0113	0.0111	0.0110
CI		$[\mathbf{0.0101}; \mathbf{0.0125}]$	$[\mathbf{0.0107}; \mathbf{0.0115}]$	
	0.7	0.0377	0.0373	0.0370
CI		$[\mathbf{0.0353}; \mathbf{0.0401}]$	$[\mathbf{0.0366}; \mathbf{0.0381}]$	
	0.8	0.0883	0.0876	0.0871
CI		$[\mathbf{0.0843}; \mathbf{0.0923}]$	$[\mathbf{0.0863}; \mathbf{0.0886}]$	
	0.9	0.1696	0.1659	0.1649
CI		$[\mathbf{0.1635}; \mathbf{0.1756}]$	$[\mathbf{0.1640}; \mathbf{0.1678}]$	
	1.0	0.267	0.271	0.271
CI		$[\mathbf{0.259}; \mathbf{0.275}]$	$[\mathbf{0.269}; \mathbf{0.274}]$	
	2.0	2.423	2.433	2.426
CI		$[\mathbf{2.387}; \mathbf{2.459}]$	$[\mathbf{2.422}; \mathbf{2.445}]$	
	3.0	5.424	5.492	5.478
CI		$[\mathbf{5.424}; \mathbf{5.512}]$	$[\mathbf{5.471}; \mathbf{5.512}]$	
	4.0	8.893	8.806	8.808
CI		$[\mathbf{8.801}; \mathbf{8.986}]$	$[\mathbf{8.777}; \mathbf{8.835}]$	

Table 1: (Pseudo-CEV model) Comparison of the Put prices obtained from Monte Carlo (MC) simulations (followed by the size of the MC in brackets) with associated confidence intervals (CI) and from the marginal quantization (MQ) method. The parameters are: $r = 0.15$; $\delta = 0.5$; $n = 120$; $N_k = 400$, $\forall k = 1, \ldots, n$; $T = 1$; $K = 100$; $X_0 = 100$; and for varying values of ϑ.

We aim to compute the price of a European Put option with payoff $(K - X_T)^+ = \max(K - X_T, 0)$, where K corresponds to the strike of the option and T to its maturity. Then we have to approximate the quantity

$$e^{-rT}\mathbb{E}(K - X_T)^+$$

where $\mathbb{E}$ stands for the expectation under the risk neutral probability. If the process $(\tilde{X}_{t_k})_k$ denotes the discrete Euler process at regular time discretization steps t_k, with $0 = t_0 < \ldots < t_n = T$, associated to the diffusion process $(X_t)_{t \geq 0}$, this turns out to estimate

$$e^{-rT}\mathbb{E}(K - \tilde{X}_T)^+$$

by optimal quantization. We estimate this quantity by the marginal quantization method proposed in this paper and compare the numerical results to those obtained from Monte Carlo simulations.

K	MC (10^5)	MC (10^6)	MC (10^7)	MQ
100	08.89	08.81	08.81	08.81
CI	$[\mathbf{08.80}; \mathbf{08.99}]$	$[\mathbf{08.78}; \mathbf{08.84}]$	$[\mathbf{08.80}; \mathbf{08.82}]$	
105	10.61	10.60	10.59	10.59
CI	$[\mathbf{10.51}; \mathbf{10.72}]$	$[\mathbf{10.57}; \mathbf{10.63}]$	$[\mathbf{10.58}; \mathbf{10.60}]$	
110	12.53	12.57	12.57	12.57
CI	$[\mathbf{12.42}; \mathbf{12.64}]$	$[\mathbf{12.53}; \mathbf{12.60}]$	$[\mathbf{12.56}; \mathbf{12.58}]$	
115	14.72	14.74	14.75	14.75
CI	$[\mathbf{14.60}; \mathbf{14.84}]$	$[\mathbf{14.70}; \mathbf{14.78}]$	$[\mathbf{14.75}; \mathbf{14.77}]$	
120	17.18	17.10	17.13	17.12
CI	$[\mathbf{17.04}; \mathbf{17.31}]$	$[\mathbf{17.06}; \mathbf{17.15}]$	$[\mathbf{17.11}; \mathbf{17.14}]$	
125	19.64	19.69	19.67	19.67
CI	$[\mathbf{19.50}; \mathbf{19.78}]$	$[\mathbf{19.64}; \mathbf{19.73}]$	$[\mathbf{19.65}; \mathbf{19.68}]$	
130	22.41	22.32	22.40	22.40
CI	$[\mathbf{22.26}; \mathbf{22.56}]$	$[\mathbf{22.32}; \mathbf{22.41}]$	$[\mathbf{22.38}; \mathbf{22.41}]$	

Table 2: (Pseudo-CEV model) Comparison of the Put prices obtained from Monte Carlo (MC) simulations (followed by the size of the MC in brackets) with associated confidence intervals (CI) and from the marginal quantization (MQ) method. The parameters are: $r = 0.15$; $n = 120$; $N_k = 400$, $\forall k = 1, \ldots, n$; $T = 1$; $\vartheta = 4$; $X_0 = 100$; and for varying values of K.

4.1.1 Numerical results

To deal with numerical examples we set $\delta = 0.5$, $X_0 = 100$, and choose the interest rate $r = 0.15$. We discretize the price process using the Euler scheme with $n = 120$ (regular) discretization steps and quantize the Euler marginal processes by our proposed method. We put all the marginal quantization grid sizes N_k equals to 400 except for $\hat{X}_0^{x_0} = X_0 = 100$ which grid size is $N_0 = 1$. We estimate the price of the Put option by

$$\mathbb{E}\big[\big(K - \hat{X}_{t_n}^{x^{N_n}}\big)^+\big] = \sum_{i=1}^{N_n} (K - x_i^{N_n})^+ \, \mathbb{P}\big(\hat{X}_{t_n}^{x^{N_n}} = x_i^{N_n}\big) \tag{38}$$

where $t_n = T$, and where $x^{N_n} = \{x_1^{N_n}, \ldots, x_{N_n}^{N_n}\}$ is the optimal quantizer of size N_n computed from the Newton algorithm (with 5 iterations) and where the associated weight are estimated from (34).

We compare the prices obtained from the marginal quantization (MQ) method with those obtained by the Monte Carlo (MC) simulations even for various values of ϑ with a fixed strike $K = 100$ (see Table 1) or for varying the values of the strike K with a fixed $\vartheta = 4$ (see Table 2). For the Monte Carlo simulations we set the sample size M equal to 10^5 and 10^6 for $K = 100$ and we set it equal to 10^5, 10^6 and 10^7 when making varying the strike K.

Remark 4.1. (*On the computation time*) a) Remark that all the quantization grids x^{N_k} of sizes $N_k = 400$, for every $k = 1, \ldots, n = 120$, and there associated weights are obtained in about 1 minute from the Newton algorithm with 5 iterations. Computations are performed using *Scilab* software on a CPU 2.7 GHz and 4 Go memory computer.

b) With the same grid sizes as in a), all the optimal grids and the associated weights may be obtained quite instantaneously from the Newton algorithm using the *C (or C++) programming language*. In this case we can compute directly the inverse of the hessian matrix using the results in [21].

c) It is clear that once the grids and the associated weights are available the estimation of the price by MQ method using the sum (38) is instantanuous.

Remark 4.2. (*Initialization of the Newton algorithm*) Let $0 = t_0 < \ldots < t_n$ be the time discretization steps, let $X_0 = x$ be the present value of the stock price process and suppose that the grid sizes N_k are equal. Since

the random variable $\tilde{X}_{t_1} \sim \mathcal{N}(m_0(x); v_0^2(x))$, in order to compute the (optimal) N_1-quantizer for $\tilde{X}_{t_1}$ we initialize the algorithm to $v_0(x)z^{N_1} + m_0(x)$, where z^{N_1} is the optimal N_1-quantizer of the $\mathcal{N}(0;1)$. Once we get the optimal N_1-quantization x^{N_1} for $\tilde{X}_{t_1}$ and its associated weights, we initialize the algorithm to x^{N_1} to perform the optimal N_2-quantizer for $\tilde{X}_{t_2}$ and its associated weights, ..., and so on, until we get the optimal N_n-quantizer for $\tilde{X}_{t_n}$ and the associated weights. Notice that doing so we observe no failure of convergence in all the considered examples.

Remark 4.3. We show in Figure 1 and Figure 2 two graphics where we depict in the abscissa axis the optimal grids (of sizes $N_k = 150$) and in ordinate axis the corresponding weights. The dynamics of the price process in Figure 1 is given by

$$dX_t = rX_t dt + \sigma X_t dW_t, \quad X_0 = 86.3$$

with $r = 0.03$, $\sigma = 0.05$ whereas its dynamics in Figure 2 is given by

$$dX_t = rX_t dt + \vartheta \frac{X_t^{\delta+1}}{\sqrt{1+X_t^2}} dW_t, \quad X_0 = 100$$

with $r = 0.15$, $\vartheta = 0.7$, $\delta = 0.5$. We observe that the forms of the distributions are close to the Gaussian distribution as indicated by the Gaussian concentration bounds for some processes (see e.g. [11]). This suggests that the Gaussian concentration results may exist for this model, at least for some choice of the parameter δ.

Let us come back to the estimation of the Put price. Since the payoff function $f(x) = (K-x)^+$ is Lipschitz and since the quantization $\hat{X}_{t_n}^{x^{N_n}}$ is stationary, it follows from Zador Theorem that

$$\left| \mathbb{E}f(X_{t_n}) - \mathbb{E}f(\hat{X}_{t_n}^{x^{N_n}}) \right| = \mathcal{O}(N_n^{-1}).$$

Keeping in mind that the estimation error induced by the MC method satisfies

$$\left| \mathbb{E}f(X_{t_n}) - \frac{1}{M}\sum_{i=1}^{M} f(X_{t_n}^i) \right| = \mathcal{O}(M^{-1/2})$$

for a sample $X_{t_n}^1, \ldots, X_{t_n}^M$ of size M, we deduce that the MQ method provides a precision as good as the MC method when $N_n \approx \sqrt{M}$. For our numerical examples, we remark first that in all examples the prices obtained by MQ stay in the confidence interval induced by the MC price estimates. On the other hand the prices obtained by the MQ method are more precise (more specifically when $\vartheta = 4$ and K increasing away from 100) than those obtained by the MC method when the sample size M equals 10^5 or 10^6. Consequently, the MQ method seems to be more efficient than the MC when the sample size is less than 10^6. However, when increasing the sample size to $M = 10^7$ the two prices because closer (up to 10^{-2}); despite everything the confidence interval associated to the MC method stays large.

Remark 4.4. We remark that when the Monte Carlo sample sample size $M = 10^7$ it takes about 2 minutes and 30 seconds to get a price using the *C programming language* on the same computer described previously. Then, in this situation, it takes more time to obtain a price by MC method than carrying it out by MQ. Furthermore, the MQ method would become much more fast when programming the Newton algorithm using the *C programming language* instead of the *Scilab* software.

To strengthen the previous conclusions related to the local volatility model we compare the two methods in the Black-Scholes framework where the stock price process evolves following the dynamics:

$$dX_t = rX_t dt + \sigma X_t dW_t, \quad X_0 = 100.$$

In this setting the true prices are available and will serve us as the support of comparison. The parameters are chosen such that the model is close to the Pseudo-CEV model: $r = 0.15$ and $\sigma \approx \vartheta X_0^{\delta-1}$. Numerical results

are printed in Tables 3 and Table 4 and heighten our conclusions on the Pseudo-CEV model. We notice that in the Black-Scholes method, the estimated prices from the MQ method are close to the true prices (the best absolute error is of order 10^{-5} for a volatility $\sigma = 5\%$ and the worse absolute error 2.10^{-2} is achieved in high volatility: $\sigma = 40\%$). This show the robustness of the MQ method even for reasonably high values of the volatility.

	σ	MC (10^5)	MC (10^6)	MQ	True price	Abs. error
	0.05	0.0015	0.00178	0.00176	0.00177	$\mathbf{1.10^{-5}}$
CI		$[\mathbf{0.0012}; \mathbf{0.0019}]$	$[\mathbf{0.0017}; \mathbf{0.0019}]$			
	0.06	0.0116	0.0109	0.0109	0.0112	$\mathbf{3.10^{-4}}$
CI		$[\mathbf{0.0104}; \mathbf{0.0128}]$	$[\mathbf{0.0106}; \mathbf{0.0113}]$			
	0.07	0.0365	0.0370	0.0369	0.0373	$\mathbf{4.10^{-4}}$
CI		$[\mathbf{0.0342}; \mathbf{0.0387}]$	$[\mathbf{0.0363}; \mathbf{0.0378}]$			
	0.08	0.0876	0.0876	0.0869	0.0875	$\mathbf{6.10^{-4}}$
CI		$[\mathbf{0.0836}; \mathbf{0.0915}]$	$[\mathbf{0.0863}; \mathbf{0.0888}]$			
	0.09	0.1666	0.1644	0.1647	0.1654	$\mathbf{7.10^{-4}}$
CI		$[\mathbf{0.1607}; \mathbf{0.1724}]$	$[\mathbf{0.1622}; \mathbf{0.1658}]$			
	0.10	0.269	0.271	0.271	0.272	$\mathbf{1.10^{-3}}$
CI		$[\mathbf{0.261}; \mathbf{0.277}]$	$[\mathbf{0.271}; \mathbf{0.273}]$			
	0.20	2.444	2.431	2.424	2.427	$\mathbf{3.10^{-3}}$
CI		$[\mathbf{2.410}; \mathbf{2.479}]$	$[\mathbf{2.420}; \mathbf{2.442}]$			
	0.30	5.455	5.469	5.470	5.474	$\mathbf{4.10^{-3}}$
CI		$[\mathbf{5.395}; \mathbf{5.515}]$	$[\mathbf{5.450}; \mathbf{5.549}]$			
	0.40	8.680	8.787	8.790	8.792	$\mathbf{2.10^{-3}}$
CI		$[\mathbf{8.598}; \mathbf{8.763}]$	$[\mathbf{8.760}; \mathbf{8.813}]$			

Table 3: (Black-Scholes model) Comparison of the Put prices obtained from Monte Carlo (MC) simulations (followed by the size of the MC in brackets) with associated confidence intervals (CI) and from the marginal quantization (MQ) method with the associated absolute error (Abs. error) w. r. t. the true price. The parameters are: $r = 0.15$; $n = 120$; $N_k = 400$, $\forall k = 1, \ldots, n$; $T = 1$; $K = 100$; $X_0 = 100$; and for varying values of σ.

	K	MC (10^5)	MC (10^6)	MQ	True price	Abs. error
	100	8.680	8.787	8.790	8.792	$\mathbf{2.10^{-3}}$
CI		$[\mathbf{8.598}; \mathbf{8.763}]$	$[\mathbf{8.760}; \mathbf{8.813}]$			
	105	10.805	10.739	10.744	10.750	$\mathbf{6.10^{-3}}$
CI		$[\mathbf{10.71}; \mathbf{10.90}]$	$[\mathbf{10.71}; \mathbf{10.90}]$			
	110	12.86	12.89	12.90	12.91	$\mathbf{1.10^{-2}}$
CI		$[\mathbf{12.76}; \mathbf{12.96}]$	$[\mathbf{12.86}; \mathbf{12.93}]$			
	115	15.29	15.24	15.26	15.27	$\mathbf{1.10^{-2}}$
CI		$[\mathbf{15.18}; \mathbf{15.40}]$	$[\mathbf{15.21}; \mathbf{15.28}]$			
	120	17.66	17.81	17.79	17.81	$\mathbf{1.10^{-2}}$
CI		$[\mathbf{17.54}; \mathbf{17.79}]$	$[\mathbf{17.78}; \mathbf{17.85}]$			
	125	20.56	20.50	20.50	20.52	$\mathbf{1.10^{-2}}$
CI		$[\mathbf{20.43}; \mathbf{20.69}]$	$[\mathbf{20.46}; \mathbf{20.54}]$			
	130	23.28	23.37	23.37	23.39	$\mathbf{2.10^{-2}}$
CI		$[\mathbf{23.14}; \mathbf{23.42}]$	$[\mathbf{23.34}; \mathbf{23.43}]$			

Table 4: (Black-Scholes model) Comparison of the Put prices obtained from Monte Carlo (MC) simulations (followed by the size of the MC in brackets) with associated confidence intervals (CI) and from the marginal quantization (MQ) method with the associated absolute error (Abs. error) w. r. t. the true price. The parameters are: $r = 0.15$; $n = 120$; $N_k = 400$, $\forall k = 1, \ldots, n$; $T = 1$; $\sigma = 40\%$; $X_0 = 100$; and for varying values of K.

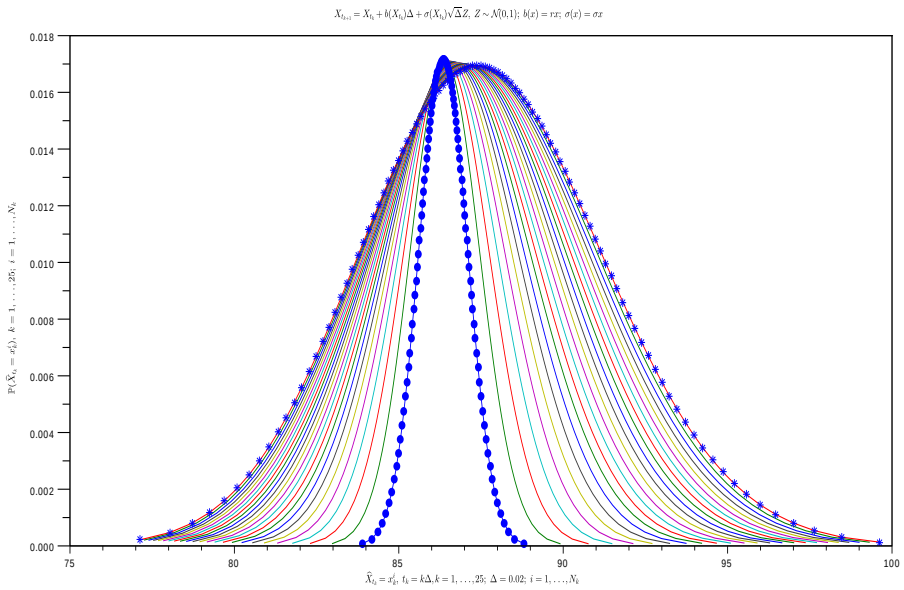


Figure 1: ("Black-Scholes model ") $dX_t = rX_t dt + \sigma X_t dW_t$, $X_0 = 86.3$, $r = 0.03$, $\sigma = 0.05$. Abscissa axis: the optimal grids, $\hat{X}_{t_k} = x_k^i$, $t_k = k\Delta$, $\Delta = 0.02$, $k = 1, \ldots, 25$, $i = 1, \ldots, N_k$. Ordinate axis: the associated weights, $\mathbb{P}(\hat{X}_{t_k} = x_k^i)$, $k = 1, \ldots, 25$, $i = 1, \ldots, N_k$. $\hat{X}_{t_1}$ is depicted in dots '•', $\hat{X}_{t_{25}}$ is represented by the symbol '*', $t_1 = 0.02$ and $t_{25} = 0.5$ and the remaining in continuous line

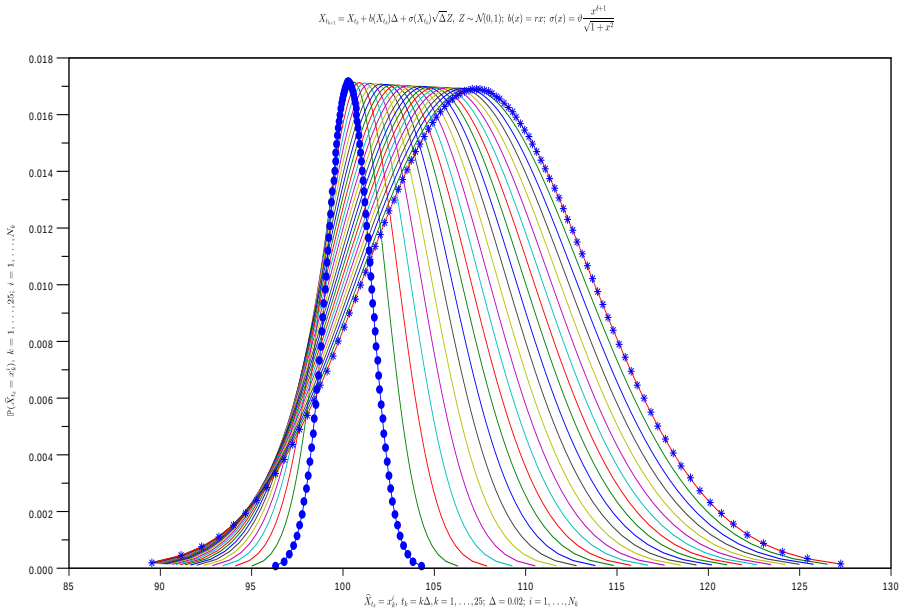


Figure 2: ("Pseudo-CEV model") $dX_t = rX_t dt + \vartheta(X_t^{\delta+1}/(1+X_t^2)^{-1/2})dW_t$, $X_0 = 100$, $r = 0.15$, $\vartheta = 0.7$, $\delta = 0.5$. Abscissa axis: the optimal grids, $\hat{X}_{t_k} = x_k^i$, $t_k = k\Delta$, $\Delta = 0.02$, $k = 1,\ldots,25$, $i = 1,\ldots,N_k$. Ordinate axis: the associated weights. $\hat{X}_{t_1}$ is depicted in dots '•', $\hat{X}_{t_{25}}$ is represented by the symbol '*', $t_1 = 0.02$ and $t_{25} = 0.5$ and the remaining in continuous line.

References

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

[2] V. Bally, G. Pagès and J. Printems. A quantization tree method for pricing and hedging multidimensional American options, *Mathematical Finance*, 15(1), 119-168, 2005.

[3] O. Bardou, S. Bouthemy and G. Pagès. Optimal quantization for pricing of swing options. *Applied Mathematical Finance*, 16(2), 183-217, 2009.

[4] J. A. Bucklew and G. L. Wise. Multidimensional asymptotic quantization theory with r-th power distribution measures. *IEEE Trans. Inform. Theory*, 28:239-247, 1982.

[5] G. Callegaro, A. Sagna. An application to credit risk of optimal quantization methods for non- linear filtering. *The Journal of Computational Finance*. To appear.

[6] M. Corsi, H. Pham and W. Runggaldier. Numerical approximation by quantization of control problems in finance under partial observations, to appear in *Mathematical modeling and numerical methods in finance, special volume of Handbook of Numerical analysis*, 2009.

[7] N. Frikha and A. Sagna. Quantization based recursive Importance Sampling. *Monte Carlo Methods and Applications*, 18, 287-326, 2012.

[8] A. Gersho and R. Gray. *Vector Quantization and Signal Compression*. Kluwer Academic Press, Boston, 1992.

[9] S. Graf and H. Luschgy. *Foundations of Quantization for Probability Distributions*. Lect. Notes in Math., 1730, 2000. Berlin: Springer.

[10] V. Lemaire and G. Pagès. Unconstrained Recursive Importance Sampling, *Annals of Applied Probability*, 20, 1029-1067, 2010.

[11] V. Lemaire and S. Menozzi. On some non asymptotic bounds for the Euler Scheme, *Electronic Journal of Probability*, 15, 1645-1681, 2010.

[12] G. Pagès. A space quantization method for numerical integration. *Journal of Computational and Applied Mathematics*, 89:1-38, 1998.

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

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

[15] G. Pagès and J. Printems. Optimal quadratic quantization for numerics: the gaussian case. *Monte Carlo Methods and Applications*, 9(2):135-165, 2003.

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

[17] H. Pham, W. Runggaldier and A. Sellami. Approximation by quantization of the filter process and applications to optimal stopping problems under partial observation, *Monte Carlo Methods and Applications*, 11, 57-82, 2005.

[18] C. Profeta and A. Sagna. Conditional hitting time estimation in a nonlinear filtering model by the Brownian bridge method. Preprint. Submitted work.

[19] A. Sagna. Pricing of barrier options by marginal functional quantization method. *Monte Carlo Methods and Applications*, 17(4), 2012.

[20] W. F. Sheppard. On the Calculation of the Most Probable Values of Frequency-Constants, for Data Arranged According to Equidistant Divisions of Scale, Proc. London Math. Soc. 29: 353Ð380, 1898.

[21] R. Usmani, Inversion of a tridiagonal Jacobi matrix, *Linear Algebra Appl.*, 212/213, 413-414, 1994.

[22] P. Zador. Asymptotic quantization error of continuous signals and the quantization dimension. *IEEE Trans. Inform. Theory*, 28:139-149, 1982.