



**HAL**  
open science

# Improved error bounds for quantization based numerical schemes for BSDE and nonlinear filtering

Gilles Pagès, Abass Sagna

► **To cite this version:**

Gilles Pagès, Abass Sagna. Improved error bounds for quantization based numerical schemes for BSDE and nonlinear filtering. 2015. hal-01211285v1

**HAL Id: hal-01211285**

**<https://hal.science/hal-01211285v1>**

Preprint submitted on 4 Oct 2015 (v1), last revised 19 Jul 2017 (v3)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Improved error bounds for quantization based numerical schemes for BSDE and nonlinear filtering

GILLES PAGÈS \* ABASS SAGNA † ‡

## Abstract

We take advantage of recent results on optimal quantization theory (see [24, 44]) to improve the quadratic optimal quantization error bounds for backward stochastic differential equations (BSDE in short) and nonlinear filtering problems. To achieve it, one of the main ideas used for both problems is the use of the orthogonality property of the conditional expectation for the mean-quadratic norm. When permitting some involving functions to be less regular than what is usually needed, the analysis of the nonlinear filtering error bounds brings into play the so-called mismatch property, namely the fact that the quadratic optimal quantizers of size  $N$  used to approximate  $\mathbb{R}^d$ -valued random vectors in  $L^2$  by a nearest neighbor projection (Voronoi quantization) at a  $N^{-\frac{1}{d}}$  still perform this approximation at the same rate in  $L^s$ ,  $2 \leq s \leq 2 + d$ .

## 1 Introduction

In this work we propose improved error bounds for quantization based numerical schemes introduced in [4] and [41] to solve BSDEs and nonlinear filtering problems. For BSDE, we consider equations where the driver depends on the “ $Z$ ” term (see Equation (1) below) and for nonlinear filtering, we extend existing results to locally Lipschitz continuous densities (see Section 5). For both problems, we also improve the error bounds themselves by using a Pythagoras like theorem for the approximation of conditional expectations introduced in [44] (see also [40]). These problems have a wide range of applications, in particular in Financial Mathematics, when modeling the price of financial derivatives or in stochastic control, in credit risk modeling, etc.

BSDEs were first introduced in [9] but raised a wide interest mostly after the extension in work [47]. In this latter paper, the existence and the uniqueness of a solution have been established for the following backward stochastic differential equation with Lipschitz continuous driver  $f$  (valued in  $\mathbb{R}^d$ ) and terminal condition  $\xi$ :

$$Y_t = \xi + \int_t^T f(s, Y_s, Z_s) ds - \int_t^T Z_s dW_s, \quad 0 \leq t \leq T, \quad (1)$$

where  $W$  is a  $q$ -dimensional brownian motion. We mean by a solution a pair  $(Y_t, Z_t)_{t \leq T}$  (valued in  $\mathbb{R}^d \times \mathbb{R}^{d \times q}$ ) of square integrable progressively measurable (with respect to the augmented Brownian

---

\*Laboratoire de Probabilités et Modèles aléatoires (LPMA), UPMC-Sorbonne Université, UMR 7599, case 188, 4, pl. Jussieu, F-75252 Paris Cedex 5, France. E-mail: gilles.pages@upmc.fr

†Laboratoire de Mathématiques et Modélisation d’Evry (LaMME), UMR 8071, 23 Boulevard de France, 91037 Evry, & ENSIIE. E-mail: abass.sagna@ensiie.fr.

‡The first author benefited from the support of the Chaire “Risques financier”, a joint initiative of École Polytechnique, ENPC-ParisTech and UPMC, under the aegis of the Fondation du Risque. The second author has benefited from the support of the Chaire “Markets in Transition”, under the aegis of Louis Bachelier Laboratory, a joint initiative of École polytechnique, Université d’Évry Val d’Essonne and Fédération Bancaire Française.

filtration  $(\mathcal{F}_t)_{t \geq 0}$ ) and satisfying Equation (1). Extensions of these existence and uniqueness results have been investigated in more general situations (less regular drivers  $f$  (see [19] for driver having a little regularity in time, called rough path driver, [1, 26] for locally Lipschitz driver, [32] for quadratic BSDEs and [33] for superlinear quadratic BSDEs), randomized horizon (see [46]), introduction of Poisson random measure component subject to constraints on the jump component (see [31, 30]), extension to second order BSDEs (see [37])).

Since the pioneering work [38] in which the link between BSDE and hedging portfolio of European (and American) derivatives has been first established, various other applications have been developed, as risk-sensitive control problems, risk measure theory, etc.

However, even if it can be established in many cases that a BSDE has a unique solution, this solution admits no closed form in general. This led to devise tractable approximation schemes of the solution. In the Markovian case (see (2) below) for example, where the terminal condition is of the form  $\xi = h(X_T)$  for some forward diffusion  $X$ , a first numerical method has been proposed in [20] for a class of forward-backward stochastic differential equations, based a four step scheme developed later on in [35].

In [51], a numerical scheme for BSDEs with possible path-dependent terminal condition has been investigated. Many others approximation methods of a solution of some classes of BSDEs such as coupled BSDE, Reflected BSDE, BSDE for quasilinear PDEs, BSDE applied to control problems or nonlinear PDEs, etc, have also be considered (we refer for e.g. to [2, 13, 14, 18]). Note in fact that in [14], the authors consider a slightly modified usual dynamical programming equation to propose a numerical approximation of (1) when the generator  $f$  has a quadratic growth with respect to  $z$ . They investigate the time discretization error and use optimal quantization to implement their algorithm. However, they do not study the induced quantization error.

In the present work, we consider the following Markovian BSDE

$$Y_t = \xi + \int_t^T f(s, X_s, Y_s, Z_s) ds - \int_t^T Z_s \cdot dW_s, \quad t \in [0, T], \quad (2)$$

where  $W$  is a  $q$ -dimensional Brownian motion,  $(Z_t)_{t \in [0, T]}$  is a square integrable progressively measurable process taking values in  $\mathbb{R}^q$ ,  $f : [0, T] \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^q \rightarrow \mathbb{R}$ . We suppose a terminal condition of the form  $\xi = h(X_T)$ , for a given Borel function  $h : \mathbb{R}^d \rightarrow \mathbb{R}$ , where  $X_T$  is the value at time  $T$  of a Brownian diffusion process  $(X_t)_{t \geq 0}$ , strong solution to the stochastic differential equation:

$$X_t = x + \int_0^t b(s, X_s) ds + \int_0^t \sigma(s, X_s) dW_s, \quad x \in \mathbb{R}^d. \quad (3)$$

In this case, the approximating methods of the solution of the BSDE are written (for a given time discretization at instants  $t_0 = 0, \dots, t_n = T$ ) as a functional of the paths of  $(X_{t_k})_{k=0, \dots, n}$  and involve in particular conditional expectations  $\mathbb{E}(g_{k+1}(X_{t_{k+1}}) | X_{t_k})$ , where  $g_{k+1}$  is a known function. The sequence  $(X_{t_{k+1}})_{0 \leq k \leq n}$  is either a ‘‘sampling’’ of the diffusion  $X$  at times  $(t_k)_{0 \leq k \leq n}$  or, most often, a discretization scheme of  $(X_t)_{t \geq 0}$ , typically the Euler scheme, when the solution of (3) is not explicit enough to be simulated in an exact way.

In this paper, we will consider an explicit time discretization scheme, recursively defined in a backward way as:

$$\tilde{Y}_{t_n} = h(\bar{X}_{t_n}) \quad (4)$$

$$\tilde{Y}_{t_k} = \mathbb{E}(\tilde{Y}_{t_{k+1}} | \mathcal{F}_{t_k}) + \Delta_n f(t_k, \bar{X}_{t_k}, \mathbb{E}(\tilde{Y}_{t_{k+1}} | \mathcal{F}_{t_k}), \tilde{\zeta}_{t_k}), \quad k = 0, \dots, n-1, \quad (5)$$

with

$$\tilde{\zeta}_{t_k} = \frac{1}{\Delta_n} \mathbb{E}(\tilde{Y}_{t_{k+1}} (W_{t_{k+1}} - W_{t_k}) | \mathcal{F}_{t_k}). \quad (6)$$

The process  $(\bar{X}_{t_k})_{k=0,\dots,n}$  is the discrete time Euler scheme of the diffusion process  $(X_t)_{t \in [0,T]}$  with step  $\Delta_n = \frac{T}{n}$ . It is defined recursively by

$$\bar{X}_{t_k} = \bar{X}_{t_{k-1}} + \Delta_n b(t_{k-1}, \bar{X}_{t_{k-1}}) + \sigma(t_{k-1}, \bar{X}_{t_{k-1}})(W_{t_k} - W_{t_{k-1}}), \quad k = 1, \dots, n, \quad \bar{X}_0 = x.$$

Under some smooth assumptions on the coefficients of the diffusions one shows (see Theorem 3.1 further on for a precise statement, see also [10]) that there is a real constant  $\tilde{C}_{b,\sigma,f,T} > 0$  such that, for every  $n \geq 1$ ,

$$\max_{k \in \{0, \dots, n\}} \mathbb{E}|Y_{t_k} - \tilde{Y}_{t_k}|^2 + \int_0^T \mathbb{E}|Z_t - \tilde{Z}_t|^2 dt \leq \tilde{C}_{b,\sigma,f,T} \Delta_n,$$

where  $\tilde{Z} = \tilde{Z}^{(n)}$  comes from the martingale representation of  $\sum_{k=1}^n \tilde{Y}_{t_k} - \mathbb{E}(\tilde{Y}_{t_k} | \mathcal{F}_{t_{k-1}})$ . A proof of this result is provided in the appendix for self completeness.

At this stage, since the scheme (4)-(5) involves the computation of conditional expectations for which no analytical expression is available, its solution  $(\tilde{Y}, \tilde{\zeta})$  has in turn to be approximated. A possible approach is to rely on regression methods involving the Monte Carlo simulations (see *e.g.* [10, 22]). Other method using on line Monte Carlo simulations has been developed in a Malliavin calculus framework (conditional expectations are “regularized” by integration by parts from which “Malliavin” weights come out, see [10, 16, 27]).

In this paper, we propose to use an *optimal quantization tree* approach originally introduced in [5] and developed in [4, 3, 6] (in fact for Reflected BSDEs). This approach is based on an optimally fitting approximation of the Markovian dynamics of the discrete time Markov chain  $(\bar{X}_{t_k})_{0 \leq k \leq n}$  (or a sampling of  $X$  at discrete times  $(t_k)_{k=0,\dots,n}$ ) with random variables having a finite support. More precisely, the quantization tree is defined recursively by mimicking (4)-(5) as follows:

$$\hat{Y}_{t_n} = h(\hat{X}_{t_n}) \quad (7)$$

$$\hat{Y}_{t_k} = \hat{\mathbb{E}}_k(\hat{Y}_{t_{k+1}}) + \Delta_n f(t_k, \hat{X}_{t_k}, \hat{\mathbb{E}}_k(\hat{Y}_{t_{k+1}}), \hat{\zeta}_{t_k}) \quad (8)$$

$$\text{with} \quad \hat{\zeta}_{t_k} = \frac{1}{\Delta_n} \hat{\mathbb{E}}_k(\hat{Y}_{t_{k+1}} \Delta W_{t_{k+1}}), \quad k = 0, \dots, n-1,$$

where  $\Delta W_{t_{k+1}} = W_{t_{k+1}} - W_{t_k}$ ,  $\hat{\mathbb{E}}_k = \mathbb{E}(\cdot | \hat{X}_{t_k})$ , and  $\hat{X}_{t_k}$  is a quantization of  $\bar{X}_{t_k}$  on a finite grid  $\Gamma_k$ , *i.e.*,  $\hat{X}_{t_k} = \pi_k(\bar{X}_{t_k})$ , where  $\pi_k : \mathbb{R}^d \rightarrow \Gamma_k$  are Borel functions,  $k = 0, \dots, n$ . Our aim is to include the  $Z$  term in the driver and to significantly improve the error bounds in [3, 6].

So, the question of interest will be to estimate the quadratic quantization error  $(\mathbb{E}|\tilde{Y}_{t_k} - \hat{Y}_{t_k}|^2)^{1/2}$  induced by the approximation of  $\tilde{Y}_{t_k}$  by  $\hat{Y}_{t_k}$ , for every  $k = 0, \dots, n$ , where  $\hat{Y}_{t_k}$  is the quantized version of  $\tilde{Y}_{t_k}$  given by (7)-(8). Under more general assumptions than [4, 5], we show in Theorem 3.2 that at every step  $k$  of the procedure,

$$\|\tilde{Y}_{t_k} - \hat{Y}_{t_k}\|_2^2 \leq \sum_{i=k}^n \tilde{K}_i \|\bar{X}_{t_i} - \hat{X}_{t_i}\|_2^2, \quad (9)$$

for positive real constants  $\tilde{K}_i$  depending on  $T$  and on the regularity of the coefficients of  $b, \sigma$  and the driver  $f$ . The presence of the squared norms in (9) improve the control of the time discretization effect, compared with [4, 5]. In fact, we switch from a global error of order  $n \times \max_{0 \leq k \leq n} \|\bar{X}_{t_k} - \hat{X}_{t_k}\|_2$  to an order  $\sqrt{n} \max_{0 \leq k \leq n} \|\bar{X}_{t_k} - \hat{X}_{t_k}\|_2$  without damaging the constants.

In the second part of the paper, we consider a (discrete time) nonlinear filtering problem and improve (in the quadratic setting) the results obtained in [41], where the error bounds associated to the estimation of a discrete filter by optimal quantization has been investigated.

The discrete time nonlinear filtering problem involves two processes: the signal process and the observation process. The signal process denoted in this paper by  $(X_k)_{k \geq 0}$  is modeled by an  $\mathbb{R}^d$ -valued discrete time random process whereas the observation process  $(Y_k)_{k \geq 0}$  is an  $\mathbb{R}^q$ -valued discrete time random process, both defined on a probability space  $(\Omega, \mathcal{A}, \mathbb{P})$ . We assume that both sequences  $(X_k)_{0 \leq k \leq n}$  and  $(X_k, Y_k)_{0 \leq k \leq n}$  share the Markov property. The nonlinear filtering problem consists in computing the conditional distribution  $\Pi_{y,n}$  of  $X_n$  given the observation  $(y_0, \dots, y_n) = (y_0, \dots, y_n)$ . If we assume that the distribution of  $Y_k$  given  $(X_{k-1}, Y_{k-1}, X_k)$  has a density given by  $y' \mapsto g_k(X_{k-1}, Y_{k-1}, X_k, y')$  for every  $k \in \{1, \dots, n\}$ , then it is well-known (see e.g. [41]) that, for every bounded Borel function  $f : \mathbb{R}^d \rightarrow \mathbb{R}$ ,

$$\Pi_{y,n}f = \frac{\pi_{y,n}f}{\pi_{y,n}\mathbf{1}} \quad (10)$$

where  $\pi_{y,n}f$  is computed by the following recursive formula:

$$\pi_{y,k}f = \pi_{y,k-1}H_{y,k}f, \quad k = 1, \dots, n, \quad \pi_{y,0} = H_{y,0},$$

with

$$H_{y,k}f(x) = \mathbb{E}(f(X_k)g_k(x, y_{k-1}, X_k, y_k) | X_{k-1} = x)$$

and

$$H_{y,0}f(x) = \mathbb{E}(f(X_0)).$$

In spite of the previous recursive formula,  $\Pi_{y,n}f$  cannot be computed explicitly in general and has to be approximated using numerical approximation methods. One efficient method to solve numerically this problem is to rely again on an optimal quantization scheme as proposed in [41]. It consists in quantizing the process  $(X_k)_{k=0, \dots, n}$  by a process  $(\widehat{X}_k)$  where, for every  $k = 0, \dots, n$ ,  $\widehat{X}_k$  is a marginal quantization of  $X_k$  of size  $N_k$ . The idea at this stage is to substitute in the formula which defines the above kernels  $H_{y,k}f(x)$  the variables  $X_k$  by their marginal quantizations  $\widehat{X}_k$ . We denote by  $\widehat{\Pi}_{y,n}$  the resulting approximation of the conditional distribution  $\Pi_{y,n}$ . It has been shown in [41] that for every bounded function  $f$ , the absolute error  $|\Pi_{y,n}f - \widehat{\Pi}_{y,n}f|$  is bounded (up to a positive constant) by the cumulated quantization errors  $\|X_k - \widehat{X}_k\|_r$  from  $k = 0$  to  $n$ , for  $r \geq 1$  (in practice, for every  $r = 1, 2$ ).

We propose to improve this result in the quadratic framework ( $r = 2$ ) in two ways. We show that for every bounded function  $f$ , the square-absolute error  $|\Pi_{y,n}f - \widehat{\Pi}_{y,n}f|^2$  is bounded by the sum of the square-quadratic quantization errors  $\|X_k - \widehat{X}_k\|_2^2$  from  $k = 0$  to  $n$ . Secondly, taking advantage of recent new results on optimal vector quantization, namely, the distortion mismatch problem (see [24] and Theorem 4.3 Appendix B), we show the same improved result (asymptotically with respect to the quantization grid sizes  $N_k$ ) under local Lipschitz instead of global Lipschitz continuity assumption on the density functions  $g_k$ .

The paper is divided into two parts. The first part is devoted to the analysis of the optimal quantization error associated to the BSDE of consideration. We recall first, in Section 2, the discretization scheme we consider for the BSDE. Then, in Section 3, we investigate error analysis for the time discretization and the quantization scheme. We perform some numerical tests in dimension one and illustrate the rate of convergence of the quantization error for a Call price in the Black-Scholes model. Then, some results about optimal quantization are recalled in Section 4. Section 5 is devoted to the nonlinear filtering problem analysis when estimating the nonlinear filter by optimal quantization.

NOTATIONS:  $\bullet \|\cdot\|$  denotes the canonical Euclidean norm on  $\mathbb{R}^d$ .

- For every  $f : \mathbb{R}^d \rightarrow \mathbb{R}$ , set  $[f]_{\text{Lip}} = \sup_{x \neq y} \frac{|f(x) - f(y)|}{|x - y|} \leq +\infty$ .
- If  $A \in \mathcal{M}(d, q)$  we define the norm  $\|A\| = \sqrt{\text{Tr}(AA^*)}$ .
- For every  $r \geq 0$ , we define  $L^{r+}(\mathbb{P}) = \bigcup_{\eta > 0} L^{r+\eta}(\mathbb{P})$ .

## 2 Discretization of the BSDE

Let  $(W_t)_{t \geq 0}$  be a  $q$ -dimensional Brownian motion defined on a probability space  $(\Omega, \mathcal{A}, \mathbb{P})$  and let  $(\mathcal{F}_t)_{t \geq 0}$  be its augmented natural filtration. We consider the following stochastic differential equation:

$$X_t = x + \int_0^t b(s, X_s) ds + \int_0^t \sigma(s, X_s) dW_s, \quad (11)$$

where the drift coefficient  $b : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$  and the matrix diffusion coefficient  $\sigma : [0, T] \times \mathbb{R}^d \rightarrow \mathcal{M}(d, q)$  are Lipschitz continuous in  $(t, x)$ . For a fixed horizon (the maturity)  $T > 0$ , we consider the following Markovian Backward Stochastic Differential Equation (BSDE):

$$Y_t = h(X_T) + \int_t^T f(s, X_s, Y_s, Z_s) ds - \int_t^T Z_s \cdot dW_s, \quad t \in [0, T], \quad (12)$$

where the function  $h : \mathbb{R}^d \rightarrow \mathbb{R}$  is  $[h]_{\text{Lip}}$ -Lipschitz continuous, the driver  $f(t, x, y, z) : [0, T] \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^q \rightarrow \mathbb{R}$  is Lipschitz continuous with respect to  $(x, y, z)$ , uniformly in  $t \in [0, T]$ , *i.e.* satisfies

$$(\text{Lip}_f) \equiv |f(t, x, y, z) - f(t, x', y', z')| \leq [f]_{\text{Lip}}(|x - x'| + |y - y'| + |z - z'|). \quad (13)$$

Under the previous assumptions on  $b, \sigma, h, f$ , the BSDE (12) has a unique  $\mathbb{R} \times \mathbb{R}^q$ -valued,  $\mathcal{F}_t$ -adapted solution  $(Y, Z)$  satisfying

$$\mathbb{E} \left( \sup_{t \in [0, T]} |Y_t|^2 + \int_0^T |Z_s|^2 ds \right) < +\infty$$

(see [47], see also [36]).

Let us consider now  $(\bar{X}_{t_k})_{k=0, \dots, n}$  the discrete time Euler scheme with step  $\Delta_n = \frac{T}{n}$  of the diffusion process  $(X_t)_{t \in [0, T]}$ , where  $(t_k)_{0 \leq k \leq n}$  is the uniform mesh of the interval  $[0, T]$  defined by  $t_k := \frac{kT}{n}$ ,  $k = 0, \dots, n$ . It reads

$$\bar{X}_{t_k} = \bar{X}_{t_{k-1}} + \Delta_n b(t_{k-1}, \bar{X}_{t_{k-1}}) + \sigma(t_{k-1}, \bar{X}_{t_{k-1}})(W_{t_k} - W_{t_{k-1}}), \quad k = 1, \dots, n, \quad \bar{X}_0 = x.$$

We will need to extend this Euler schemes into a continuous time process  $(\bar{X}_t^n)_{t \in [0, T]}$ , sometimes called *genuine Euler scheme* by setting (we drop the dependence in  $n$  when no ambiguity):

$$\bar{X}_t = \bar{X}_{t_{k-1}} + (t - t_{k-1})b(t_{k-1}, \bar{X}_{t_{k-1}}) + \sigma(t_{k-1}, \bar{X}_{t_{k-1}})(W_t - W_{t_{k-1}}), \quad t \in [t_{k-1}, t_k], \quad k = 1, \dots, n. \quad (14)$$

If we set  $\underline{t} = \frac{kT}{n}$  when  $t \in [t_k, t_{k+1})$ , one easily checks that  $(\bar{X}_t)_{t \in [0, T]}$  is an Itô process satisfying

$$d\bar{X}_t = (t - \underline{t})b(\underline{t}, \bar{X}_{\underline{t}}) + \sigma(\underline{t}, \bar{X}_{\underline{t}})dW_t, \quad \bar{X}_0 = x.$$

In particular  $(\bar{X}_t)_{t \in [0, T]}$  is a continuous  $\mathcal{F}_t$ -adapted process and, it is classical background (see *e.g.* [11]), that, under the assumptions made on  $b$  and  $\sigma$ , one has for every  $p \in (0, +\infty)$ ,

$$\left\| \sup_{t \in [0, T]} |X_t| \right\|_p + \sup_{n \geq 1} \left\| \sup_{t \in [0, T]} |\bar{X}_t^n| \right\|_p \leq C_{b, \sigma, p, T}(1 + |x|),$$

for a positive constant  $C_{b, \sigma, p, T}$ . On the other hand, there exists a real constant  $C'_{b, \sigma, p, T} > 0$  such that, for every  $n \geq 1$  and for every  $p \in (0, +\infty)$ ,

$$\left\| \sup_{t \in [0, T]} |X_t - \bar{X}_t^n| \right\|_p \leq C'_{b, \sigma, p, T} \sqrt{\Delta_n}(1 + |x|).$$

As a consequence, general existence-uniqueness results for BSDEs ensure (see [48]) that there exists a unique solution  $(\bar{Y}, \bar{Z})$  to the Markovian BSDE associated to the genuine Euler scheme, namely

$$\bar{Y}_t = h(\bar{X}_T) + \int_t^T f(s, \bar{X}_s, \bar{Y}_s, \bar{Z}_s) ds - \int_t^T \bar{Z}_s dW_s, \quad t \in [0, T], \quad (15)$$

such that

$$\mathbb{E} \left( \sup_{t \in [0, T]} |\bar{Y}_t|^2 + \int_0^T |\bar{Z}_s|^2 ds \right) < +\infty.$$

Then, we can apply the classical comparison result (Proposition 2.1 from [38]) with  $f^1(\omega, t, y, z) = f(t, \bar{X}_t(\omega), y, z)$  and  $f^2(\omega, t, x, y, z) = f(t, X_t(\omega), y, z)$  which immediately yields the existence of real constants  $C_{b, \sigma, f, T}^{(i)} > 0$ ,  $i = 1, 2$ , such that

$$\begin{aligned} \mathbb{E} \left( \sup_{t \in [0, T]} |Y_t - \bar{Y}_t|^2 + \int_0^T |Z_t - \bar{Z}_t|^2 dt \right) &\leq C^{(1)} \left( \mathbb{E}(h(X_T) - h(\bar{X}_T))^2 + [f]_{\text{Lip}}^2 \mathbb{E} \int_0^T |X_t - \bar{X}_t|^2 dt \right) \\ &\leq C^{(1)} ([h]_{\text{Lip}}^2 + T[f]_{\text{Lip}}^2) \mathbb{E} \left( \sup_{t \in [0, T]} |X_t - \bar{X}_t|^2 \right) \\ &\leq C_{b, \sigma, f, T}^{(2)} \Delta_n. \end{aligned}$$

Unfortunately, at this stage, the couple  $(\bar{Y}_t, \bar{Z}_t)_{t \in [0, T]}$  is still “untractable” for numerical purposes (it satisfies no Dynamic Programming Principle due to its continuous time nature and there is no possible exact simulation, etc). This is mainly due to  $\bar{Z}$  on which little is known (by contrast with  $Z$  which is closely connected to a PDE as it will be recalled further on). So we will need to go deeper in the time discretization, by discretizing the  $Z$  term itself.

Consequently, we need to perform again a time discretization, this time on (15), only involving discrete instants  $t_k$ ,  $k = 0, \dots, n$ .

We consider an explicit scheme recursively defined in a backward way as follows:

$$\tilde{Y}_{t_n} = h(\bar{X}_{t_n}) \quad (16)$$

$$\tilde{Y}_{t_k} = \mathbb{E}(\tilde{Y}_{t_{k+1}} | \mathcal{F}_{t_k}) + \Delta_n f(t_k, \bar{X}_{t_k}, \mathbb{E}(\tilde{Y}_{t_{k+1}} | \mathcal{F}_{t_k}), \tilde{\zeta}_{t_k}) \quad (17)$$

and

$$\tilde{\zeta}_{t_k} = \frac{1}{\Delta_n} \mathbb{E}(\tilde{Y}_{t_{k+1}} (W_{t_{k+1}} - W_{t_k}) | \mathcal{F}_{t_k}), \quad k = 0, \dots, n-1. \quad (18)$$

Note that in many situations, one uses the following alternative formula

$$\tilde{\zeta}_{t_k} = \frac{1}{\Delta_n} \mathbb{E}((\tilde{Y}_{t_{k+1}} - \tilde{Y}_{t_k})(W_{t_{k+1}} - W_{t_k}) | \mathcal{F}_{t_k}),$$

which is clearly more natural when thinking to  $\zeta_k$  as a hedging term in discrete time. One easily shows by a backward induction that, for every  $k \in \{0, \dots, n\}$ ,  $\tilde{Y}_{t_k} \in L^2(\Omega, \mathcal{A}, \mathbb{P})$  since  $\sup_{t \in [0, T]} |\bar{X}_{t_k}| \in L^2(\mathbb{P})$ .

Our first aim is to adapt standard comparison theorems to compare the above purely discrete scheme  $(\tilde{Y}_{t_k}, \tilde{Z}_{t_k})$  with the original BSDE to derive error bounds similar to those recalled above between  $(Y, Z)$  and  $(\tilde{Y}, \tilde{Z})$ . To this end, like for the Euler scheme, we need to extend  $\tilde{Y}$  into a continuous time process by an appropriate interpolation. We proceed as follows: let

$$M_T = \sum_{k=1}^n \tilde{Y}_{t_k} - \mathbb{E}(\tilde{Y}_{t_k} | \mathcal{F}_{t_{k-1}}).$$

This random variable is in  $L^2(\mathbb{P})$ . Hence, by the Brownian representation theorem, there exists an  $(\mathcal{F}_t)$ -progressively measurable  $\tilde{Z} \in L^2([0, T] \times \Omega, \mathbb{P} \otimes dt)$  such that

$$M_T = \int_0^T \tilde{Z}_t dW_t.$$

Then  $\tilde{Y}_{t_k} - \mathbb{E}(\tilde{Y}_{t_k} | \mathcal{F}_{t_{k-1}}) = \int_{t_{k-1}}^{t_k} \tilde{Z}_s dW_s$ . In particular

$$\tilde{\zeta}_{t_k} = \frac{1}{\Delta_n} \mathbb{E}(\tilde{Y}_{t_{k+1}}(W_{t_{k+1}} - W_{t_k}) | \mathcal{F}_{t_k}) = \frac{1}{\Delta_n} \mathbb{E}\left(\int_{t_k}^{t_{k+1}} \tilde{Z}_s ds | \mathcal{F}_{t_k}\right), \quad k = 0, \dots, n-1,$$

so that we may define a continuous extension of  $(\tilde{Y}_{t_k})_{0 \leq k \leq n}$  as follows:

$$\tilde{Y}_t = \tilde{Y}_{t_k} - (t - t_k)f(t_k, \bar{X}_{t_k}, \mathbb{E}(\tilde{Y}_{t_{k+1}} | \mathcal{F}_{t_k}), \tilde{\zeta}_{t_k}) + \int_{t_k}^t \tilde{Z}_s dW_s, \quad t \in [t_k, t_{k+1}]. \quad (19)$$

### 3 Error analysis

#### 3.1 The time discretization error

We consider the time discretization scheme  $(\tilde{Y}, \tilde{Z})$  defined by (16)-(33) and (19) and compute below the discretization error bound. The result is proved in the Appendix for self-completeness. It closely follows, like most results of this type, the classical method of proof devised for comparison theorems in [38].

**Theorem 3.1.** (a) Assume the function  $f : [0, T] \times \mathbb{R} \times \mathbb{R}^d \times \mathbb{R} \rightarrow \mathbb{R}$  is Lipschitz continuous in  $(t, x, y, z)$  and that

$$\forall t \geq 0, \quad |f(t, x, y, z)| \leq C(f)(1 + |x| + |y| + |z|). \quad (20)$$

Then, there exists a real constant  $C_{b,\sigma,f,T} > 0$  such that, for every  $n \geq 1$ ,

$$\max_{k \in \{0, \dots, n\}} \mathbb{E}|Y_{t_k} - \tilde{Y}_{t_k}|^2 + \int_0^T \mathbb{E}|Z_t - \tilde{Z}_t|^2 dt \leq C_{b,\sigma,f,T} \left( \Delta_n + \int_0^T \mathbb{E}|Z_s - Z_{\underline{s}}|^2 ds \right).$$

(b) Assume that the functions  $b, \sigma, h, f$  are bounded in  $x$ , uniformly Lipschitz continuous in  $(x, y, z)$  and Hölder continuous of parameter  $1/2$  with respect to  $t$ . Suppose furthermore that  $h$  is of class  $C_b^{2+\alpha}$ ,  $\alpha \in (0, 1)$  and that  $\sigma\sigma^*$  is uniformly elliptic. Then

$$\int_0^T \mathbb{E}|Z_s - Z_{\underline{s}}|^2 ds \leq C_{b,\sigma,f,T}^{(1)} \Delta_n, \quad (21)$$

so that there exists a real constant  $\tilde{C}_{b,\sigma,f,T} > 0$  such that, for every  $n \geq 1$ ,

$$\sup_{k \in \{0, \dots, n\}} \mathbb{E}|Y_{t_k} - \tilde{Y}_{t_k}|^2 + \int_0^T \mathbb{E}|Z_t - \tilde{Z}_t|^2 dt \leq \tilde{C}_{b,\sigma,f,T} \Delta_n. \quad (22)$$

NOTATIONS (CHANGE OF). The previous schemes (16)-(17) involve some quantities and operators which will be the core of what follows and are of discrete time nature. So, in order to simplify the proofs and alleviate the notations, we will identify every time step  $t_k^n$  by  $k$  and we will denote  $\mathbb{E}_k = \mathbb{E}(\cdot | \mathcal{F}_{t_k})$ . Thus, we will switch to

$$\bar{X}_k := \bar{X}_{t_k}, \quad \tilde{Y}_k := \tilde{Y}_{t_k}, \quad f_k(x, y, z) = f(t_k, x, y, z).$$

However note that, in the proof of the Appendix, we still use continuous time notations.



### 3.2 Error bound for the quantization scheme

In this section, we consider the quantization scheme (7)-(8) and compute the quadratic quantization error  $(\mathbb{E}|\tilde{Y}_{t_k} - \hat{Y}_{t_k}|^2)^{1/2}$  induced by the approximation of  $\tilde{Y}_{t_k}$  by  $\hat{Y}_{t_k}$ , for every  $k = 0, \dots, n$ . This leads to the following result.

**Theorem 3.2.** *Assume that the drift  $b$  and the diffusion coefficient  $\sigma$  of the diffusion  $(X_t)_{t \in [0, T]}$  defined by (11) are Lipschitz continuous, that the driver function  $f$  satisfies  $(Lip_f)$  (Assumption(13)) and that the function  $h$  is  $[h]_{Lip}$ -Lipschitz continuous.*

(a) For every  $k = 0, \dots, n$ ,

$$\|\tilde{Y}_k - \hat{Y}_k\|_2^2 \leq \sum_{i=k}^n e^{(1+[f]_{Lip})t_i} K_i(b, \sigma, T, f) \|\bar{X}_i - \hat{X}_i\|_2^2, \quad (23)$$

where  $K_n(b, \sigma, T, f) := [h]_{Lip}^2$  and, for every  $i = 0, \dots, n-1$ , one can choose (provided  $n \geq n_0$ ),

$$K_i(b, \sigma, T, f) := \kappa_1^2 e^{2\kappa_0(T-t_k)} + \left(1 + \Delta_{n_0}\right) \left(C_{1,k}(b, \sigma, T, f) \Delta_{n_0} + C_{2,k}(b, \sigma, T, f)\right),$$

with

$$C_{2,k}(b, \sigma, T, f) = q \kappa_1^2 [f]_{Lip}^2 e^{2\Delta_n C_{b,\sigma,T} + 2\kappa_0(T-t_{k+1})}$$

and

$$C_{1,k}(b, \sigma, T, f) = [f]_{Lip}^2 + \frac{C_{2,k}(b, \sigma, T, f)}{q}.$$

(b) For every  $k = 0, \dots, n$ ,

$$\Delta_n \sum_{k=0}^{n-1} \|\tilde{\zeta}_k - \hat{\zeta}_k\|_2^2 \leq \sum_{k=0}^{n-1} \frac{C_{2,k}(b, \sigma, T, f)}{[f]_{Lip}^2} \|\bar{X}_k - \hat{X}_k\|_2^2 + \sum_{k=0}^{n-1} \|\tilde{Y}_{k+1} - \hat{Y}_{k+1}\|_2^2.$$

Following the usual architecture of proofs of quantization based schemes like those developed in [4] or [41] (among others), The proof of Theorem 3.2 (which will be displayed further on) is divided in two main steps: the first one is to establish the propagation of the Lipschitz property through the functions  $y_k$  and  $z_k$  involved in the Markov representation (16)-(17) of  $\tilde{Y}_k$  and  $\tilde{\zeta}_k$ , namely  $\tilde{Y}_k = y_k(\bar{X}_k)$  and  $\tilde{\zeta}_k = \Delta_n^{-1} z_k(\bar{X}_k)$ , and to control precisely the propagation of their Lipschitz coefficients (an alternative to this phase can be to consider the Lipschitz properties of the flow of the SDE like in [28]). As a second step, we introduce the quantization based scheme which is the counterpart of (16) and (17) for which we establish a backward recursive inequality satisfied by  $\|\tilde{Y}_k - \hat{Y}_k\|_2^2$ .

#### 3.2.1 First step toward the proof of Theorem 3.2

As a first step we need to introduce several operators which appear naturally when representing  $Y_k$ . To be more precise, we set for every  $k \in \{0, \dots, n-1\}$  and every Borel function  $g : \mathbb{R}^d \rightarrow \mathbb{R}$  with polynomial growth

$$\mathcal{E}_k(x, u) = x + \Delta_n b(t_k, x) + \sqrt{\Delta_n} \sigma(t_k, x) u, \quad x \in \mathbb{R}^d, \quad u \in \mathbb{R}^q \quad (24)$$

$$P_{k+1}g(x) = \mathbb{E}g(\mathcal{E}_k(x, \varepsilon)) \quad \text{where } \varepsilon \sim \mathcal{N}(0; I_q) \quad (25)$$

$$Q_{k+1}g(x) = \frac{1}{\sqrt{\Delta_n}} \mathbb{E}\left(g(\mathcal{E}_k(x, \varepsilon))\varepsilon\right). \quad (26)$$

One immediately checks that for every  $k \in \{0, \dots, n-1\}$ ,

$$\mathbb{E}_k g(\bar{X}_{k+1}) = P_{k+1}g(\bar{X}_k) \quad \text{and} \quad \mathbb{E}_k(g(\bar{X}_{k+1})(W_{t_{k+1}^n} - W_{t_k^n})) = \Delta_n Q_{k+1}g(\bar{X}_k).$$

Note that the process  $(\bar{X}_k)_{0 \leq k \leq n}$  is an  $(\mathcal{F}_k)_{0 \leq k \leq n}$ -Markov chain with transitions  $P_{k+1}$ ,  $k = 0, \dots, n-1$ . Moreover, it shares the property to propagate the Lipschitz property as established in the Lemma below.

**Lemma 3.3.** *For every  $k = 0, \dots, n-1$ , the transition operator  $P_{k+1}$  is Lipschitz in the sense that its Lipschitz coefficient defined by  $[P_{k+1}]_{\text{Lip}} := \sup_{f, [f]_{\text{Lip}} \leq 1} [P_{k+1}f]_{\text{Lip}}$  is finite. More precisely, it satisfies:*

$$[P_{k+1}]_{\text{Lip}} \leq e^{\Delta_n C_{b, \sigma, T}} \quad (27)$$

where

$$C_{b, \sigma, T} = [b]_{\text{Lip}} + \frac{1}{2}([\sigma]_{\text{Lip}}^2 + T[b]_{\text{Lip}}^2). \quad (28)$$

Furthermore, if  $n \geq n_0$ , one can take  $C_{b, \sigma, T} = [b]_{\text{Lip}} + \frac{1}{2}([\sigma]_{\text{Lip}}^2 + \frac{T}{n_0}[b]_{\text{Lip}}^2)$ .

**Proof.** We have for every  $x, x' \in \mathbb{R}^d$ , and for every Lipschitz continuous function  $g$  with Lipschitz coefficient  $[g]_{\text{Lip}}$ ,

$$\begin{aligned} |P_{k+1}g(x) - P_{k+1}g(x')|^2 &\leq \mathbb{E} |g(\mathcal{E}_k(x, \varepsilon)) - g(\mathcal{E}_k(x', \varepsilon))|^2 \\ &\leq [g]_{\text{Lip}}^2 \mathbb{E} |\mathcal{E}_k(x, \varepsilon) - \mathcal{E}_k(x', \varepsilon)|^2 \end{aligned}$$

and elementary computations, already carried out in [4], show that

$$\begin{aligned} \mathbb{E} |\mathcal{E}_k(x, \varepsilon) - \mathcal{E}_k(x', \varepsilon)|^2 &\leq (1 + \Delta_n(2[b(t_k^n, \cdot)]_{\text{Lip}} + [\sigma(t_k^n, \cdot)]_{\text{Lip}}^2) + \Delta_n^2[b(t_k^n, \cdot)]_{\text{Lip}}^2) |x - x'|^2 \\ &\leq (1 + \Delta_n(2[b]_{\text{Lip}} + [\sigma]_{\text{Lip}}^2) + \Delta_n^2[b]_{\text{Lip}}^2) |x - x'|^2 \\ &\leq (1 + \Delta_n C_{b, \sigma, T})^2 |x - x'|^2 \\ &\leq e^{2\Delta_n C_{b, \sigma, T}} |x - x'|^2 \end{aligned}$$

where  $C_{b, \sigma, T}$  can be e.g. taken equal to  $[b]_{\text{Lip}} + \frac{1}{2}([\sigma]_{\text{Lip}}^2 + \frac{T}{n_0}[b]_{\text{Lip}}^2)$  provided  $n \geq n_0$ . It follows that  $P_{k+1}$  is Lipschitz with Lipschitz constant  $[P_{k+1}]_{\text{Lip}} \leq e^{\Delta_n C_{b, \sigma, T}}$ .  $\square$

**Proposition 3.4.** (see [4]) (a) *The sequence of functions  $y_k$ ,  $k = 0, \dots, n$ , defined by the backward induction*

$$\begin{aligned} y_n &= h \\ y_k &= P_{k+1}y_{k+1} + \Delta_n f_k(\cdot, P_{k+1}y_{k+1}, Q_{k+1}y_{k+1}), \quad k = 0, \dots, n-1, \end{aligned}$$

satisfies  $\tilde{Y}_k = y_k(\bar{X}_k)$  for every  $k \in \{0, \dots, n\}$ .

(b) *Furthermore, assume that the function  $h$  is  $[h]_{\text{Lip}}$ -Lipschitz continuous and that the function  $f(t, x, y, z)$  is  $[f]_{\text{Lip}}$ -Lipschitz continuous in  $(x, y, z)$ , uniformly in  $t \in [0, T]$ . Then, for every  $k \in \{0, \dots, n\}$ , the function  $y_k$  is  $[y_k]_{\text{Lip}}$ -Lipschitz continuous and there exists real constants  $\kappa_0 = C_{b, \sigma, T} + [f]_{\text{Lip}}$ , where*

$C_{b, \sigma, T}$  is given by (28) in Lemma 3.3, and  $\kappa_1 = \frac{[f]_{\text{Lip}}}{\kappa_0} + [h]_{\text{Lip}}$  such that

$$\begin{aligned} [y_k]_{\text{Lip}} &\leq \frac{\Delta_n}{e^{\kappa_0 \Delta_n} - 1} (e^{\kappa_0(T-t_k^n)} - 1) [f]_{\text{Lip}} + e^{\kappa_0(T-t_k^n)} [h]_{\text{Lip}} \\ &\leq e^{\kappa_0(T-t_k^n)} \kappa_1. \end{aligned} \quad (29)$$

In particular,

$$\sup_{n \geq 1} \max_{k=0, \dots, n} [y_k]_{\text{Lip}} \leq e^{\kappa_0 T} \kappa_1 < +\infty. \quad (30)$$

*Proof.* (a) We proceed by a backward induction using (16) and (17), relying on the fact that  $(\bar{X}_k)_{k=0, \dots, n}$  is a Markov chain which propagates Lipschitz continuity. In fact,  $\tilde{Y}_n = h(\bar{X}_n) := y_n(\bar{X}_n)$ . Assuming that  $\tilde{Y}_{k+1} = y_{k+1}(\bar{X}_{k+1})$  and using Equation (17) and the Markov property, we get

$$\begin{aligned}\tilde{Y}_k &= \mathbb{E}(y_{k+1}(\bar{X}_{k+1})|\bar{X}_k) + \Delta_n f_k(\bar{X}_k, \mathbb{E}(y_{k+1}(\bar{X}_{k+1})|\bar{X}_k), \zeta_{t_k}^n) \\ &= P_{k+1}y_{k+1}(\bar{X}_k) + \Delta_n f_k(\bar{X}_k, P_{k+1}y_{k+1}(\bar{X}_k), Q_{k+1}y_{k+1}(\bar{X}_k)) \\ &= y_k(\bar{X}_k).\end{aligned}$$

(b) We also show this claim by a backward induction. In fact,  $\tilde{Y}_n = h(\bar{X}_n) := y_n(\bar{X}_n)$  and  $h$  is  $[h]_{\text{Lip}}$ -Lipschitz. Suppose that  $y_{k+1}$  is  $[y_{k+1}]_{\text{Lip}}$ -Lipschitz continuous. Then, for every  $x, x' \in \mathbb{R}^d$ , we can write

$$\begin{aligned}y_k(x) - y_k(x') &= \mathbb{E}(y_{k+1}(\mathcal{E}_k(x, \varepsilon)) - y_{k+1}(\mathcal{E}_k(x', \varepsilon))) \\ &\quad + \Delta_n [A_{x,x'}(x - x') + B_{x,x'}\mathbb{E}(y_{k+1}(\mathcal{E}_k(x, \varepsilon)) - y_{k+1}(\mathcal{E}_k(x', \varepsilon))) \\ &\quad + C_{x,x'}\mathbb{E}((y_{k+1}(\mathcal{E}_k(x, \varepsilon)) - y_{k+1}(\mathcal{E}_k(x', \varepsilon)))\varepsilon)]\end{aligned}$$

where  $\varepsilon \sim \mathcal{N}(0, I_q)$  and

$$\begin{aligned}A_{x,x'} &= \frac{f_k(x, P_{k+1}y_{k+1}(x), Q_{k+1}y_{k+1}(x)) - f_k(x', P_{k+1}y_{k+1}(x), Q_{k+1}y_{k+1}(x))}{x - x'} \mathbf{1}_{\{x \neq x'\}}, \\ B_{x,x'} &= \frac{f_k(x', P_{k+1}y_{k+1}(x), Q_{k+1}y_{k+1}(x)) - f_k(x', P_{k+1}y_{k+1}(x'), Q_{k+1}y_{k+1}(x))}{P_{k+1}y_{k+1}(x) - P_{k+1}y_{k+1}(x')} \mathbf{1}_{P_{x,x'}}, \\ C_{x,x'} &= \frac{f_k(x', P_{k+1}y_{k+1}(x'), Q_{k+1}y_{k+1}(x)) - f_k(x', P_{k+1}y_{k+1}(x'), Q_{k+1}y_{k+1}(x'))}{Q_{k+1}y_{k+1}(x) - Q_{k+1}y_{k+1}(x')} \mathbf{1}_{Q_{x,x'}},\end{aligned}$$

with  $P_{x,x'} = \{P_{k+1}y_{k+1}(x) \neq P_{k+1}y_{k+1}(x')\}$  and  $Q_{x,x'} = \{Q_{k+1}y_{k+1}(x) \neq Q_{k+1}y_{k+1}(x')\}$ . The function  $f_k$  being Lipschitz continuous, one clearly has  $|A_{x,x'}|, |B_{x,x'}|, |C_{x,x'}| \leq [f]_{\text{Lip}}$ . Now, taking advantage of the linearity of expectation, we get

$$y_k(x) - y_k(x') = \mathbb{E} \left( (y_{k+1}(\mathcal{E}(x, \varepsilon)) - y_{k+1}(\mathcal{E}(x', \varepsilon))) \left( 1 + \Delta_n (B_{x,x'} + C_{x,x'}\varepsilon) \right) \right) + A_{x,x'}(x - x').$$

Then Schwarz's Inequality yields

$$|y_k(x) - y_k(x')| \leq \|y_{k+1}(\mathcal{E}(x, \varepsilon)) - y_{k+1}(\mathcal{E}(x', \varepsilon))\|_2 \|1 + \Delta_n (B_{x,x'} + C_{x,x'}\varepsilon)\|_2 + \Delta_n [f]_{\text{Lip}} |x - x'|.$$

Now,

$$\|y_{k+1}(\mathcal{E}_k(x, \varepsilon)) - y_{k+1}(\mathcal{E}_k(x', \varepsilon))\|_2 \leq [y_{k+1}]_{\text{Lip}} \|\mathcal{E}_k(x, \varepsilon) - \mathcal{E}_k(x', \varepsilon)\|_2 \leq [y_{k+1}]_{\text{Lip}} e^{\Delta_n C_{b,\sigma,T}} |x - x'|,$$

by Lemma 3.3. On the other hand,

$$\begin{aligned}\|1 - \Delta_n (B_{x,x'} + C_{x,x'}\varepsilon)\|_2^2 &= (1 + \Delta_n B_{x,x'})^2 + \Delta_n^2 C_{x,x'}^2 \\ &\leq 1 + 2\Delta_n [f]_{\text{Lip}} + 2\Delta_n^2 [f]_{\text{Lip}}^2 \\ &\leq e^{2\Delta_n [f]_{\text{Lip}}}.\end{aligned}$$

Finally, owing to the definition of  $\kappa_0$ , we get

$$|y_k(x) - y_k(x')| \leq (e^{\Delta_n \kappa_0} [y_{k+1}]_{\text{Lip}} + \Delta_n [f]_{\text{Lip}}) |x - x'|$$

i.e.  $y_k$  is Lipschitz continuous with Lipschitz coefficient  $[y_k]_{\text{Lip}}$  satisfying

$$[y_k]_{\text{Lip}} \leq e^{\kappa_0 \Delta_n} [y_{k+1}]_{\text{Lip}} + \Delta_n [f]_{\text{Lip}}.$$

The conclusion follows by induction.  $\square$

### 3.2.2 Second step of the proof of Theorem 3.2

Let  $(\widehat{X}_k)_{k=0, \dots, n}$  be the quantization of the Markov chain  $\bar{X}$ , where every quantizer  $\widehat{X}_k$  is of size  $N_k$ , for every  $k \in \{0, \dots, n\}$ . Recall that the discrete time quantized BSDE process  $(\widehat{Y}_k)_{k=0, \dots, n}$  is defined by the following recursive algorithm:

$$\begin{aligned} \widehat{Y}_n &= h(\widehat{X}_n) \\ \widehat{Y}_k &= \widehat{\mathbb{E}}_k(\widehat{Y}_{k+1}) + \Delta_n f_k(\widehat{X}_k, \widehat{\mathbb{E}}_k(\widehat{Y}_{k+1}), \widehat{\zeta}_k) \\ \text{with} \quad \widehat{\zeta}_k &= \frac{1}{\Delta_n} \widehat{\mathbb{E}}_k(\widehat{Y}_{k+1} \Delta W_{t_{k+1}}), k = 0, \dots, n-1, \end{aligned}$$

where  $\widehat{\mathbb{E}}_k = \mathbb{E}(\cdot | \widehat{X}_k)$ . Owing to the previous section, we are now in position to prove Theorem 3.2.

**Proof of Theorem 3.2.** (a) Using the fact that, for every  $k \in \{0, \dots, n\}$ ,  $\sigma(\widehat{X}_k) \subset \sigma(X_k)$ , we have

$$\widetilde{Y}_k - \widehat{Y}_k = \widetilde{Y}_k - \widehat{\mathbb{E}}_k(\widetilde{Y}_k) + \widehat{\mathbb{E}}_k(\widetilde{Y}_k - \widehat{Y}_k) \quad (31)$$

where  $\widetilde{Y}_k - \widehat{\mathbb{E}}_k(\widetilde{Y}_k)$  and  $\widehat{\mathbb{E}}_k(\widetilde{Y}_k - \widehat{Y}_k)$  are square integrable and orthogonal in  $L^2(\sigma(\widehat{X}_k))$ . As a consequence, using the Pythagoras theorem for conditional expectation yield

$$\|\widetilde{Y}_k - \widehat{Y}_k\|_2^2 = \|\widetilde{Y}_k - \widehat{\mathbb{E}}_k(\widetilde{Y}_k)\|_2^2 + \|\widehat{\mathbb{E}}_k(\widetilde{Y}_k - \widehat{Y}_k)\|_2^2.$$

On the other hand, it follows from the definition of the conditional expectation  $\widehat{\mathbb{E}}_k(\cdot)$  as the best approximation in  $L^2$  among square integrable  $\sigma(\widehat{X}_k)$ -measurable random vectors that

$$\|\widetilde{Y}_k - \widehat{\mathbb{E}}_k(\widetilde{Y}_k)\|_2^2 = \|y_k(\bar{X}_k) - \widehat{\mathbb{E}}_k(y_k(\bar{X}_k))\|_2^2 \leq \|y_k(\bar{X}_k) - y_k(\widehat{X}_k)\|_2^2 \leq [y_k]_{\text{Lip}}^2 \|\bar{X}_k - \widehat{X}_k\|_2^2.$$

Let us consider now the last term of the equality (31). We have,

$$\begin{aligned} \widehat{\mathbb{E}}_k(\widetilde{Y}_k - \widehat{Y}_k) &= \widehat{\mathbb{E}}_k[\widetilde{Y}_{k+1} - \widehat{Y}_{k+1} + \Delta_n (f_k(\bar{X}_k, \mathbb{E}_k(\widetilde{Y}_{k+1}), \widetilde{\zeta}_k) - f_k(\widehat{X}_k, \widehat{\mathbb{E}}_k(\widehat{Y}_{k+1}), \widehat{\zeta}_k))] \\ &= \widehat{\mathbb{E}}_k[\widetilde{Y}_{k+1} - \widehat{Y}_{k+1} + \Delta_n (f_k(\bar{X}_k, \mathbb{E}_k(\widetilde{Y}_{k+1}), \widetilde{\zeta}_k) - f_k(\widehat{X}_k, \widehat{\mathbb{E}}_k(\widetilde{Y}_{k+1}), \widehat{\mathbb{E}}_k(\widetilde{\zeta}_k))) \\ &\quad + \Delta_n (f_k(\widehat{X}_k, \widehat{\mathbb{E}}_k(\widetilde{Y}_{k+1}), \widehat{\mathbb{E}}_k(\widetilde{\zeta}_k)) - f_k(\widehat{X}_k, \widehat{\mathbb{E}}_k(\widehat{Y}_{k+1}), \widehat{\zeta}_k))] \\ &= \widehat{\mathbb{E}}_k[\widetilde{Y}_{k+1} - \widehat{Y}_{k+1} + \Delta_n \widehat{B}_k \widehat{\mathbb{E}}_k(\widetilde{Y}_{k+1} - \widehat{Y}_{k+1}) - \Delta_n \widehat{C}_k \widehat{\mathbb{E}}_k(\widetilde{\zeta}_k - \widehat{\zeta}_k)] \\ &\quad + \Delta_n \widehat{\mathbb{E}}_k(f_k(\bar{X}_k, \mathbb{E}_k(\widetilde{Y}_{k+1}), \widetilde{\zeta}_k) - f_k(\widehat{X}_k, \widehat{\mathbb{E}}_k(\widetilde{Y}_{k+1}), \widehat{\mathbb{E}}_k(\widetilde{\zeta}_k))), \end{aligned}$$

where

$$\widehat{B}_k := \frac{f_k(\widehat{X}_k, \widehat{\mathbb{E}}_k(\widetilde{Y}_{k+1}), \widehat{\mathbb{E}}_k(\widetilde{\zeta}_k)) - f_k(\widehat{X}_k, \widehat{\mathbb{E}}_k(\widehat{Y}_{k+1}), \widehat{\mathbb{E}}_k(\widetilde{\zeta}_k))}{\widehat{\mathbb{E}}_k(\widetilde{Y}_{k+1}) - \widehat{\mathbb{E}}_k(\widehat{Y}_{k+1})} \mathbf{1}_{\{\widehat{\mathbb{E}}_k(\widetilde{Y}_{k+1}) \neq \widehat{\mathbb{E}}_k(\widehat{Y}_{k+1})\}}$$

$$\text{and} \quad \widehat{C}_k := \frac{f_k(\widehat{X}_k, \widehat{\mathbb{E}}_k(\widehat{Y}_{k+1}), \widehat{\mathbb{E}}_k(\widetilde{\zeta}_k)) - f_k(\widehat{X}_k, \widehat{\mathbb{E}}_k(\widehat{Y}_{k+1}), \widehat{\zeta}_k)}{\widehat{\mathbb{E}}_k(\widetilde{\zeta}_k) - \widehat{\mathbb{E}}_k(\widehat{\zeta}_k)} \mathbf{1}_{\{\widehat{\mathbb{E}}_k(\widetilde{\zeta}_k) \neq \widehat{\mathbb{E}}_k(\widehat{\zeta}_k)\}}.$$

As

$$\widehat{\mathbb{E}}_k(\widetilde{\zeta}_k) - \widehat{\mathbb{E}}_k(\widehat{\zeta}_k) = \frac{1}{\Delta_n} \widehat{\mathbb{E}}_k((\widetilde{Y}_{k+1} - \widehat{Y}_{k+1}) \Delta W_{t_{k+1}}),$$

we deduce that

$$\begin{aligned} \widehat{\mathbb{E}}_k(\widetilde{Y}_k - \widehat{Y}_k) &= \widehat{\mathbb{E}}_k[(\widetilde{Y}_{k+1} - \widehat{Y}_{k+1})(1 - \Delta_n \widehat{B}_k - \widehat{C}_k \Delta W_{t_{k+1}})] \\ &\quad + \Delta_n (f_k(\bar{X}_k, \mathbb{E}_k(\widetilde{Y}_{k+1}), \widetilde{\zeta}_k) - f_k(\widehat{X}_k, \widehat{\mathbb{E}}_k(\widetilde{Y}_{k+1}), \widehat{\mathbb{E}}_k(\widetilde{\zeta}_k))). \end{aligned} \quad (32)$$

So, it remains to control each term of the above equality. Considering its last term, it follows from the Lipschitz assumption on the driver  $f_k$  that

$$\begin{aligned} \|f_k(\bar{X}_k, \mathbb{E}_k(\tilde{Y}_{k+1}), \tilde{\zeta}_k) - f_k(\hat{X}_k, \hat{\mathbb{E}}_k(\tilde{Y}_{k+1}), \hat{\mathbb{E}}_k(\tilde{\zeta}_k))\|_2^2 &\leq [f]_{\text{Lip}}^2 (\|\bar{X}_k - \hat{X}_k\|_2^2 \\ &\quad + \|\mathbb{E}_k(\tilde{Y}_{k+1}) - \hat{\mathbb{E}}_k(\mathbb{E}_k(\tilde{Y}_{k+1}))\|_2^2 \\ &\quad + \|\tilde{\zeta}_k - \hat{\mathbb{E}}_k(\tilde{\zeta}_k)\|_2^2). \end{aligned}$$

First, the contraction property of the conditional expectation yields

$$\begin{aligned} \|\mathbb{E}_k(\tilde{Y}_{k+1}) - \hat{\mathbb{E}}_k(\mathbb{E}_k(\tilde{Y}_{k+1}))\|_2^2 &\leq \|P_{k+1}y_{k+1}(\bar{X}_k) - P_{k+1}y_{k+1}(\hat{X}_k)\|_2^2 \\ &\leq [P_{k+1}]_{\text{Lip}}^2 [y_{k+1}]_{\text{Lip}}^2 \|\bar{X}_k - \hat{X}_k\|_2^2. \end{aligned}$$

On the other hand, set

$$z_k(x) = \mathbb{E}(y_{k+1}(\mathcal{E}_k(x, \varepsilon))\varepsilon), \quad k = 0, \dots, n-1,$$

where  $\varepsilon \sim \mathcal{N}(0; I_q)$ . We get

$$\begin{aligned} \|\tilde{\zeta}_k - \hat{\mathbb{E}}_k \tilde{\zeta}_k\|_2^2 &= \frac{1}{\Delta_n^2} \|\mathbb{E}_k(\tilde{Y}_{k+1} \Delta W_{t_{k+1}}) - \hat{\mathbb{E}}_k(\mathbb{E}_k(\tilde{Y}_{k+1} \Delta W_{t_{k+1}}))\|_2^2 \\ &= \frac{1}{\Delta_n} \|z_k(\bar{X}_k) - \hat{\mathbb{E}}_k(z_k(\bar{X}_k))\|_2^2 \\ &\leq \frac{1}{\Delta_n} \|z_k(\bar{X}_k) - z_k(\hat{X}_k)\|_2^2 \\ &\leq \frac{1}{\Delta_n} [z_k]_{\text{Lip}}^2 \|\bar{X}_k - \hat{X}_k\|_2^2. \end{aligned} \tag{33}$$

Let us determine  $[z_k]_{\text{Lip}}$ . We have, for every  $k = 0, \dots, n-1$ ,

$$z_k(x) - z_k(x') = \mathbb{E}\left(\left(y_{k+1}(\mathcal{E}_k(x, \varepsilon)) - y_{k+1}(\mathcal{E}_k(x', \varepsilon))\right)\varepsilon\right)$$

so that

$$\begin{aligned} |z_k(x) - z_k(x')| &\leq [y_{k+1}]_{\text{Lip}} \mathbb{E}\left|\left((x - x') + \Delta_n(b(x) - b(x')) + \sqrt{\Delta_n}(\sigma(x) - \sigma(x'))\varepsilon\right)\varepsilon\right| \\ &\leq [y_{k+1}]_{\text{Lip}} \mathbb{E}\left|(x - x') + \Delta_n(b(x) - b(x')) + \sqrt{\Delta_n}(\sigma(x) - \sigma(x'))\varepsilon\right| |\varepsilon| \\ &\leq [y_{k+1}]_{\text{Lip}} \left\| (x - x') + \Delta_n(b(x) - b(x')) + \sqrt{\Delta_n}(\sigma(x) - \sigma(x'))\varepsilon \right\|_2 |\varepsilon|_2 \\ &\leq [y_{k+1}]_{\text{Lip}} e^{\Delta_n C_{b, \sigma, T}} \sqrt{q} |x - x'| \\ &\leq \sqrt{q} e^{\Delta_n C_{b, \sigma, T}} \kappa_1 e^{\kappa_0(T - t_{k+1}^n)} |x - x'|. \end{aligned}$$

We finally deduce that

$$\begin{aligned} \|f_k(\bar{X}_k, \mathbb{E}_k(\tilde{Y}_{k+1}), \tilde{\zeta}_k) - f_k(\hat{X}_k, \hat{\mathbb{E}}_k(\tilde{Y}_{k+1}), \hat{\mathbb{E}}_k(\tilde{\zeta}_k))\|_2 &\leq \left(C_{1,k}(b, \sigma, T, f) \right. \\ &\quad \left. + \frac{C_{2,k}(b, \sigma, T, f)}{\Delta_n}\right)^{\frac{1}{2}} \|\bar{X}_k - \hat{X}_k\|_2 \end{aligned} \tag{34}$$

since, owing to (27) and (29), we have

$$[f]_{\text{Lip}}^2 (1 + [P_{k+1}]_{\text{Lip}}^2 [y_{k+1}]_{\text{Lip}}^2) \leq C_{1,k}(b, \sigma, T, f) \quad \text{and} \quad [f]_{\text{Lip}}^2 [z_k]_{\text{Lip}}^2 \leq C_{1,k}(b, \sigma, T, f),$$

$k = 0, \dots, n-1$  where

$$C_{2,k}(b, \sigma, T, f) = q\kappa_1^2 [f]_{\text{Lip}}^2 e^{2\Delta_n C_{b,\sigma,T} + 2\kappa_0(T-t_{k+1})}$$

and

$$C_{1,k}(b, \sigma, T, f) = [f]_{\text{Lip}}^2 + \frac{C_{2,k}(b, \sigma, T, f)}{q}.$$

To complete the proof, it suffices to control the remaining terms in Equation (32). Using the (conditional) Schwarz's inequality yields

$$\left| \widehat{\mathbb{E}}_k \left[ (\widetilde{Y}_{k+1} - \widehat{Y}_{k+1})(1 - \Delta_n \widehat{B}_k - \widehat{C}_k \Delta W_{t_{k+1}}) \right] \right| \leq \left[ \widehat{\mathbb{E}}_k (\widetilde{Y}_{k+1} - \widehat{Y}_{k+1})^2 \right]^{\frac{1}{2}} \left[ \widehat{\mathbb{E}}_k (1 - \Delta_n \widehat{B}_k - \widehat{C}_k \Delta W_{t_{k+1}})^2 \right]^{\frac{1}{2}}.$$

Furthermore, using the fact that  $\widehat{\mathbb{E}}_k(\Delta W_{t_{k+1}}) = \widehat{\mathbb{E}}_k(\mathbb{E}_k(\Delta W_{t_{k+1}})) = 0$  and owing to the measurability of  $\widehat{B}_k$  and  $\widehat{C}_k$  with respect to  $\sigma(\widehat{X}_k)$ , we get

$$\begin{aligned} \widehat{\mathbb{E}}_k \left[ (1 - \Delta_n \widehat{B}_k - \widehat{C}_k \Delta W_{t_{k+1}})^2 \right] &= (1 - \Delta_n \widehat{B}_k)^2 + \widehat{C}_k^2 \widehat{\mathbb{E}}_k ((\Delta W_{t_{k+1}})^2) \\ &= (1 - \Delta_n \widehat{B}_k)^2 + \widehat{C}_k^2 \Delta_n \\ &\leq (1 + \Delta_n [f]_{\text{Lip}})^2 + (\Delta_n [f]_{\text{Lip}})^2 \\ &\leq e^{2\Delta_n [f]_{\text{Lip}}}. \end{aligned}$$

Then, using the conditional Schwarz inequality and agains the contraction property of conditional expectation, we get

$$\left\| \widehat{\mathbb{E}}_k \left[ (\widetilde{Y}_{k+1} - \widehat{Y}_{k+1})(1 - \Delta_n \widehat{B}_k - \widehat{C}_k \Delta W_{t_{k+1}}) \right] \right\|_2 \leq e^{\Delta_n [f]_{\text{Lip}}} \|\widetilde{Y}_{k+1} - \widehat{Y}_{k+1}\|_2. \quad (35)$$

Using Schwarz's Inequality for the  $L^2$ -norm, we derive from (31), (32), (34) and (35) that

$$\begin{aligned} \|\widetilde{Y}_k - \widehat{Y}_k\|_2^2 &= \|\widetilde{Y}_k - \widehat{\mathbb{E}}_k(\widetilde{Y}_k)\|_2^2 + \|\widehat{\mathbb{E}}_k(\widetilde{Y}_k - \widehat{Y}_k)\|_2^2 \\ &\leq [y_k]_{\text{Lip}}^2 \|\bar{X}_k - \widehat{X}_k\|_2^2 + \left( e^{\Delta_n [f]_{\text{Lip}}} \|\widetilde{Y}_{k+1} - \widehat{Y}_{k+1}\|_2 + \Delta_n \|f_k(\bar{X}_k, \mathbb{E}_k(\widetilde{Y}_{k+1}), \widetilde{\zeta}_k) \right. \\ &\quad \left. - f_k(\widehat{X}_k, \widehat{\mathbb{E}}_k(\widetilde{Y}_{k+1}), \widehat{\mathbb{E}}_k(\widetilde{\zeta}_k)) \right)_2^2 \\ &\leq [y_k]_{\text{Lip}}^2 \|\bar{X}_k - \widehat{X}_k\|_2^2 + \left( e^{\Delta_n [f]_{\text{Lip}}} \|\widetilde{Y}_{k+1} - \widehat{Y}_{k+1}\|_2 \right. \\ &\quad \left. + \Delta_n \left( C_{1,k}(b, \sigma, T, f) + \frac{C_{2,k}(b, \sigma, T, f)}{\Delta_n} \right)^{\frac{1}{2}} \|\bar{X}_k - \widehat{X}_k\|_2 \right)^2. \end{aligned} \quad (36)$$

Relying on the classical identity

$$(a + b)^2 \leq a^2(1 + \Delta_n) + b^2(1 + \Delta_n^{-1}),$$

we derive that

$$\begin{aligned} &\left( e^{\Delta_n [f]_{\text{Lip}}} \|\widetilde{Y}_{k+1} - \widehat{Y}_{k+1}\|_2 + \Delta_n \left( C_{1,k}(b, \sigma, T, f) + \frac{C_{2,k}(b, \sigma, T, f)}{\Delta_n} \right)^{\frac{1}{2}} \|\bar{X}_k - \widehat{X}_k\|_2 \right)^2 \\ &\leq e^{\Delta_n [f]_{\text{Lip}}} (1 + \Delta_n) \|\widetilde{Y}_{k+1} - \widehat{Y}_{k+1}\|_2^2 \\ &\quad + \left( 1 + \frac{1}{\Delta_n} \right) \Delta_n^2 \left( C_{1,k}(b, \sigma, T, f) + \frac{C_{2,k}(b, \sigma, T, f)}{\Delta_n} \right) \|\bar{X}_k - \widehat{X}_k\|_2^2 \\ &\leq e^{\Delta_n(1+[f]_{\text{Lip}})} \|\widetilde{Y}_{k+1} - \widehat{Y}_{k+1}\|_2^2 + \left( 1 + \Delta_n \right) \left( C_{1,k}(b, \sigma, T, f) \Delta_n + C_{2,k}(b, \sigma, T, f) \right) \|\bar{X}_k - \widehat{X}_k\|_2^2. \end{aligned}$$

Hence (using an upper-bound for  $\Delta_n$  e.g. like  $T$  or  $T/n_0$ , if  $n \geq n_0$ ), we obtain

$$\|\widetilde{Y}_k - \widehat{Y}_k\|_2^2 \leq e^{\Delta_n(1+[f]_{\text{Lip}})} \|\widetilde{Y}_{k+1} - \widehat{Y}_{k+1}\|_2^2 + K_k(b, \sigma, T, f) \|\bar{X}_k - \widehat{X}_k\|_2^2. \quad (37)$$

It follows that, for every  $k \in \{0, \dots, n-1\}$ ,

$$e^{\Delta_n k(1+[f]_{\text{Lip}})} \|\tilde{Y}_k - \hat{Y}_k\|_2^2 \leq e^{\Delta_n(k+1)(1+[f]_{\text{Lip}})} \|\tilde{Y}_{k+1} - \hat{Y}_{k+1}\|_2^2 + e^{\Delta_n k(1+[f]_{\text{Lip}})} \tilde{K}_k(b, \sigma, T, f) \|\bar{X}_k - \hat{X}_k\|_2^2$$

where

$$\begin{aligned} \tilde{K}_k(b, \sigma, T, f) &:= [y_{k+1}]_{\text{Lip}}^2 + \left(1 + \frac{T}{n}\right) \left(C_{1,k}(b, \sigma, T, f) \frac{T}{n_0} + C_{2,k}(b, \sigma, T, f)\right), \quad k = 0, \dots, n-1, \\ &\leq K_k(b, \sigma, T, f) \end{aligned}$$

(if  $n \geq n_0$ ). Keeping in mind that  $\|\tilde{Y}_n - \hat{Y}_n\|_2^2 \leq [h]_{\text{Lip}}^2 \|\bar{X}_n - \hat{X}_n\|_2^2$ , we finally derive by a backward induction that

$$\|\tilde{Y}_k - \hat{Y}_k\|_2^2 \leq \sum_{i=k}^n K_i(b, \sigma, T, f) \|\bar{X}_i - \hat{X}_i\|_2^2.$$

(b) We derive from the very definition of  $\tilde{\zeta}_k$  and  $\hat{\zeta}_k$  that

$$\tilde{\zeta}_k - \hat{\zeta}_k = (\tilde{\zeta}_k - \hat{\mathbb{E}}_k(\tilde{\zeta}_k)) \perp (\hat{\mathbb{E}}_k(\tilde{\zeta}_k) - \hat{\zeta}_k).$$

We know from (33) in the the proof of claim (a) that

$$\|\hat{\mathbb{E}}_k(\tilde{\zeta}_k - \hat{\mathbb{E}}_k(\tilde{\zeta}_k))\|_2^2 \leq \frac{[z_k]_{\text{Lip}}^2}{\Delta_n} \|\bar{X}_k - \hat{X}_k\|_2^2.$$

On the other hand, as  $\sigma(\hat{X}_k) \subset \sigma(X\bar{X}_k)$ , it is clear that  $\hat{\mathbb{E}}_k(\tilde{\zeta}_k) = \frac{1}{\Delta_n} \hat{\mathbb{E}}_k(\tilde{Y}_{k+1} \Delta W_{t_{k+1}})$  so that

$$\|\hat{\mathbb{E}}_k(\tilde{\zeta}_k) - \hat{\zeta}_k\|_2^2 = \frac{1}{\Delta_n^2} \|\hat{\mathbb{E}}_k((\tilde{Y}_{k+1} - \hat{Y}_{k+1}) \Delta W_{t_{k+1}})\|_2^2.$$

Conditional Schwarz's Inequality applied with  $\hat{\mathbb{E}}_k$  implies that

$$\hat{\mathbb{E}}_k((\tilde{Y}_{k+1} - \hat{Y}_{k+1}) \Delta W_{t_{k+1}})^2 \leq (\hat{\mathbb{E}}_k(\tilde{Y}_{k+1} - \hat{Y}_{k+1})^2) \Delta_n$$

which in turn implies that

$$\|\hat{\mathbb{E}}_k(\tilde{\zeta}_k) - \hat{\zeta}_k\|_2^2 = \frac{1}{\Delta_n} \|\tilde{Y}_{k+1} - \hat{Y}_{k+1}\|_2^2$$

so that finally

$$\Delta_n \|\hat{\mathbb{E}}_k(\tilde{\zeta}_k) - \hat{\zeta}_k\|_2^2 \leq \frac{C_{2,k}(b, \sigma, T, f)}{[f]_{\text{Lip}}^2} \|\bar{X}_k - \hat{X}_k\|_2^2 + \|\tilde{Y}_{k+1} - \hat{Y}_{k+1}\|_2^2.$$

□

**Remark 3.1.** Remark that the key property leading to Theorem 3.2 and allowing to improve the existing results for similar problems (see *e.g.* [4]) is the Pythagoras like equality (36) which is true only for the quadratic norm. This equality is the key to get the sharp constant equal to 1 before the term  $\|\hat{\mathbb{E}}_k(\tilde{Y}_k - \hat{Y}_k)\|_2^2$ .

### 3.3 Computing the $\widehat{\zeta}_k$ terms

Recall that for every  $k \in \{0, \dots, n-1\}$ , the  $\mathbb{R}^q$ -valued random vector  $\widehat{\zeta}_k = (\widehat{\zeta}_k^1, \dots, \widehat{\zeta}_k^q)$  reads

$$\widehat{\zeta}_k = \frac{1}{\Delta_n} \widehat{z}_k(\widehat{X}_k) \quad \text{where} \quad \widehat{z}_k(\widehat{X}_k) = \widehat{\mathbb{E}}_k(\widehat{Y}_{k+1} \Delta W_{t_{k+1}})$$

with  $\widehat{z}_k : \Gamma_k \rightarrow \mathbb{R}^q$  is a Borel function ( $\Gamma_k$  is the grid used to quantize  $\bar{X}_k$ ). As  $\widehat{Y}_{k+1} = \widehat{y}_{k+1}(\widehat{X}_{k+1})$  we easily derive that the function  $\widehat{z}_k$  is defined on  $\Gamma_k = \{x_1^k, \dots, x_{N_k}^k\}$  by the ( $\mathbb{R}^q$ -valued) weighted sum

$$\widehat{z}_k(x_i^k) = \sum_{j=1}^{N_{k+1}} \widehat{y}_{k+1}(x_j^{k+1}) \pi_{ij}^{W,k}$$

where, for every  $(i, j) \in \{1, \dots, N_k\} \times \{1, \dots, N_{k+1}\}$ ,  $\pi_{ij}^{W,k}$  is an  $\mathbb{R}^q$ -valued vector given by

$$\pi_{ij}^{W,k} = \frac{1}{\mathbb{P}(\widehat{X}_k = x_i^k)} \times \mathbb{E} \left( \Delta W_{t_{k+1}} \mathbf{1}_{\{\widehat{X}_{k+1} = x_j^{k+1}, \widehat{X}_k = x_i^k\}} \right).$$

These vector valued “weights” appear as new companion parameters (as well as the original weights  $\pi_{ij}^k$  of the quantized transition matrices) which can be computed *on line* when simulating the Euler scheme of the diffusion by a Monte Carlo simulation.

Note that, for every  $k \in \{0, \dots, n-1\}$  and for every  $i \in \{1, \dots, N_k\}$ ,

$$\begin{aligned} \sum_{j=1}^{N_{k+1}} \pi_j^{W,k} &= \widehat{\mathbb{E}}_k(\Delta W_{t_{k+1}} \mathbf{1}_{\{\widehat{X}_k = x_i^k\}}) \\ &= \widehat{\mathbb{E}}_k \left( \mathbb{E}_k(\Delta W_{t_{k+1}} \mathbf{1}_{\{\widehat{X}_k = x_i^k\}}) \right) \\ &= \widehat{\mathbb{E}}_k \left( \mathbb{E}_k(\Delta W_{t_{k+1}}) \mathbb{P}(\widehat{X}_k = x_i^k) \right) = \widehat{\mathbb{E}}_k 0 = 0. \end{aligned}$$

As a consequence, an alternative formula for  $\widehat{z}_k$  can be

$$\widehat{z}_k(x_i^k) = \sum_{j=1}^{N_{k+1}} \pi_{ij}^{W,k} (\widehat{y}_{k+1}(x_j^{k+1}) - \widehat{y}_k(x_i^k)).$$

## 4 Background on optimal vector quantization

It is important to have in mind that all what precedes holds true for any quantizations  $\widehat{X}_k$  of the Euler scheme  $X_{t_k}$  *i.e.* for any sequence of the form  $\widehat{X}_k = \pi_k(X_{t_k})$ . In fact the theory of optimal vector quantization starts when tackling the problem of optimizing the  $L^2$  (and more generally the  $L^r$ ) mean error induce by this substitution, namely  $\|X_{t_k} - \widehat{X}_k\|_2$  which in turn will provide the best possible error bound for quantization based numerical schemes. This question is in fact a very old question that goes back to te 1940s motivated by Signal transmission and processing. These techniques have been imported in Numerical Probability, originally for numerical integration by cubature formulas in the early 1990’s (see [39] or [15]).

Let  $X : (\Omega, \mathcal{A}, \mathbb{P}) \rightarrow \mathbb{R}^d$  be a random vector lying in  $L^r(\mathbb{P})$ ,  $r \in (0, +\infty)$ . The  $L^r$ -optimal quantization problem of size  $N$  for  $X$  (or equivalently for its distribution  $\mathbb{P}_X$ ) consists in finding the best  $L^r(\mathbb{P})$ -approximation of  $X$  by a random variable  $\pi(X)$  taking at most  $N$  values. The integer  $N$  is called the quantization *level*.



First, we associate to every Borel function  $\pi : \mathbb{R}^d \rightarrow \mathbb{R}$  taking at most  $N$  values the induced  $L^r(\mathbb{P})$ -mean error  $\|X - \pi(X)\|_r$  (where  $\|X\|_r := (\mathbb{E}|X|^r)^{1/r}$  is the usual  $L^r$  norm induced by the norm  $|\cdot|$  on  $\mathbb{R}^d$  and the probability  $\mathbb{P}$  on  $(\Omega, \mathcal{A})$ ). Note that when  $r \in (0, 1)$ , the terms “norm” is an abuse of language since  $L^r(\mathbb{P})$  is only a metric space metrized by  $\|\cdot\|_r$ . As a consequence, finding the best approximation of  $X$  in the earlier described sense amounts to find the solution to the following minimization problem:

$$e_{N,r}(X) = \inf \left\{ \|X - \pi(X)\|_r, v : \mathbb{R}^d \rightarrow \Gamma, \Gamma \subset \mathbb{R}^d, \text{card}(\Gamma) \leq N \right\},$$

where  $\text{card}(\Gamma)$  denotes the cardinality of the set  $\Gamma$  (sometimes called *grid*). It is clear that for every grid  $\Gamma = \{x_1, \dots, x_N\} \subset \mathbb{R}^d$ , for any Borel function  $\pi : \mathbb{R}^d \rightarrow \Gamma$ ,

$$|\xi - \pi(\xi)| \geq \text{dist}(\xi, \Gamma) = \min_{1 \leq i \leq N} |\xi - x_i|.$$

Equality holds if and only if  $\pi$  is a Borel nearest neighbor projection  $\pi_\Gamma$  defined by

$$\pi_\Gamma(\xi) = \sum_{i=1}^N x_i \mathbf{1}_{C_i(\Gamma)}(\xi),$$

where  $(C_i(\Gamma))_{i=1, \dots, N}$  is a Borel partition of  $\mathbb{R}^d$  satisfying

$$\forall i \in \{1, \dots, N\}, \quad C_i(\Gamma) \subset \left\{ \xi \in \mathbb{R}^d : |\xi - x_i| = \min_{j=1, \dots, N} |\xi - x_j| \right\}.$$

Such a Borel partition is called a *Voronoi partition* (induced by  $\Gamma$ ). One defines the Voronoi quantization  $\widehat{X}^\Gamma$  of  $X$  induced by  $\Gamma$  as  $\pi_\Gamma(X)$ . It follows that for every  $r > 0$ ,  $\|X - \widehat{X}^\Gamma\|_r = \|\text{dist}(X, \Gamma)\|_r$  so that the  $L^r$ -optimal quantization finally reads

$$e_{N,r}(X) = \inf \left\{ \|X - \widehat{X}^\Gamma\|_r, \Gamma \subset \mathbb{R}^d, \text{card}(\Gamma) \leq N \right\}. \quad (38)$$

Note that for every level  $N \geq 1$ , the infimum in (38) is in fact a minimum *i.e.* is attained at least at one grid (or codebook)  $\Gamma_N$  at least (see *e.g.* (see *e.g.* [23] or [39])). Any such grid or any of the resulting Borel nearest neighbor projections is called an  $L^r$ -optimal  $N$ -quantizer.

One shows that if  $\text{card}(\text{supp}(\mathbb{P}_X)) \geq N$  then any optimal  $N$ -quantizer is of full size  $N$ . Furthermore (see again [23] or [39]), the  $L^r$ -mean quantization error  $e_{N,r}(X)$  at level  $N$  decreases to 0 as  $N$  goes to infinity. Its rate of convergence is ruled by the so-called Zador Theorem recalled below, in which,  $|\cdot|$  temporarily denotes any norm on  $\mathbb{R}^d$ .

**Theorem 4.1. Zador’s Theorem** (a) *Sharp asymptotic rate (see [23]): Let  $X$  be an  $\mathbb{R}^d$ -valued random vector such that  $X \in L^{r+\delta}(\mathbb{P})$  for soem real number  $\delta > 0$  and let  $\mathbb{P}_X = \varphi \cdot \lambda_d + P_s$  where  $\varphi$  denotes the absolutely continuous part of  $\mathbb{P}_X$  with respect to the Lebesgue measure  $\lambda_d$  on  $\mathbb{R}^d$  and  $P_s$  its singular part. Then*

$$\lim_{N \rightarrow +\infty} N^{r/d} (e_{N,r}(P))^r = J_{r,d} \|\varphi\|_{\frac{d}{d+r}} \in [0, +\infty) \quad (39)$$

with

$$\|\varphi\|_{\frac{d}{d+r}} = \left( \int_{\mathbb{R}^d} \varphi^{\frac{d}{d+r}} d\lambda_d \right)^{\frac{d+r}{d}} \quad \text{and} \quad J_{r,d,|\cdot|} = \inf_{N \geq 1} N^{r/d} e_{N,r}^r(U([0, 1]^d)) \in (0, +\infty)$$

( $U([0, 1]^d)$  denotes the uniform distribution on the hypercube  $[0, 1]^d$ ).

(b) *Non-asymptotic bound (see [34]): Let  $r' > r$ . There exists a universal real constant  $C_{r,r',d} \in (0, +\infty)$  such that, for every  $\mathbb{R}^d$ -valued  $X$  random vector*

$$\forall N \geq 1, \quad e_{p,N}(X) \leq C_{p,p',d} \sigma_{p'}(X) \cdot N^{-\frac{1}{d}}$$

where  $\sigma_{r'}(X) := \inf_{a \in \mathbb{R}^d} \|X - a\|_{r'} \leq +\infty$  is the  $L^{r'}$ -(pseudo-)standard deviation of  $X$ .

**Numerical aspects (few words about)** From the Numerical Probability viewpoint, finding an optimal  $N$ -quantizer  $\Gamma$  is a challenging task, especially in higher dimension ( $d \geq 2$ ). In this paper as in many applications we will mainly focus on the quadratic case  $r = 2$ . Note that, in practice,  $|\cdot|$  will be the canonical Euclidean norm on  $\mathbb{R}^d$  for numerical implementations.

The key property to devise procedures to search for optimal quantizers rely on the following differentiability property of the squared quadratic quantization error (also known as quadratic distortion function) for a fixed level  $N$  (and with respect to the canonical Euclidean norm). First we define the distortion function  $D_{N,2}$  (which is defined on  $(\mathbb{R}^d)^N$  and not on the set of grids of size at most  $N$ ) by:

$$\forall x = (x_1, \dots, x_N) \in (\mathbb{R}^d)^N, \quad D_{N,2}^X(x) = \int_{\mathbb{R}^d} \min_{1 \leq i \leq N} |\xi - x_i|^2 d\mathbb{P}_X(\xi). \quad (40)$$

To any  $N$ -tuple  $x = (x_1, \dots, x_N) \in (\mathbb{R}^d)^N$ , we associate its grid of values  $\Gamma^x = \{x_1, \dots, x_N\}$ , so that  $D_{N,2}^X(x) = \|X - \widehat{X}^{\Gamma^x}\|_2^2$ . In particular, it is clear that

$$e_{N,2}(X) = \inf_{x \in (\mathbb{R}^d)^N} D_{N,2}(x)$$

since an  $N$ -tuples can contain repeated values.

**Proposition 4.2.** (a) *The function  $D_{N,2}$  is differentiable at any  $N$ -tuple  $x \in (\mathbb{R}^d)^N$  having pairwise distinct components and satisfying the following boundary negligibility assumption:*

$$\mathbb{P}_X(\cup_i \partial C_i(\Gamma^x)) = 0.$$

*Its gradient is given by*

$$\nabla D_{N,2}^X(x) = 2 \left( \int_{C_i(\Gamma^x)} (x_i - \xi) d\mathbb{P}_X(\xi) \right)_{i=1, \dots, N}. \quad (41)$$

(b) *The above negligibility assumption on the Voronoi partition boundaries does not depend on the selected partition. It holds in particular when the distribution of  $X$  is strongly continuous i.e. assigns no mass to hyperplanes and, for any distribution  $\mathbb{P}_X$  such that  $\text{card}(\text{supp}(\mathbb{P}_X)) \geq N$ , when  $x \in \text{argmin} D_{N,2}$  (see Theorem 4.2 in [23]).*

The result is a consequence of the interchange of the differentiation and the integral leading to (41) when formally differentiating (40) (see [23, 39]). Consequently, any  $N$ -tuple  $x \in \text{argmin} D_{N,2}$  satisfies

$$\nabla D_{N,2}(x) = 0.$$

Note that this equality also reads, still under the assumption  $\text{card}(\text{supp}(\mathbb{P}_X)) \geq N$ ,

$$\mathbb{E} \left( X | \widehat{X}^\Gamma \right) = \widehat{X}^\Gamma$$

All numerical methods to compute optimal quadratic quantizers are based on this result: recursive procedures like Newton's algorithm (when  $d = 1$ ), randomized fixed point procedures like Lloyd's I algorithms (see e.g. [21, 45]) or recursive stochastic gradient descent like the Competitive Learning Vector Quantization (CLVQ) algorithm (see [21, 39] or [42]) in the multidimensional framework. However note that in higher dimension this equation has several solutions (called *stationary quantizers*) possibly sub-optimal. Optimal quantization grids associated to the multivariate Gaussian random vector can be downloaded from the website [www.quantize.math-fi.com](http://www.quantize.math-fi.com). For more details about numerical methods we refer to the recent survey [40] and the references therein.

**Distortion mismatch** By mismatch we mean the behavior of  $e_s(\Gamma_N, X)$  where  $(\Gamma_N)_{N \geq 1}$  is a sequence of  $L^r$ -optimal quantization grids and  $s > r$ . The first result in that direction go back to [24] for various classes of distributions, in particular distributions with a radial densities satisfying a moment assumption on  $X$  (of higher order than  $s$ , see below). An extension to all random vectors satisfying this moment condition on  $X$  is proposed in AppendixB (it is part of monography in progress co-authored by the first author and H. Luschgy, see [25]).

**Theorem 4.3** ( $L^r$ - $L^s$ -distortion mismatch). *Let  $X : (\Omega, \mathcal{A}, \mathbb{P}) \rightarrow \mathbb{R}^d$  be a random vector and let  $r \in (0 + \infty)$ . Assume that the distribution  $P = \mathbb{P}_X$  of  $X$  has a non-zero absolutely continuous component with density  $\varphi$ . Let  $(\Gamma_N)_{N \geq 1}$  be an  $L^r$ -optimal sequence for grids and let  $s \in (r, r + d)$ . If*

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

for some  $\delta > 0$ , then

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

Note that  $\frac{sd}{d+r-s} > s$  so that distortion mismatch requires more than  $L^s$  integrability. A proof of this theorem is provided in Appendix B.

**Remark.** This result is in some way optimal for the following reason. It has been established in [24] (Theorem 1) that if  $X \in L^{r+}(\mathbb{P})$ , and if  $(\Gamma_n)_{n \geq 1}$  is a sequence of  $L^r$ -asymptotically optimal quantization grids, then

$$\varliminf_N N^{\frac{1}{d}} e_s(\Gamma_n, X) \geq J_{r,d}^{\frac{1}{s}} \left[ \int_{\mathbb{R}^d} \varphi^{\frac{d}{r+d}} d\lambda_d \right]^{\frac{1}{d}} \left[ \int_{\mathbb{R}^d} \varphi^{\frac{d+r-s}{r+d}} d\lambda_d \right]^{\frac{1}{s}}$$

and an elementary application of the inverse Minkowski inequality shows (see Equation (2.11) from [24]) that

$$X \notin L^{r+}(\mathbb{P}) \implies \int g \varphi^{\frac{d}{r+d}} d\lambda_d = +\infty \quad \text{and} \quad X \notin L^{\frac{ds}{d+r-s} +}(\mathbb{P}) \implies \int g \varphi^{\frac{d+r-s}{r+d}} d\lambda_d = +\infty.$$

#### 4.1 Numerical experiments for the BSDE scheme

We deal now with numerical experiments using three examples: the pricing and hedging of a call option (in a market where the risk free return for the borrower and the lender are the same), a bull-call spread option (in a market where the risk free returns for the borrower and the lender are different) and a multidimensional example with the Brownian motion. Numerical tests are performed using our quantized BSDE algorithm. In fact, let us set for  $i = 1, \dots, N_k, j = 1, \dots, N_{k+1}$ ,

$$\begin{aligned} p_i^k &= \mathbb{P}(X_k = x_i^k), k = 0, \dots, n \\ \text{and} \quad p_{ij}^k &= \mathbb{P}(\widehat{X}_{k+1} = x_j^{k+1} | \widehat{X}_k = x_i^k), k = 0, \dots, n-1. \end{aligned}$$

Setting  $\widehat{Y}_k = \widehat{y}_k(\widehat{X}_k)$ , for every  $k \in \{0, \dots, n\}$ , the quantized BSDE scheme reads as

$$\begin{cases} \widehat{y}_n(x_i^n) = h(x_i^n) & i = 1, \dots, N_n \\ \widehat{y}_k(x_i^k) = \widehat{\alpha}_k(x_i^k) + \Delta_n f(t_k, x_i^k, \widehat{\alpha}_k(x_i^k), \widehat{\beta}_k(x_i^k)) & i = 1, \dots, N_k \end{cases}$$

where for  $k = 0, \dots, n-1$ ,

$$\widehat{\alpha}_k(x_i^k) = \sum_{j=1}^{N_{k+1}} \widehat{y}_{k+1}(x_j^{k+1}) p_{ij}^k \quad \text{and} \quad \widehat{\beta}_k(x_i^k) = \frac{1}{\Delta_n} \sum_{j=1}^{N_{k+1}} \widehat{y}_{k+1}(x_j^{k+1}) \pi_{ij}^{W,k}, \quad (42)$$

with

$$\pi_{ij}^{W,k} = \frac{1}{p_i^k} \times \mathbb{E} \left( \Delta W_{t_{k+1}} \mathbf{1}_{\{\widehat{X}_{k+1}=x_j^{k+1}, \widehat{X}_k=x_i^k\}} \right).$$

For all examples, we use a time discretization mesh of length  $n = 20$ . For the optimal quantization method, the optimal quantizers  $\widehat{X}_k$ ,  $k = 1, \dots, n$  (with  $\widehat{X}_0 = X_0$ ) are computed using the Lloyd's algorithm (see e.g. [21, 42] for details on Lloyd's algorithm). The size of the Monte Carlo trials in the Lloyd algorithm is set to  $7 \times 10^5$  (for the first two examples, and, to  $5 \times 10^6$  for the third example) and the number of iterations of the Lloyd procedure is set to 5 (except for the last example where the grids are obtained by a scaling of the optimal grid of Gaussian random variables). We use a uniform dispatching for the quantizers where we assign the same grid size  $N_k$  to the  $\widehat{X}_k$ 's, for every discretization step  $t_k$ ,  $k = 1, \dots, n$ .

#### 4.1.1 Call price in the Black-Scholes model

We consider a call option with maturity  $T$  and strike  $K$  on a stock price  $(X_t)_{t \in [0, T]}$  with dynamics

$$dX_t = \mu X_t dt + \sigma X_t dW_t.$$

Considering a self financing portfolio  $Y_t$  with  $\varphi_t$  assets and bonds with risk free return  $r$ . We know that (see [38]) the portfolio evolves according to the following dynamics:

$$Y_t = Y_T + \int_t^T f(Y_s, Z_s) ds - \int_t^T Z_s dW_s \quad (43)$$

where the payoff  $Y_T = (X_T - K)^+$ , the hedging strategy  $Z_t = \sigma \varphi_t X_t$  and  $f(y, z) = -ry - \frac{\mu-r}{\sigma} z$ . It is clear that the function  $f$  is linear with respect to  $y$  and  $z$  and, it is Lipschitz continuous with  $[f]_{\text{Lip}} = \max(r, \frac{\mu-r}{\sigma})$ . We perform the numerical tests from the algorithm we propose with the following parameters

$$X_0 = 100, \quad r = 0.1, \quad \mu = 0.2, \quad \sigma = 0.25, \quad K = 100, \quad T = 0.1.$$

Setting  $v(t_0, X_0)$  the price of the option at  $t_0 = 0$ , the corresponding exact solution of the BDSE (43) at  $t_0 = 0$  is given by

$$(Y_0, Z_0) = (v(t_0, X_0), \sigma X_0 v(t_0, X_0)) = (3.65997, 14.14823).$$

We have to notice that the solution  $(Y_t, Z_t)$  of (43) provides both the price  $Y_t$  of the option and the hedging strategy  $Z_t$  at time  $t$ , under the historical probability  $\mathbb{P}$ . The Black-Scholes price is given under the risk neutral probability and does not depend on  $\mu$ .

The estimated solution of the BSDE by optimal quantization is  $(\widehat{Y}_0^N, \widehat{Z}_0^N) = (3.66, 13.83)$  for  $N_k = N = 40$ , and,  $(\widehat{Y}_0^N, \widehat{Z}_0^N) = (3.655, 14.147)$  for  $N_k = N = 100$ . The numerical convergence rate of the error  $|Y_0 - \widehat{Y}_0^N|$  is depicted in Figure 1 for  $\mu = 0.05$  and  $\mu = 0.2$  and bring to light a convergence rate of order  $N^{-1}$ .

#### 4.1.2 Bid-ask spread for interest rate

Let us consider now a model for different interest rates introduces in [8]: a borrowing rate  $R$  and a lending rate  $r \leq R$ . Let  $\varphi_t$  still be the amount of assets held at time  $t$ . Then, the dynamics of the replicating portfolio is given by

$$Y_t = Y_T + \int_t^T f(Y_s, Z_s) ds - \int_t^T Z_s dW_s \quad (44)$$

where  $Z_t = \sigma \varphi_t X_t$  and the driver function  $f$  is given by

$$f(y, z) = -ry - \frac{\mu - r}{\sigma} z - (R - r) \min\left(y - \frac{z}{\sigma}, 0\right).$$

As in [7], we consider a bull-call spread comprising a long call with strike  $K_1 = 95$  and two short call with strike  $K_2 = 105$ , with payoff function

$$(X_T - K_1)^+ - 2(X_T - K_2)^+ = Y_T.$$

Furthermore, we consider the set of parameters:

$$X_0 = 100, \quad R = 0.06, \quad r = 0.01, \quad \mu = 0.05, \quad \sigma = 0.2, \quad T = 0.25.$$

The BSDE (44) has no analytical solution. We refer to the reference prices given in [7, 49]:  $(Y_0, Z_0) = (2.96, 0.55)$ . The estimated solution of the BSDE by optimal quantization is  $(\widehat{Y}_0^N, \widehat{Z}_0^N) = (2.97, 0.63)$  for an uniform dispatching with the same grid size  $N_k = N = 40$ .

### 4.1.3 Multidimensional example

We consider the following example due to J.-F. Chassagneux:

$$dX_t = dW_t, \quad -dY_t = f(t, Y_t, Z_t)dt - Z_t \cdot dW_t$$

where  $f(t, y, z) = (z_1 + \dots + z_d)(y - \frac{2+d}{2d})$  and where  $W$  is a  $d$ -dimensional Brownian motion. The solution of this BSDE reads

$$Y_t = \frac{e_t}{1 + e_t}, \quad Z_t = \frac{e_t}{(1 + e_t)^2}, \quad (45)$$

with

$$e_t = \exp(x_1 + \dots + x_d + t).$$

For the numerical experiments, we put the (regular) time discretization mesh to  $n = 20$ , with discretization step  $\Delta$ . We use the uniform dispatching grid allocation and define the quantization  $(\widehat{W}_{t_k})_{0 \leq k \leq n}$  of the Brownian trajectories  $(W_{t_k})_{0 \leq k \leq n}$  from the following recursive procedure

$$\widehat{W}_{t_{k+1}} = \widehat{W}_{t_k} + \sqrt{\Delta} \widehat{\varepsilon}, \quad (46)$$

$\widehat{W}_0 = 0$  and where  $\widehat{\varepsilon}$  is the optimal quantization of the  $d$ -dimensional standard Gaussian random variable. We choose  $t = 0.5$ ,  $d = 2, 3$ , so that  $Y_0 = 0.5$  and  $Z_0^i = 0.24$ , for every  $i = 1, \dots, d$ . We show in Figure 2, the rates of convergence of  $|\widehat{Y}_0^N - 0.5|$  towards 0, for the grid sizes  $N_k = N = 5, \dots, 100$ . The graphics show a rate of convergence of order  $N^{-1/d}$ . In particular, when  $d = 2$ , we get

$$\begin{aligned} & (\widehat{Y}_0, \widehat{Z}_0^1, \widehat{Z}_0^2) = (0.50, 0.28, 0.28) \text{ for } N_k = 40 \\ \text{and } & (\widehat{Y}_0, \widehat{Z}_0^1, \widehat{Z}_0^2) = (0.50, 0.23, 0.23) \text{ for } N_k = 100. \end{aligned}$$

When  $d = 3$ , we get

$$\begin{aligned} & (\widehat{Y}_0, \widehat{Z}_0^1, \widehat{Z}_0^2, \widehat{Z}_0^3) = (0.51, 0.08, 0.06, 0.06) \text{ for } N_k = 40 \\ \text{and } & (\widehat{Y}_0, \widehat{Z}_0^1, \widehat{Z}_0^2, \widehat{Z}_0^3) = (0.51, 0.18, 0.16, 0.11) \text{ for } N_k = 100. \end{aligned}$$

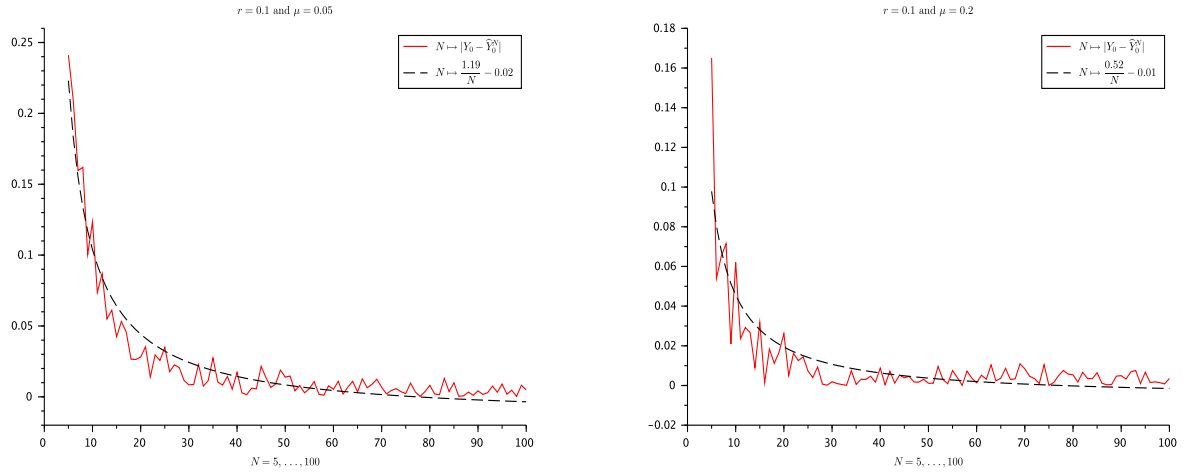


Figure 1: Convergence rate of the quantization error for the Call price in the Black-Scholes model). Abscissa axis: the size  $N = 5, \dots, 100$  of the quantization. Ordinate axis: The error  $|Y_0 - \hat{Y}_0^N|$  and the graph  $N \mapsto \hat{a}/N + \hat{b}$ , where  $\hat{a}$  and  $\hat{b}$  are the regression coefficients. The left hand side graphic corresponds to  $r = 0.1$  and  $\mu = 0.05$  and the right hand side to  $r = 0.1$  and  $\mu = 0.2$ .

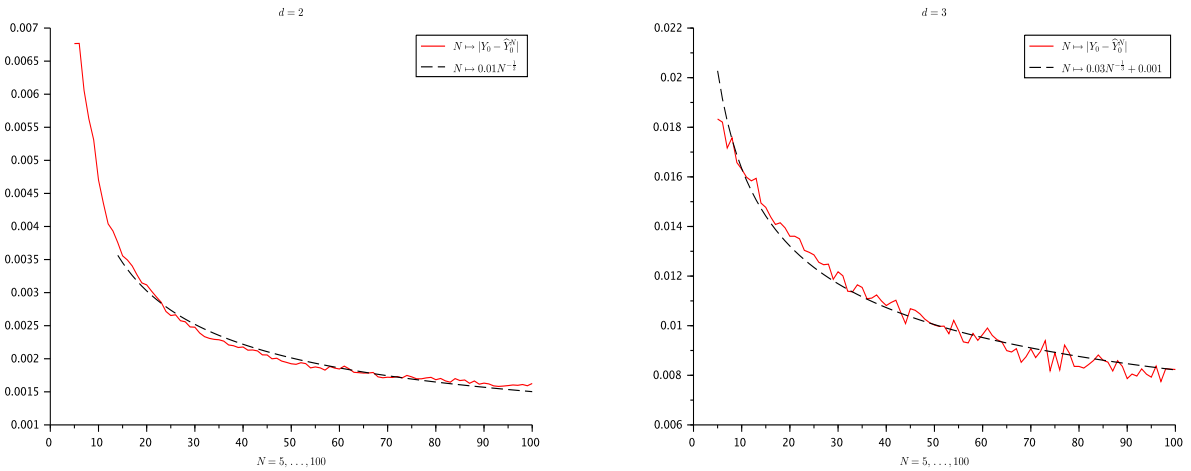


Figure 2: Convergence rate of the quantization error for the multidimensional example). Abscissa axis: the size  $N = 5, \dots, 100$  of the quantization. Ordinate axis: The error  $|Y_0 - \hat{Y}_0^N|$  and the graph  $N \mapsto \hat{a}/N + \hat{b}$ , where  $\hat{a}$  and  $\hat{b}$  are the regression coefficients. The left hand side graphic corresponds to the dimension  $d = 2$  and the right hand side to  $d = 3$ .

## 5 Nonlinear filtering problem

### 5.1 A brief overview of discrete time nonlinear filtering

We consider a discrete time nonlinear filtering problem where the signal process  $(X_k)_{k \geq 0}$  is an  $\mathbb{R}^d$ -valued random vector and the observation process  $(Y_k)_{k \geq 0}$  is an  $\mathbb{R}^q$ -valued random vector, both defined on a probability space  $(\Omega, \mathcal{A}, \mathbb{P})$ . The distribution  $\mu$  of  $X_0$  is given, as well as the transition probabilities  $P_k(x, dx')$  of the process  $(X_k)_{k \geq 0}$ . We also suppose that the process  $(X_k, Y_k)_{k \geq 0}$  is a Markov chain and that for every  $k \geq 1$ , the conditional distribution of  $Y_k$ , given  $(X_{k-1}, Y_{k-1}, X_k)$  has a density  $g_k(X_{k-1}, Y_{k-1}, X_k, \cdot)$ . Having a fixed observation  $(y_0, \dots, y_n) = (y_0, \dots, y_n)$ , for  $n \geq 1$ , we aim at computing the conditional distribution  $\Pi_{y,n}$  of  $X_n$  given  $y = (y_0, \dots, y_n)$ . It is well-known that for any bounded and measurable function  $f$ , we can derive the following formula (see e.g. [41])

$$\Pi_{y,n} f = \frac{\pi_{y,n} f}{\pi_{y,n} \mathbf{1}} \quad (47)$$

where the so-called un-normalized filter  $\pi_{y,n}$  is defined for every bounded or non-negative Borel function  $f$  by

$$\pi_{y,n} f = \mathbb{E}(f(X_n) L_{y,n})$$

with

$$L_{y,n} = \prod_{k=1}^n g_k(X_{k-1}, y_{k-1}, X_k, y_k).$$

Defining the family of transition kernels  $H_{y,k}$ ,  $k = 1, \dots, n$ , by

$$H_{y,k} f(x) = \mathbb{E}(f(X_k) g_k(x, y_{k-1}, X_k, y_k) | X_{k-1} = x) \quad (48)$$

for every bounded Borel function  $f : \mathbb{R}^d \rightarrow \mathbb{R}$  and setting

$$H_{y,0} f(x) = \mathbb{E}(f(X_0)),$$

one shows that the un-normalized filter may be computed by the following *forward* induction formula:

$$\pi_{y,k} f = \pi_{y,k-1} H_{y,k} f, \quad k = 1, \dots, n, \quad (49)$$

with  $\pi_{y,0} = H_{y,0}$ . A useful formulation in order to establish the quantization error bound turns out to be the *backward* induction formula defined by setting

$$\pi_{y,n} f = u_{y,-1}(f)$$

where  $u_{y,-1}$  is the final value of the induction formula:

$$\begin{aligned} u_{y,n}(f)(x) &= f(x), \\ u_{y,k-1}(f) &= H_{y,k} u_{y,k}(f), \quad k = 0, \dots, n. \end{aligned} \quad (50)$$

In order to compute the normalized filter  $\Pi_{y,n}$ , we just have to compute the transition kernels  $H_{y,k}$  and to use the recursive formula (49) or (50). However these kernels have no closed formula in general so that we have to approximate them. Optimal quantization is an adequate method to proceed (see e.g. [12, 41, 50]) owing to its tractability. Then, denoting by  $\widehat{X}_k$  a quantization of  $X_k$  at level  $N_k$  by the grid  $\Gamma_k = \{x_k^1, \dots, x_k^{N_k}\}$ , for every  $k = 0, \dots, n$ , we will formally replace  $X_k$  in (49) or (50) by  $\widehat{X}_k$ . As a consequence the (optimally) quantized approximation  $\widehat{\pi}_{y,n}$  of  $\pi_{y,n}$  will be defined by the recursive formula

$$\widehat{\pi}_{y,k} = \widehat{\pi}_{y,k-1} \widehat{H}_{y,k}, \quad k = 1, \dots, n \quad \text{and} \quad \widehat{\pi}_{y,0} = \widehat{H}_{y,0} \quad (51)$$

where

$$\begin{cases} \hat{\pi}_{y,k} = \hat{\pi}_{y,k-1} \hat{H}_{y,k} := \left[ \sum_{i=1}^{N_{k-1}} \hat{H}_{y,k}^{i,j} \hat{\pi}_{y,k-1}^i \right]_{j=1,\dots,N_k}, & k = 1, \dots, n, \\ \text{with } \hat{H}_{y,k}^{ij} = g_k(x_{k-1}^i, y_{k-1}; x_k^j, y_k) \hat{p}_k^{ij}, & i = 1, \dots, N_{k-1}; j = 1, \dots, N_k. \end{cases} \quad (52)$$

As a final step, we approximate the normalized filter  $\Pi_{y,n}$  by  $\hat{\Pi}_{y,n}$  which is given by

$$\hat{\Pi}_{y,n} = \sum_{i=1}^{N_n} \hat{\Pi}_{y,n}^i \delta_{x_n^i}$$

with

$$\hat{\Pi}_{y,n}^i = \frac{\hat{\pi}_{y,n}^i}{\sum_{j=1}^{N_n} \hat{\pi}_{y,n}^j}, \quad i = 1, \dots, N_n.$$

Our aim is then to estimate the quantization error induced by the approximation of  $\Pi_{y,n}$  by  $\hat{\Pi}_{y,n}$ . Note that this problem has been considered in [41] where it has been shown that for every bounded function  $f$ , the absolute error  $|\Pi_{y,n}f - \hat{\Pi}_{y,n}f|$  is bounded (up to a constant depending in particular on  $n$ ) by the cumulated  $L^r$ -quantization errors  $\|X_k - \hat{X}_k\|_r$  from  $k = 0$  to  $n$ . In this work, we improve this result in the particular case of the quadratic quantization framework (*i.e.*  $r = 2$ ) in two directions. In fact, we first show that, for every bounded Borel function  $f$ , the squared-absolute error  $|\Pi_{y,n}f - \hat{\Pi}_{y,n}f|^2$  is bounded by the cumulated square-quadratic quantization errors  $\|X_k - \hat{X}_k\|_2^2$  from  $k = 0$  to  $n$ . Secondly, we show this improved result (asymptotically with respect to the quantization grid sizes  $N_k$ ) under weaker assumptions on the density functions  $g_k$  taking advantage of a so-called distortion mismatch property, originally introduced and investigated in [24], and improved Theorem 4.3 proved in Appendix B hereafter, which shows that a sequence of  $L^r$ -optimal quantizers can be “rate-optimal” (in the Zador sense) when viewed as  $L^{r+\nu}$ -quantizer when  $\nu \in (0, d)$ .

## 5.2 Error analysis

Let us consider first the following assumptions from [41].

( $\mathcal{H}_0$ )  $\equiv$  For every  $k \in \{1, \dots, n\}$  their exists  $[g_k^1]_{\text{Lip}}, [g_k^2]_{\text{Lip}} : \mathbb{R}^q \times \mathbb{R}^q \mapsto \mathbb{R}_+$  such that

$$|g_k(x, y, x', y') - g_k(\hat{x}, y, \hat{x}', y')| \leq [g_k^1]_{\text{Lip}}(y, y')|x - \hat{x}| + [g_k^2]_{\text{Lip}}(y, y')|x' - \hat{x}'|.$$

( $\mathcal{A}_1$ )  $\equiv$  (i) The Markov transition operators  $P_k(x, dx')$ ,  $k = 1, \dots, n$  propagate Lipschitz continuity (in the sense of Lemma 3.3) and

$$[P]_{\text{Lip}} := \max_{k=1,\dots,n} [P_k]_{\text{Lip}} < +\infty.$$

(ii) For every  $k = 1, \dots, n$ , the functions  $g_k$  are bounded on  $\mathbb{R}^d \times \mathbb{R}^q \times \mathbb{R}^d \times \mathbb{R}^q$  and we set

$$K_g := \max_{k=1,\dots,n} \|g_k\|_{\text{sup}} < +\infty.$$

Furthermore, let us consider, for a fixed Borel function  $\theta : \mathbb{R}^d \mapsto \mathbb{R}_+$ , the following  $\theta$ -local Lipschitz continuity assumption (which is weaker than ( $\mathcal{H}_0$ )) on the growth of the functions  $g_k$ :

( $\mathcal{H}_{\text{Liploc}}$ )  $\equiv$  There exists a non-negative convex function  $\theta : \mathbb{R}^d \mapsto \mathbb{R}_+$  such that, for every  $k \in \{1, \dots, n\}$ ,  $\mathbb{E}(\theta(X_k)) < +\infty$  and there exists  $[g_k^1]_{\text{Liploc}}, [g_k^2]_{\text{Liploc}} : \mathbb{R}^q \times \mathbb{R}^q \mapsto \mathbb{R}_+$  such that

$$\begin{aligned} |g_k(x, y, x', y') - g_k(\hat{x}, y, \hat{x}', y')| &\leq [g_k^1]_{\text{Liploc}}(y, y')(1 + \theta(x) + \theta(\hat{x}) + \theta(x') + \theta(\hat{x}'))|x - \hat{x}| \\ &+ [g_k^2]_{\text{Liploc}}(y, y')(1 + \theta(x) + \theta(\hat{x}) + \theta(x') + \theta(\hat{x}'))|x' - \hat{x}'|. \end{aligned}$$



Remark that if the following assumption holds:

$(\mathcal{H}_\alpha) \equiv$  For every  $k \in \{1, \dots, n\}$  their exists  $[g_k^1]_{\text{pol}}, [g_k^2]_{\text{pol}} : \mathbb{R}^q \times \mathbb{R}^q \mapsto \mathbb{R}_+$  such that

$$\begin{aligned} |g_k(x, y, x', y') - g_k(\widehat{x}, y, \widehat{x}', y')| &\leq [g_k^1]_{\text{pol}}(y, y')(1 + |x|^\alpha + |\widehat{x}|^\alpha + |x'|^\alpha + |\widehat{x}'|^\alpha)|x - \widehat{x}| \\ &\quad + [g_k^2]_{\text{pol}}(y, y')(1 + |x|^\alpha + |\widehat{x}|^\alpha + |x'|^\alpha + |\widehat{x}'|^\alpha)|x' - \widehat{x}'|, \end{aligned}$$

for  $\alpha > 1$ , then  $\mathcal{H}_{\text{Liploc}}$  holds with  $\theta : x \mapsto \theta(x) = |x|^\alpha$ .

The following classical lemma is borrowed from [41] (Lemma 3.1).

**Lemma 5.1.** *Let  $\mu_y$  and  $\vartheta_y$  be two families of finite and positive measure on a measurable space  $(E, \mathcal{E})$ . Suppose that their exists two symmetric functions  $R$  and  $S$  defined on the set of positive finite measures such that for every bounded Lipschitz function  $f$ ,*

$$\left| \int f d\mu_y - \int f d\vartheta_y \right| \leq \|f\|_\infty^2 R(\mu_y, \vartheta_y) + [f]_{\text{Lip}}^2 S(\mu_y, \vartheta_y). \quad (53)$$

Then,

$$\left| \frac{\int f d\mu_y}{\mu_y(E)} - \frac{\int f d\vartheta_y}{\vartheta_y(E)} \right| \leq 2\|f\|_\infty R(\mu_y, \vartheta_y)^2 + 2\|f\|_\infty^2 R(\mu_y, \vartheta_y) + 2[f]_{\text{Lip}}^2 S(\mu_y, \vartheta_y). \quad (54)$$

In Theorem 5.3 below we will consider the assumption  $(\mathcal{H}_{\text{Liploc}})$  in place of Assumption  $(\mathcal{H}_0)$  (which has been considered [41]) to derive an error bound. This less stringent assumption is counter-balanced by taking advantage of the *distortion mismatch* (see Theorem 4.3 and Appendix B satisfied by sequences of optimal quantizers for wide classes of distributions. More precisely, we need that the  $L^{2+\nu}$ -mean quantization error associated to any sequence of optimal quadratic quantizers at level  $N$  still goes to zero at the optimal rate  $N^{-\frac{1}{d}}$ .

We first need first to control the  $\theta$ -local Lipschitz constants of  $[u_k]_{\text{Liploc}}$ , for every  $k \geq 0$  (where  $\theta : \mathbb{R}^d \rightarrow \mathbb{R}_+$  is defined as above). If we suppose that  $f : \mathbb{R}^d \rightarrow \mathbb{R}$  is  $\theta$ -locally Lipschitz with a local Lipschitz coefficient  $[f]_{\text{Liploc}}$  defined by

$$[f]_{\text{Liploc}} = \sup_{x \neq x'} \frac{|f(x) - f(x')|}{(1 + \theta(x) + \theta(x'))|x - x'|} < +\infty \quad (55)$$

then for every  $k = 0, \dots, n$ ,  $P_k f$  is  $\theta$ -locally Lipschitz with  $[P_k f]_{\text{Liploc}} \leq [P_k]_{\text{Liploc}} [f]_{\text{Liploc}}$ . The following lemma provides a control of the  $\theta$ -local Lipschitz coefficients of the  $u_k$ 's.

**Lemma 5.2.** *Suppose that Assumption  $(\mathcal{H}_{\text{Liploc}}) \equiv$  holds and that for every  $k = 0, \dots, n-1$ ,*

$$\mathbb{E}(\theta(X_{k+1}) | X_k = x) \leq C_{\theta, X}(1 + \theta(x)), \quad (56)$$

then, for every  $k = 0, \dots, n-1$ ,

$$[u_k]_{\text{Liploc}} \leq ([g_k^1]_{\text{Liploc}} C_{\theta, X} + [P_{k+1}]_{\text{Liploc}} [g_k^2]_{\text{Liploc}}) K_g^{n-k} \|f\|_\infty + K_g [P_{k+1}]_{\text{Liploc}} [u_{k+1}]_{\text{Liploc}}. \quad (57)$$

If furthermore,

$$X_{k+1} = F_k(X_k, \varepsilon_{k+1}), \quad k = 0, \dots, n-1$$

where  $(\varepsilon_k)$  is an i.i.d sequence of random variables independent from  $X_0$  and for every  $x \in \mathbb{R}^d$ ,

$$\mathbb{E}\theta(F_k(x, \varepsilon_{k+1})) \leq C'_{\theta, X}(1 + \theta(x)), \quad \forall k = 0, \dots, n-1, \quad (58)$$

for a positive real valued and convex function  $\theta$ , then

$$[P_{k+1}]_{\text{Liploc}} = C'_{\theta, X}. \quad (59)$$

*Proof.* We have for every  $k = 0, \dots, n-1$ ,

$$u_k(X_k) = \mathbb{E}_k(u_{k+1}(X_{k+1})g_{k+1}(X_k, X_{k+1})) = (P_{k+1}u_{k+1}g_{k+1}(x, \cdot))(x).$$

It follows that for every  $k = 0, \dots, n-1$ ,  $\|u_k\|_\infty \leq K_g \|u_{k+1}\|_\infty$ , so that,

$$\|u_k\|_\infty \leq K_g^{n-k} \|f\|_\infty$$

since  $\|u_n\|_\infty \leq \|f\|_{\text{infly}}$ . On the other hand we have for every  $k = 0, \dots, n-1$ ,  $x, x' \in \mathbb{R}^d$ ,

$$\begin{aligned} |u_k(x) - u_k(x')| &\leq [g_k^1]_{\text{Liploc}} \|u_{k+1}\|_\infty (1 + \theta(x) + \theta(x') + \mathbb{E}(\theta(X_{k+1})|X_k = x)) |x - x'| \\ &\quad + [P_{k+1}]_{\text{Liploc}} [u_{k+1}g_{k+1}(x', \cdot)]_{\text{Liploc}} (1 + \theta(x) + \theta(x')) |x - x'|. \end{aligned}$$

Now, for every  $k = 0, \dots, n-1$ ,

$$\begin{aligned} |u_{k+1}(z)g_{k+1}(x', z) - u_{k+1}(z')g_{k+1}(x', z')| &\leq |u_{k+1}(z) - u_{k+1}(z')|g_{k+1}(x', z) \\ &\quad + |g_{k+1}(x', z) - g_{k+1}(x', z')| |u_{k+1}(z')| \\ &\leq K_g [u_{k+1}]_{\text{Liploc}} (1 + \theta(z) + \theta(z')) |z - z'| \\ &\quad + \|u_{k+1}\|_{\text{sup}} [g_k^2]_{\text{Liploc}} (1 + \theta(z) + \theta(z')) |z - z'| \end{aligned}$$

so that

$$[u_{k+1}g_{k+1}(x', \cdot)]_{\text{Liploc}} \leq K_g [u_{k+1}]_{\text{Liploc}} + \|u_{k+1}\|_\infty [g_k^2]_{\text{Liploc}}.$$

Therefore, we deduce from assumption (56) that

$$[u_k]_{\text{Liploc}} \leq ([g_k^1]_{\text{Liploc}} C_{\theta, X} + [P_{k+1}]_{\text{Liploc}} [g_k^2]_{\text{Liploc}}) \|u_{k+1}\|_\infty + K_g [P_{k+1}]_{\text{Liploc}} [u_{k+1}]_{\text{Liploc}}.$$

On the other hand, as  $f$  is  $\theta$ -locally Lipschitz with  $\theta$ -local Lipschitz constant  $[f]_{\text{Lip}}$  then, for every  $x, x' \in \mathbb{R}^d$  and  $k = 0, \dots, n-1$ ,

$$\begin{aligned} |P_{k+1}f(x) - P_{k+1}f(x')| &= |\mathbb{E}f(F_k(x, \varepsilon_{k+1})) - \mathbb{E}f(F_k(x', \varepsilon_{k+1}))| \\ &\leq [f]_{\text{Liploc}} |x - x'| (1 + \mathbb{E}\theta(F_k(x, \varepsilon_{k+1})) + \mathbb{E}\theta(F_k(x', \varepsilon_{k+1}))). \end{aligned}$$

Then, owing to Assumption (58), we deduce that  $[P_{k+1}]_{\text{Liploc}} = C'_{\theta, X}$ , for every  $k = 0, \dots, n-1$ .  $\square$

Notice that assumptions (56) and (58) hold when  $\theta$  is a polynomial and convex function and when the process  $(X_k)_{0 \leq k \leq n}$  is the Euler scheme (where  $n$  is the length of the time discretization mesh) associate with a stochastic differential equation of the form (11). In the latter case, the transition operator  $P_{k+1} = P$ ,  $k = 0, \dots, n-1$ , is time homogenous. We suppose in the sequel that  $(P_k)$  is time homogenous.

**Theorem 5.3.** *Let  $(\mathcal{A}_{\text{liploc}})$  holds and assume that  $\mathcal{A}_1$  is fulfilled, as well as assumptions of Lemma 5.2. Suppose that for every  $k = 0, \dots, n$ ,  $X_k$  has an  $(2, 2 + \nu_k)$ -distribution,  $\nu_k > 0$ , and set  $\nu_n = \min_{k=0, \dots, n} \nu_k/2$ . Then for every  $\nu \in (0, \nu_n)$ ,*

$$|\Pi_{y,n}f - \widehat{\Pi}_{y,n}f|^2 \leq \frac{2M_\nu^n (K_g^n)^2}{\phi_n^2(y) \vee \widehat{\phi}_n^2(y)} \sum_{k=0}^n B_k^n(f, y, \alpha) \|X_k - \widehat{X}_k\|_{2+2\nu}^2 \quad (60)$$

with

$$\phi_n(y) = \pi_{y,n} \mathbf{1} \quad \text{and} \quad \widehat{\phi}_n(y) = \widehat{\pi}_{y,n} \mathbf{1}$$

and where

$$B_k^n(f, y) := 2[P]_{\text{Liploc}}^{2(n-k)} [f]_{\text{Liploc}}^2 + 2\|f\|_\infty^2 R_{n,k} + \|f\|_\infty R_{n,k}^2,$$

with

$$R_{n,k} = \frac{8^{q\nu}}{K_g^2} \left[ [g_{k+1}^1]_{\text{Liploc}}^2 + [g_k^2]_{\text{Liploc}}^2 + \left( \sum_{m=1}^{n-k} [P]_{\text{Liploc}}^{m-1} ([g_{k+m}^1]_{\text{Liploc}} + [P]_{\text{Liploc}} [g_{k+m}^2]_{\text{Lip}}) \right)^2 \right],$$

$q_\nu = 1 + 1/\nu$  and

$$M_\nu^n := \max_{k=0, \dots, n-1} (\mathbb{E}(\theta(X_k)^{2q\nu}) + \mathbb{E}(\theta(\widehat{X}_k)^{2q\nu}) + \mathbb{E}(\theta(X_{k+1})^{2q\nu}) + \mathbb{E}(\theta(\widehat{X}_{k+1})^{2q\nu})).$$

Before dealing with the proof let us remark that if  $\widehat{X}_k$  is an optimal quantizer then it is a stationary quantizer so that, for every  $k = 0, \dots, n$ , we have

$$\mathbb{E} \theta(\widehat{X}_k)^{2q\nu} \leq \mathbb{E} \theta(X_k)^{2q\nu} < +\infty.$$

*Proof.* To simplify the notations, we will omit the dependence of the used functions from the observation parameter  $y$ . For every  $k \in \{0, \dots, n-1\}$ , we define the function  $\varphi_{k+1}$  by

$$\varphi_{k+1}(x_k, x_{k+1}, x'_{k+1}) := g_{k+1}(x_k, x_{k+1})u_{k+1}(x'_{k+1}), \quad x_k, x_{k+1}, x'_{k+1} \in \mathbb{R}^d.$$

Following the lines of the proof of Theorem 3.1 in [41], one shows by a backward induction taking advantage of the Markov property that there exists functions  $u_k : \mathbb{R}^d \rightarrow \mathbb{R}$ ,  $k = 0, \dots, n-1$  such that  $u_n = f$  and

$$u_k(X_k) = \mathbb{E}_k(\varphi_{k+1}(X_k, X_{k+1}, X_{k+1})) = \mathbb{E}_k(u_{k+1}(X_{k+1})g_{k+1}(X_k, X_{k+1})). \quad (61)$$

Now, using the definition of conditional expectation  $\widehat{\mathbb{E}}_k$  as an orthogonal projector (hence an  $L^2$ -contraction as well), we have

$$\begin{aligned} \|u_k(X_k) - \widehat{u}_k(\widehat{X}_k)\|_2^2 &= \|\mathbb{E}_k(u_k(X_k) - \widehat{\mathbb{E}}_k(\varphi_{k+1}(\widehat{X}_k, \widehat{X}_{k+1}, X_{k+1}))\|_2^2 \\ &\quad + \|\widehat{\mathbb{E}}_k(\varphi_{k+1}(\widehat{X}_k, \widehat{X}_{k+1}, X_{k+1})) - \widehat{\mathbb{E}}_k(\varphi_{k+1}(\widehat{X}_k, \widehat{X}_{k+1}, \widehat{X}_{k+1}))\|_2^2 \\ &\leq \|\mathbb{E}_k(u_k(X_k) - \widehat{\mathbb{E}}_k(\varphi_{k+1}(\widehat{X}_k, \widehat{X}_{k+1}, X_{k+1}))\|_2^2 \\ &\quad + K_g^2 \|u_{k+1}(X_{k+1}) - \widehat{u}_{k+1}(\widehat{X}_{k+1})\|_2^2. \end{aligned}$$

Considering the first term on the right hand side of the previous inequality and using once again the orthogonality property of the conditional expectation, we obtain

$$\|\mathbb{E}_k(u_k(X_k) - \widehat{\mathbb{E}}_k(\varphi_{k+1}(\widehat{X}_k, \widehat{X}_{k+1}, X_{k+1}))\|_2^2 \leq \|u_k(X_k) - \widehat{\mathbb{E}}_k(u_k(X_k))\|_2^2 + \|\widehat{\mathbb{E}}_k(u_k(X_k)) - \widehat{\mathbb{E}}_k(\varphi_{k+1}(\widehat{X}_k, \widehat{X}_{k+1}, X_{k+1}))\|_2^2.$$

It follows now from the definition of the conditional expectation  $\widehat{\mathbb{E}}_k(\cdot)$  as the best approximation in  $L^2$  among square integrable  $\sigma(\widehat{X}_k)$ -measurable random vectors that

$$\|u_k(X_k) - \widehat{\mathbb{E}}_k(u_k(X_k))\|_2^2 \leq \|u_k(X_k) - u_k(\widehat{X}_k)\|_2^2 \leq [u_k]_{\text{Liploc}}^2 \|(1 + \theta(X_k) + \theta(\widehat{X}_k))(X_k - \widehat{X}_k)\|_2^2.$$

Let  $\nu \in (0, \nu_n)$ , so that for every  $k = 0, \dots, n$ ,  $2 + 2\nu \leq 2 + \nu_k$ . The Hölder inequality with exponent  $p_\nu = 1 + \nu$  and  $q_\nu = 1 + \frac{1}{\nu}$  gives

$$\|u_k(X_k) - \widehat{\mathbb{E}}_k(u_k(X_k))\|_2^2 \leq [u_k]_{\text{Liploc}}^2 (1 + \mathbb{E}\theta(X_k)^{2q\nu} + \mathbb{E}\theta(\widehat{X}_k)^{2q\nu}) \|X_k - \widehat{X}_k\|_{2+2\nu}^2.$$

On the other hand, setting  $R_k = \widehat{\mathbb{E}}_k(u_k(X_k)) - \widehat{\mathbb{E}}_k(\varphi_{k+1}(\widehat{X}_k, \widehat{X}_{k+1}, X_{k+1}))$ , it follows from Assumption ( $\mathcal{A}_{\text{liploc}}$ ) that

$$\begin{aligned} \|R_k\|_2^2 &\leq 2\|u_{k+1}\|_\infty^2 \left( [g_{k+1}^2]_{\text{Liploc}}^2 \mathbb{E} \left( (1 + \theta(X_k) + \theta(\widehat{X}_k) + \theta(X_{k+1}) + \theta(\widehat{X}_{k+1}))^2 |X_{k+1} - \widehat{X}_{k+1}|^2 \right) \right. \\ &\quad \left. + [g_{k+1}^1]_{\text{Liploc}}^2 \mathbb{E} \left( (1 + \theta(X_k) + \theta(\widehat{X}_k) + \theta(X_{k+1}) + \theta(\widehat{X}_{k+1}))^2 |X_k - \widehat{X}_k|^2 \right) \right). \quad (62) \end{aligned}$$

Once again, the Hölder inequality with exponent  $p_\nu = 1 + \nu$  and  $q_\nu = 1 + \frac{1}{\nu}$  yields

$$\|R_k\|_2^2 \leq 2 \times 8^{q_\nu} M_\nu^k \|u_{k+1}\|_\infty^2 \left( [g_{k+1}^1]_{\text{Liploc}}^2 \|X_k - \widehat{X}_k\|_{2+2\nu}^2 + [g_{k+1}^2]_{\text{Liploc}}^2 \|X_k - \widehat{X}_k\|_{2+2\nu}^2 \right)$$

where  $M_\nu^k := 1 + \mathbb{E}(\theta(X_k)^{2q_\nu}) + \mathbb{E}(\theta(\widehat{X}_k)^{2q_\nu}) + \mathbb{E}(\theta(X_{k+1})^{2q_\nu}) + \mathbb{E}(\theta(\widehat{X}_{k+1})^{2q_\nu})$ .

We finally get that, for every  $k = 0, \dots, n-1$ ,

$$\|u_k(X_k) - \widehat{u}_k(\widehat{X}_k)\|_2^2 \leq K_g^2 \|u_{k+1}(X_{k+1}) - \widehat{u}_{k+1}(\widehat{X}_{k+1})\|_2^2 + \alpha_k \|X_k - \widehat{X}_k\|_{2+2\nu}^2 + \beta_{k+1} \|X_{k+1} - \widehat{X}_{k+1}\|_{2+2\nu}^2$$

where

$$\alpha_k := [u_k]_{\text{Liploc}}^2 M_\nu^k + 2 \times 8^{q_\nu} M_\nu^k \|u_{k+1}\|_\infty^2 [g_{k+1}^1]_{\text{Liploc}}^2, \quad 0 \leq k \leq n-1$$

and

$$\beta_k := 2 \times 8^{q_\nu} M_\nu^k \|u_{k+1}\|_\infty^2 [g_{k+1}^2]_{\text{Liploc}}^2, \quad 1 \leq k \leq n.$$

Setting,  $\alpha_n := [f]_{\text{Liploc}}^2$ , we show by induction that, for every  $k \in \{0, \dots, n\}$

$$|\pi_{y,k} f - \widehat{\pi}_{y,k} f|^2 \leq \sum_{\ell=k}^n C_\ell^n(f, y) \|X_\ell - \widehat{X}_\ell\|_{2+2\nu}^2$$

where

$$\begin{aligned} C_\ell^n(f, y) &:= K_g^{2(\ell-1)} (\alpha_\ell K_g^2 + \beta_\ell) \\ &= M_\mu^k \left( [u_\ell]_{\text{Liploc}}^2 K_g^{2\ell} + 2 \times 8^{q_\nu} K_g^{2(\ell-1)} \|f\|_\infty^2 ([g_{\ell+1}^1]_{\text{Liploc}}^2 + [g_\ell^2]_{\text{Liploc}}^2) \right) \\ &\leq 2M_\mu^k K_g^{2n} \left[ [P]_{\text{Liploc}}^{2(n-\ell)} [f]_{\text{Lip}}^2 + 8^{q_\nu} \frac{\|f\|_\infty^2}{K_g^2} \left( [g_{\ell+1}^1]_{\text{Lip}}^2 + [g_\ell^2]_{\text{Lip}}^2 \right) \right. \\ &\quad \left. + \left( \sum_{m=1}^{n-\ell} [P]_{\text{Liploc}}^{m-1} ([g_{\ell+m}^1]_{\text{Liploc}} C_{\theta, X} + [P]_{\text{Liploc}} [g_{\ell+m}^2]_{\text{Liploc}}) \right)^2 \right]. \end{aligned}$$

We conclude by Lemma 5.1.  $\square$

The previous theorem shows the usefulness of the distortion mismatch result, which in different contexts may be used to improve several results involving the quantization errors in the asymptotic framework. In our context, it allows us to weaken the assumptions on the functions  $g_k$  whereas for the maximal radius problem (see [43, 29]), its use is crucial to derive the sharp constant for the asymptotic of the maximal radius sequence of quantizers when considering distributions with radial exponential tails.

**Remark 5.1.** Note that if we consider Assumption  $(\mathcal{H}_0)$  instead of Assumption  $\mathcal{H}_{\text{liploc}}$  in Theorem 5.3, one shows that for every bounded Lipschitz continuous function  $f$  on  $\mathbb{R}^d$ ,

$$|\Pi_{y,n} f - \widehat{\Pi}_{y,n} f| \leq \frac{2K_g^n}{\phi_n(y) \vee \widehat{\phi}_n(y)} \left( \sum_{k=0}^n B_k^n(f, y) \|X_k - \widehat{X}_k\|_2^2 \right)^{\frac{1}{2}} \quad (63)$$

where  $\phi_n(y) = \pi_{y,n} \mathbf{1}$  and  $\widehat{\phi}_n(y) = \widehat{\pi}_{y,n} \mathbf{1}$  and

$$B_k^n(f, y) := 2[P]_{\text{Lip}}^{2(n-k)} [f]_{\text{Lip}}^2 + 2\|f\|_\infty^2 R_{n,k} + \|f\|_\infty R_{n,k}^2,$$

with

$$R_{n,k} = \frac{1}{K_g^2} \left[ [g_{k+1}^1]_{\text{Lip}}^2 + [g_k^2]_{\text{Lip}}^2 + \left( \sum_{m=1}^{n-k} [P]_{\text{Lip}}^{m-1} ([g_{m+k}^1]_{\text{Lip}} + [P]_{\text{Lip}} [g_{m+k}^2]_{\text{Lip}}) \right)^2 \right].$$

In fact, in this case, the upper bound (62) may be replaced by

$$\begin{aligned} \|R_k\|_2^2 &\leq \|\varphi_{k+1}(X_k, X_{k+1}, X_{k+1}) - \varphi_{k+1}(\widehat{X}_k, \widehat{X}_{k+1}, X_{k+1})\|_2^2 \\ &\leq 2\|u_{k+1}\|_\infty^2 \left( [g_{k+1}^1]_{\text{Lip}}^2 \|X_k - \widehat{X}_k\|_2^2 + [g_{k+1}^2]_{\text{Lip}}^2 \|X_{k+1} - \widehat{X}_{k+1}\|_2^2 \right). \end{aligned}$$

## References

- [1] K. Bahlali. Existence and uniqueness of solutions for BSDEs with locally lipschitz coefficients. *Elect. Comm. in Probab.*, 7:169–179, 2002.
- [2] V. Bally. An approximation scheme for BSDEs and applications to control and nonlinear PDEs. *Prepublication 95-15 du Laboratoire de Statistique et Processus de l'Université du Maine*, 1995.
- [3] V. Bally and G. Pagès. Error analysis of the quantization algorithm for obstacle problems. *Stochastic Processes & Their Applications*, 1:1–40, 2003.
- [4] V. Bally and G. Pagès. A quantization algorithm for solving discrete time multidimensional optimal stopping problems. *Bernoulli*, 6:1003–1049, 2003.
- [5] V. Bally, G. Pagès, and J. Printems. A stochastic quantization method for non-linear problems. *Monte Carlo Methods and Appl.*, 1:21–34, 2001.
- [6] V. Bally, G. Pagès, and J. Printems. A quantization tree method for pricing and hedging multidimensional american options. *Mathematical Finance*, 15:119–168, 2005.
- [7] C. Bender and J. Steiner. Least-squares Monte Carlo for backward SDEs. *Numerical Methods in Finance*, 12:257–289, 2012.
- [8] Y. Z. Bergman. Option pricing with differential interest rates. *Review of Financial Studies.*, 8(2):475–500, 1995.
- [9] J. M. Bismut. Conjugate convex functions in optimal stochastic control. *J. Math. Anal. Appl.*, 44:383–404, 1973.
- [10] B. Bouchard and N. Touzi. Discrete-time approximation and Monte-Carlo simulation of backward stochastic differential equations. *Stochastic Process. Appl.*, 111(2):175–206, 2004.
- [11] N. Bouleau and D. Lépingle. *Numerical Methods for Stochastic Processes*. Wiley-Interscience, 1993.
- [12] G. Callegaro and A. Sagna. An application to credit risk of a hybrid Monte Carlo-optimal quantization method. *The Journal of Computational Finance*, 16:123–156, 2013.
- [13] J. F. Chassagneux and D. Crisan. Runge kutta schemes for BSDEs. *Annals of Applied Probability*, To appear.
- [14] J. F. Chassagneux and A. Richou. Numerical simulation of quadratic BSDEs. *Preprint*, 2013.
- [15] E.V. Chernaya. On the optimization of weighted cubature formulae on certain classes of continuous functions. *East J. Approx.*, 1:47–60, 1995.
- [16] D. Crisan, K. Manolarakis, and N. Touzi. On the monte carlo simulation of bsde's: an improvement on the malliavin weights. *Stochastic Processes and their Applications*, 120:1133–1158, 2010.
- [17] F. Delarue and S. Menozzi. A forward-backward stochastic algorithm for quasi-linear pdes. *The Annals of Applied Probabilities*, 16(1):140–184, 2006.
- [18] F. Delarue and S. Menozzi. An interpolated stochastic algorithm for quasi-linear pdes. *Math. Comp.*, 77(26):125–158, 2008.

- [19] J. Diehl and P. Friz. Backward stochastic differential equations with rough drivers. *The Annals of Probability*, 40(4):1715–1758, 2012.
- [20] J. Jr. Douglas, J. Ma, and P. Protter. Numerical methods for forward-backward stochastic differential equations. *Ann. Appl. Probab.*, 6:940–968, 1996.
- [21] A. Gersho and R. Gray. *Vector Quantization and Signal Compression*. Kluwer Academic Press, Boston, 1992.
- [22] E. Gobet, J. P. Lemor, and X. Warin. A regression-based Monte Carlo method to solve backward stochastic differential equations. *Ann. Appl. Probab.*, 15(3):2172–2202, 2005.
- [23] S. Graf and H. Luschgy. *Foundations of Quantization for Probability Distributions*. Lect. Notes in Math. 1730. Springer, Berlin., 2000.
- [24] Siegfried Graf, Harald Luschgy, and Gilles Pagès. Distortion mismatch in the quantization of probability measures. *ESAIM Probab. Stat.*, 12:127–153, 2008.
- [25] Luschgy H. and Pagès G. *Marginal and Functional Quantization of Stochastic Processes*. book, in progress.
- [26] S. Hamadène. Equations différentielles stochastiques rétrogrades: le cas localement lipschitz. *Ann. Inst. Henri Poincaré*, 32(5):645–659, 1996.
- [27] Y. Hu, D. Nualart, and X. Song. Malliavin calculus for backward stochastic differential equations and applications to numerical solutions. *The Annals of Applied Probability*, 21(6):2379–2423, 2011.
- [28] C. Illand. *Contrôle stochastique par quantification et application à la finance*. PhD thesis, Université Pierre et Marie Curie, 2012.
- [29] S. Junglen. Quantization balls and asymptotics of quantization radii for probability distributions with radial exponential tails. *Elect. Comm. in Probab.*, 16:283–295, 2011.
- [30] I. Kharroubi and R. Elie. Adding constraints to BSDEs with jumps: an alternative to multidimensional reflections. *ESAIM, Probability and Statistics*, 18:233–250, 2014.
- [31] I. Kharroubi, M. Jin, H. Pham, and J. Zhang. Backward SDEs with constrained jumps and quasi-variational inequalities. *The Annals of Probability*, 38(2):794–840, 2010.
- [32] M. Kobylanski. Backward stochastic differential equations and partial differential equations with quadratic growth. *Annals of Probability*, 28:558–602, 2000.
- [33] J.-P. Lepeltier and J. San Martin. On the existence of BSDE with superlinear quadratic coefficient. *Stochastics and Stochastic Report*, 63:227–240, 1998.
- [34] H. Luschgy and G. Pagès. Functional quantization rate and mean regularity of processes with an application to lévy processes. *Annals of Applied Probability*, 18(2):427–469, 2008.
- [35] J. Ma, P. Protter, and J. Yong. Solving forward-backward stochastic differential equations explicitly-a four step scheme. *Probab. Theory Related Fields*, 98:339–359, 1994.
- [36] J. Ma and J. Yong. *Forward-Backward Stochastic Differential Equations and their Applications*. Lect. Notes in Math. 1702. Springer., 2007.

- [37] H. Meter Soner, N. Touzi, and J. Zhang. Wellposedness of second order backward SDEs. *Probab. Theory Relat. Fields*, 153:149–190, 2012.
- [38] El Karoui N., Peng S., and M. C. Quenez. Backward stochastic differential equations in finance. *Mathematical Finance*, 7(1):1–71, 1997.
- [39] G. Pagès. A space vector quantization method for numerical integration. *J. Computational and Applied Mathematics*, 89:1–38, 1998.
- [40] G. Pagès. Introduction to optimal vector quantization and its applications to numerics. *ESAIM Proc. & Survey*, 48:29–79, 2015.
- [41] G. Pagès and H. Pham. Optimal quantization methods for nonlinear filtering with discrete time observations. *Bernoulli*, 11(5):893–932, 2005.
- [42] G. Pagès and J. Printems. Optimal quadratic quantization for numerics: the gaussian case. *Monte Carlo Methods and Applications*, 9(2):135–165, 2003.
- [43] G. Pagès and A. Sagna. Asymptotics of the maximal radius of an  $L^r$ -optimal sequence of quantizers. *Bernoulli*, 18(1):360–389, 2012.
- [44] G. Pagès and B. Wilbertz. *Numerical Methods in Finance*, chapter Optimal Delaunay and Voronoi Quantization Schemes for Pricing American Style Options, pages 171–213. Springer Proceedings in Mathematics, 2012.
- [45] G. Pagès and J. Yu. Pointwise convergence of the lloyd algorithm in higher dimension. *Preprint LPMA*, 2013.
- [46] E. Pardoux. BSDEs, weak convergence and homogenization of semilinear PDEs. *Nonlinear Analysis, Differential Equations and Control, NATO Science Series*, 528:503–549, 1999.
- [47] E. Pardoux and S. G. Peng. Adapted solutions of backward stochastic differential equation. *Systems and Control Letters*, 14:55–61, 1990.
- [48] E. Pardoux and S. G. Peng. Backward stochastic differential equations and quasilinear parabolic partial differential equations. *Lecture Notes in CIS*, 176:200–217, 1992.
- [49] M. J. Ruijter and Oosterlee C.W. A Fourier cosine method for an efficient computation of solutions to BSDEs. *SIAM J. Sci. Computation*, 37(2):859–889, 2015.
- [50] A. Sellami. Comparative survey on nonlinear filtering methods: the quantization and the particle filtering approaches. *Journal of Statistical Computation and Simulation*, 78(2):93–113, 2008.
- [51] J. Zhang. A numerical scheme for bsdes. *Ann. Appl. Probab.*, 14(1):459–488, 2004.

## A Proof of Theorem 3.1

Temporarily set for convenience  $\bar{s} = t_k$  for  $s \in [t_k, t_{k+1})$ .

**Proof of Theorem 3.1.** STEP 1. Applying Ito’s formula we have

$$\begin{aligned}
e^{\alpha T} \tilde{Y}_T^2 &= e^{\alpha t} \tilde{Y}_t^2 + \int_t^T \alpha e^{\alpha s} \tilde{Y}_s^2 ds + 2 \int_t^T e^{\alpha s} \tilde{Y}_s d\tilde{Y}_s + \int_t^T e^{\alpha s} |\tilde{Z}_s|^2 ds \\
&= e^{\alpha t} \tilde{Y}_t^2 + \int_t^T e^{\alpha s} [\alpha \tilde{Y}_s^2 + |\tilde{Z}_s|^2 + 2\tilde{Y}_s f(\underline{s}, \bar{X}_{\underline{s}}, \mathbb{E}_{\underline{s}}(\tilde{Y}_{\bar{s}}), \zeta_{\underline{s}})] ds + 2 \int_t^T e^{\alpha s} \tilde{Z}_s dW_s.
\end{aligned}$$

Then, using assumption (20) we have

$$e^{\alpha T} \tilde{Y}_T^2 \leq e^{\alpha t} \tilde{Y}_t^2 + \int_t^T e^{\alpha s} [\alpha \tilde{Y}_s^2 + |\tilde{Z}_s|^2 + 2\tilde{Y}_s C(f)(1 + |\bar{X}_s| + |\mathbb{E}_s(\tilde{Y}_s)| + |\zeta_s|)] ds + 2 \int_t^T e^{\alpha s} \tilde{Z}_s dW_s.$$

Owing to Young's inequality ( $ab \leq \frac{a^2}{2\theta} + \frac{\theta b^2}{2}$ , for every  $\theta > 0$  and  $a, b \geq 0$ ) we get

$$\begin{aligned} e^{\alpha t} \tilde{Y}_t^2 &\leq e^{\alpha T} \tilde{Y}_T^2 - \alpha \int_t^T e^{\alpha s} |\tilde{Y}_s|^2 ds - \int_t^T e^{\alpha s} |\tilde{Z}_s|^2 ds + \theta C(f) \int_t^T e^{\alpha s} \tilde{Y}_s^2 ds \\ &\quad + \frac{C(f)}{\theta} \int_t^T e^{\alpha s} (1 + |\bar{X}_s|^2 + |\mathbb{E}_s(\tilde{Y}_s)|^2 + |\zeta_s|^2) ds - 2 \int_t^T e^{\alpha s} \tilde{Z}_s dW_s \\ &\leq e^{\alpha T} \tilde{Y}_T^2 + (\theta C(f) - \alpha) \int_t^T e^{\alpha s} \tilde{Y}_s^2 ds - \int_t^T e^{\alpha s} |\tilde{Z}_s|^2 ds \\ &\quad + \frac{C(f)}{\theta} \int_t^T e^{\alpha s} (1 + |\bar{X}_s|^2 + |\mathbb{E}_s(\tilde{Y}_s)|^2 + |\zeta_s|^2) ds - 2 \int_t^T e^{\alpha s} \tilde{Z}_s dW_s. \end{aligned}$$

After choosing  $\alpha$  and  $\theta$  such that  $\theta C(f) - \alpha < 0$ , we take the expectation in both sides of the previous inequality and use the fact that  $\mathbb{E}|\mathbb{E}_s(\tilde{Y}_s)|^2 \leq \mathbb{E}|\tilde{Y}_s|^2$  (owing to conditional Jensen inequality) to get

$$e^{\alpha t} \mathbb{E}(\tilde{Y}_t^2) + \int_t^T e^{\alpha s} \mathbb{E}|\tilde{Z}_s|^2 ds \leq e^{\alpha T} \mathbb{E}(\tilde{Y}_T^2) + \frac{C(f)}{\theta} \int_t^T e^{\alpha s} (1 + \mathbb{E}|\bar{X}_s|^2 + \mathbb{E}(\tilde{Y}_s^2) + \mathbb{E}|\zeta_s|^2) ds.$$

Owing to the fact that  $\mathbb{E}(\sup_{t \in [0, T]} |\bar{X}_t|^2) \leq C_X(1 + \mathbb{E}|X_0|^2)$  and setting  $t = t_k$ , we have

$$\begin{aligned} e^{\alpha t_k} \mathbb{E}(\tilde{Y}_{t_k}^2) + \int_{t_k}^T e^{\alpha s} \mathbb{E}|\tilde{Z}_s|^2 ds &\leq e^{\alpha T} \mathbb{E}(\tilde{Y}_T^2) + \frac{C(f)}{\theta} \left( \frac{e^{\alpha T} - e^{\alpha t_k}}{\alpha} + C_X(1 + \mathbb{E}|X_0|^2) \right) \\ &\quad + \frac{C(f)}{\theta} \sum_{\ell=k}^{n-1} e^{\alpha t_\ell} \mathbb{E}|\tilde{Y}_{t_{\ell+1}}|^2 + \frac{C(f)}{\theta} \int_{t_k}^T e^{\alpha s} \mathbb{E}|\zeta_s|^2 ds. \end{aligned}$$

On the other hand, we have

$$\tilde{\zeta}_{t_\ell} = \frac{1}{\Delta_n} \mathbb{E}_\ell \int_{t_\ell}^{t_{\ell+1}} \tilde{Z}_s ds, \quad \text{so that by Jensen's inequality, } |\tilde{\zeta}_{t_\ell}|^2 \leq \frac{1}{\Delta_n} \mathbb{E}_\ell \int_{t_\ell}^{t_{\ell+1}} |\tilde{Z}_s|^2 ds. \quad (64)$$

It follows that

$$\begin{aligned} \int_{t_k}^T e^{\alpha s} \mathbb{E}|\zeta_s|^2 ds &\leq \frac{1}{\Delta_n^2} \mathbb{E} \sum_{\ell=k}^{n-1} \int_{t_\ell}^{t_{\ell+1}} e^{\alpha s} \left( \int_{t_\ell}^{t_{\ell+1}} |\tilde{Z}_u|^2 du \right) ds \\ &\leq \frac{e^{\alpha \Delta_n} - 1}{\alpha \Delta_n} \mathbb{E} \sum_{\ell=k}^{n-1} \int_{t_\ell}^{t_{\ell+1}} e^{\alpha s} |\tilde{Z}_u|^2 du = \frac{e^{\alpha \Delta_n} - 1}{\alpha \Delta_n} \int_{t_k}^T e^{\alpha s} \mathbb{E}|\tilde{Z}_u|^2 du. \end{aligned}$$

Since  $e^{\alpha \Delta_n} - 1 \leq \alpha \Delta_n e^{\alpha \Delta_n}$ , we have

$$\begin{aligned} e^{\alpha t_k} \mathbb{E}|\tilde{Y}_{t_k}|^2 + \int_{t_k}^T e^{\alpha s} \mathbb{E}|\tilde{Z}_s|^2 ds &\leq e^{\alpha T} \mathbb{E}|\tilde{Y}_T|^2 + \frac{C(f)}{\theta \alpha} e^{\alpha T} + \frac{C(f)}{\theta} C_X(1 + \mathbb{E}|X_0|^2) \\ &\quad + \frac{\Delta C(f)}{\theta} \sum_{\ell=k}^{n-1} e^{\alpha t_{\ell+1}} \mathbb{E}|\tilde{Y}_{t_{\ell+1}}|^2 + \frac{C(f)}{\theta} e^{\alpha \Delta_n} \int_{t_k}^T e^{\alpha s} \mathbb{E}|\tilde{Z}_u|^2 ds. \end{aligned}$$

Now, let us choose  $\theta$  so that  $\frac{C(f)}{\theta} e^{\alpha \Delta_n} < 1$ . Owing to the fact that  $\theta C(f) < \alpha$ , this implies that  $C(f) e^{\alpha \Delta_n} < \theta < \frac{\alpha}{C(f)}$ . This constraint holds true if  $e^{\alpha \Delta_n} < \frac{\alpha}{C(f)^2}$ . Taking  $\alpha > C(f)^2(T \vee 1)$  and



owing to the fact that  $e^{\alpha\Delta_n} \rightarrow 1$  as  $n$  goes to infinity we may consequently choose  $\theta \in (C(f)(e^{\alpha\Delta_n} \vee T), \frac{\alpha}{C(f)})$ , for every  $n \geq n_0 \in \mathbb{N}$ . Setting

$$C^{(1,1)} = e^{\alpha T} \mathbb{E}|\tilde{Y}_T|^2 + \frac{C(f)}{\theta\alpha} e^{\alpha T} + \frac{C_f}{\theta} C_X(1 + \mathbb{E}|X_0|^2), \quad C^{(1,2)} = \frac{C(f)}{\theta} \text{ and } C^{(1,3)} = \frac{C(f)}{\theta} e^{\alpha\Delta_n},$$

it follows that, for every  $n \geq n_0$ ,

$$e^{\alpha t_k} \mathbb{E}|\tilde{Y}_{t_k}|^2 + (1 - C^{(1,3)}) \int_{t_k}^T e^{\alpha s} \mathbb{E}|\tilde{Z}_s|^2 ds \leq C^{(1,1)} + \Delta C^{(1,2)} \sum_{\ell=k+1}^n e^{\alpha t_\ell} \mathbb{E}|\tilde{Y}_{t_\ell}|^2. \quad (65)$$

In particular we have  $\mathbb{E}|\tilde{Y}_T|^2 = \mathbb{E}\xi^2 \leq C^{(1,1)}$  and

$$e^{\alpha t_k} \mathbb{E}|\tilde{Y}_{t_k}|^2 \leq C^{(1,1)} + \Delta_n C^{(1,2)} \sum_{\ell=k+1}^n e^{\alpha t_\ell} \mathbb{E}|\tilde{Y}_{t_\ell}|^2, \quad \forall k \in \{0, \dots, n-1\}. \quad (66)$$

Since  $\theta > TC(f)$  then  $TC^{(1,2)} < 1$  and we may show by induction that if  $A \geq C^{(1,1)}/(1 - TC^{(1,2)})$  then

$$\sup_{k=0, \dots, n} e^{\alpha t_k} \mathbb{E}|\tilde{Y}_{t_k}|^2 \leq A \quad \text{so that} \quad \sup_{k=0, \dots, n} \mathbb{E}|\tilde{Y}_{t_k}|^2 \leq A.$$

Now, setting  $k = 0$  in (65) we get

$$\sup_{n \geq 0} \int_0^T e^{\alpha s} \mathbb{E}|\tilde{Z}_s|^2 ds \leq \frac{C^{(1,1)}}{1 - C^{(1,3)}} + \frac{n-k}{n} \frac{C^{(1,2)}}{1 - C^{(1,3)}} A \times T \leq \frac{C^{(1,1)}}{1 - C^{(1,3)}} + \frac{C^{(1,2)}}{1 - C^{(1,3)}} A \times T.$$

Furthermore, since  $|\tilde{\zeta}_{t_k}|^2 \leq \frac{1}{\Delta_n} \mathbb{E}_k \int_{t_k}^{t_{k+1}} |\tilde{Z}_s|^2 ds$  (see (64)), we deduce that

$$\Delta_n \sum_{k=0}^{n-1} \mathbb{E}|\tilde{\zeta}_{t_k}|^2 \leq \int_{t_k}^{t_{k+1}} \mathbb{E}|\tilde{Z}_s|^2 ds \leq C^{(1,4)}$$

where  $C^{(1,4)}$  is a positive real constant not depending on  $n$ .

STEP 2. We show that  $\tilde{Y}$  satisfies

$$\forall t \in [0, T], \quad \mathbb{E}|\tilde{Y}_t - \tilde{Y}_{\underline{t}}|^2 \leq C_{b,\sigma,f,T} |t - \underline{t}|, \quad C_{b,\sigma,f,T} > 0.$$

In fact, we have for every  $t \in [t_k, t_{k+1}]$ ,

$$\tilde{Y}_t = \tilde{Y}_{t_k} - (t - t_k) f(t_k, \bar{X}_{t_k}, \mathbb{E}(\tilde{Y}_{t_{k+1}} | \mathcal{F}_{t_k}), \tilde{\zeta}_{t_k}) + \int_{t_k}^t \tilde{Z}_s dW_s.$$

Then, using the assumptions (20) yield

$$\mathbb{E}|\tilde{Y}_t - \tilde{Y}_{t_k}|^2 \leq C(f)(t - t_k)(1 + \mathbb{E}|\bar{X}_{t_k}|^2 + \mathbb{E}|\tilde{Y}_{t_{k+1}}|^2 + \mathbb{E}|\tilde{\zeta}_{t_k}|^2) + \int_{t_k}^t \mathbb{E}|\tilde{Z}_s|^2 ds.$$

Now, thanks to the previous step we know that

$$\sup_{s \in [t_k, t]} \mathbb{E}|\tilde{Z}_s|^2 < +\infty, \quad \sup_{k \in \{0, \dots, n\}} \mathbb{E}|\tilde{Y}_{t_k}|^2 < +\infty \text{ and } \sup_{n \geq 1} \sup_{k \in \{0, \dots, n\}} \mathbb{E}|\tilde{\zeta}_{t_k}|^2 < +\infty.$$

We also know that  $\sup_{n \geq 1} \sup_{k \in \{0, \dots, n\}} \mathbb{E}|\bar{X}_{t_k}|^2 < +\infty$ . As a consequence, there exists a positive real constant  $C_{b,\sigma,f,T}$  such that for every  $t \in [t_k, t_{k+1}]$ ,

$$\forall t \in [t_k, t_{k+1}], \quad \mathbb{E}|\tilde{Y}_t - \tilde{Y}_{t_k}|^2 \leq C_{b,\sigma,f,T} |t - t_k|, \quad k = 0, \dots, n-1.$$

STEP 3. Let  $t \in [0, T]$ . It follows from Ito's formula that

$$\begin{aligned}
e^{\alpha t} |Y_t - \tilde{Y}_t|^2 &= 2 \int_t^T e^{\alpha s} (Y_s - \tilde{Y}_s) (f(s, X_s, Y_s, Z_s) - f(\underline{s}, \bar{X}_{\underline{s}}, \mathbb{E}_{\underline{s}}(\tilde{Y}_{\bar{s}}), \tilde{\zeta}_{\underline{s}})) ds \\
&\quad - \alpha \int_t^T e^{\alpha s} |Y_s - \tilde{Y}_s|^2 ds - \int_t^T e^{\alpha s} |Z_s - \tilde{Z}_s|^2 ds + 2 \int_t^T e^{\alpha s} (Z_s - \tilde{Z}_s) dW_s \\
&\leq 2 \int_t^T e^{\alpha s} [f]_{\text{Lip}} |Y_s - \tilde{Y}_s| (\Delta_n^2 + |X_s - \bar{X}_{\underline{s}}|^2 + |Y_s - \mathbb{E}_{\underline{s}}(\tilde{Y}_{\bar{s}})|^2 + |Z_s - \tilde{\zeta}_{\underline{s}}|^2)^{\frac{1}{2}} ds \\
&\quad - \alpha \int_t^T e^{\alpha s} |Y_s - \tilde{Y}_s|^2 ds - \int_t^T e^{\alpha s} |Z_s - \tilde{Z}_s|^2 ds + 2 \int_t^T e^{\alpha s} (Z_s - \tilde{Z}_s) dW_s.
\end{aligned}$$

Using the Young inequality:  $ab \leq \frac{\theta}{2} a^2 + \frac{1}{2\theta} b^2$ ,  $\forall \theta > 0$ , yields

$$\begin{aligned}
e^{\alpha t} |Y_t - \tilde{Y}_t|^2 &\leq [f]_{\text{Lip}} \int_t^T e^{\alpha s} \left( \theta |Y_s - \tilde{Y}_s|^2 + \frac{1}{\theta} (\Delta_n^2 + |X_s - \bar{X}_{\underline{s}}|^2 + |Y_s - \mathbb{E}_{\underline{s}}(\tilde{Y}_{\bar{s}})|^2 + |Z_s - \tilde{\zeta}_{\underline{s}}|^2) \right) ds \\
&\quad - \alpha \int_t^T e^{\alpha s} |Y_s - \tilde{Y}_s|^2 ds - \int_t^T e^{\alpha s} |Z_s - \tilde{Z}_s|^2 ds + 2 \int_t^T e^{\alpha s} (Z_s - \tilde{Z}_s) dW_s. \quad (67)
\end{aligned}$$

The stochastic integral on the right hand side of the previous inequality is a martingale since both  $Z$  and  $\tilde{Z}$  lie in  $L^2([0, T] \times \Omega, dt \otimes d\mathbb{P})$ . On the other hand, owing to the error bound for the Euler scheme and the fact that  $X$  is an Itô process, we get

$$\mathbb{E} |X_s - \bar{X}_{\underline{s}}|^2 \leq 2(C^{(3,1)} \mathbb{E} |X_s - X_{\underline{s}}|^2 + C^{(3,2)} \mathbb{E} |X_{\underline{s}} - \bar{X}_{\underline{s}}|^2) \leq C^{(3,3)} \Delta_n,$$

for some positive real constants  $C^{(3,1)}$ ,  $C^{(3,2)}$  and  $C^{(3,3)}$ . Then, taking the expectation in (67) and using the fact that

$$\mathbb{E} |Y_s - \mathbb{E}_{\underline{s}}(\tilde{Y}_{\bar{s}})|^2 \leq 2 \mathbb{E} |Y_s - \tilde{Y}_s|^2 + 2 \mathbb{E} |\tilde{Y}_s - \mathbb{E}_{\underline{s}}(\tilde{Y}_{\bar{s}})|^2$$

yield

$$\begin{aligned}
\mathbb{E} \left( e^{\alpha t} |Y_t - \tilde{Y}_t|^2 + \int_t^T e^{\alpha s} |Z_s - \tilde{Z}_s|^2 ds \right) &\leq \left( -\alpha + [f]_{\text{Lip}} \left( \theta + \frac{2}{\theta} \right) \right) \int_t^T e^{\alpha s} \mathbb{E} |Y_s - \tilde{Y}_s|^2 ds \\
&\quad + \frac{[f]_{\text{Lip}}}{\theta} \left( \frac{e^{\alpha T} - e^{\alpha t}}{\alpha} (\Delta_n^2 + C^{(3,3)} \Delta_n) + 2 \int_t^T e^{\alpha s} \mathbb{E} |\tilde{Y}_s - \mathbb{E}_{\underline{s}}(\tilde{Y}_{\bar{s}})|^2 ds + \int_t^T e^{\alpha s} \mathbb{E} |Z_s - \tilde{\zeta}_{\underline{s}}|^2 ds \right). \quad (68)
\end{aligned}$$

We notice that for every  $k \in \{0, \dots, n-1\}$  and for every  $s \in [t_k, t_{k+1})$ ,

$$\tilde{\zeta}_{\underline{s}} = \frac{1}{\Delta_n} \mathbb{E}_k \int_{t_k}^{t_{k+1}} \tilde{Z}_s ds \in \arg \min_{a \in \mathcal{F}_{t_k}} \mathbb{E}_k \int_{t_k}^{t_{k+1}} |Z_s - a|^2 ds$$

and

$$\zeta_{\underline{s}} := \frac{1}{\Delta_n} \mathbb{E}_k \int_{t_k}^{t_{k+1}} Z_s ds \in \arg \min_{a \in \mathcal{F}_{t_k}} \mathbb{E}_k \int_{t_k}^{t_{k+1}} |Z_s - a|^2 ds,$$

where  $a \in \mathcal{F}_{t_k}$  means that  $a$  is an  $\mathbb{R}^q$ -valued  $\mathcal{F}_{t_k}$ -measurable random vector. Then, using the inequality  $\mathbb{E}_k |Z_s - \tilde{\zeta}_{\underline{s}}|^2 \leq 2 \mathbb{E}_k |Z_s - \zeta_{\underline{s}}|^2 + 2 \mathbb{E}_k |\zeta_{\underline{s}} - \tilde{\zeta}_{\underline{s}}|^2$ , we get

$$\begin{aligned}
\int_{t_k}^{t_{k+1}} e^{\alpha s} \mathbb{E}_k |Z_s - \tilde{\zeta}_{\underline{s}}|^2 ds &\leq 2 \int_{t_k}^{t_{k+1}} e^{\alpha s} \mathbb{E}_k |Z_s - \zeta_{\underline{s}}|^2 ds + 2 \int_{t_k}^{t_{k+1}} e^{\alpha s} \mathbb{E}_k |\zeta_{\underline{s}} - \tilde{\zeta}_{\underline{s}}|^2 ds \\
&\leq 2 \int_{t_k}^{t_{k+1}} e^{\alpha s} \mathbb{E}_k |Z_s - Z_{\underline{s}}|^2 ds + \frac{2}{\Delta_n^2} \int_{t_k}^{t_{k+1}} e^{\alpha s} \mathbb{E}_k \left| \mathbb{E}_k \int_{t_k}^{t_{k+1}} (Z_u - \tilde{Z}_u) du \right|^2 ds. \quad (69)
\end{aligned}$$

Now, owing to the Cauchy-Schwarz inequality, we have

$$\begin{aligned}
\int_{t_k}^{t_{k+1}} e^{\alpha s} \mathbb{E}_k \left| \mathbb{E}_k \int_{t_k}^{t_{k+1}} (Z_u - \tilde{Z}_u) du \right|^2 ds &\leq \Delta_n \int_{t_k}^{t_{k+1}} e^{\alpha s} ds \mathbb{E}_k \int_{t_k}^{t_{k+1}} |Z_u - \tilde{Z}_u|^2 du \\
&= \Delta_n \frac{e^{\alpha \Delta_n} - 1}{\alpha} e^{\alpha t_k} \mathbb{E}_k \int_{t_k}^{t_{k+1}} |Z_u - \tilde{Z}_u|^2 du \\
&\leq \Delta_n \frac{e^{\alpha \Delta_n} - 1}{\alpha} \mathbb{E}_k \int_{t_k}^{t_{k+1}} e^{\alpha s} |Z_u - \tilde{Z}_u|^2 du.
\end{aligned}$$

Consequently, taking the expectation in (69) leads to

$$\int_{t_k}^{t_{k+1}} e^{\alpha s} \mathbb{E} |Z_s - \tilde{Z}_s|^2 ds \leq 2 \int_{t_k}^{t_{k+1}} e^{\alpha s} \mathbb{E} |Z_s - Z_{\underline{s}}|^2 ds + 2 \frac{e^{\alpha \Delta_n} - 1}{\alpha \Delta_n} \int_{t_k}^{t_{k+1}} e^{\alpha s} \mathbb{E} |Z_u - \tilde{Z}_u|^2 du$$

Coming back to Inequality (68) and setting  $\alpha = \alpha(\theta) = [f]_{\text{Lip}}(\theta + \frac{2}{\theta})$  yields

$$\begin{aligned}
\mathbb{E} \left( e^{\alpha t} |Y_t - \tilde{Y}_t|^2 + \int_t^T e^{\alpha s} |Z_s - \tilde{Z}_s|^2 ds \right) &\leq \frac{[f]_{\text{Lip}}}{\theta} \left( \Delta_n C_{b,\sigma,T} + 2 \int_t^T e^{\alpha s} \mathbb{E} |\tilde{Y}_s - \mathbb{E}_{\underline{s}}(\tilde{Y}_s)|^2 ds \right. \\
&\quad \left. + 2 \frac{e^{\alpha \Delta_n} - 1}{\alpha \Delta_n} \int_t^T e^{\alpha s} \mathbb{E} |Z_u - \tilde{Z}_u|^2 du \right. \\
&\quad \left. + 2 \int_t^T e^{\alpha s} \mathbb{E} |Z_s - Z_{\underline{s}}|^2 ds \right).
\end{aligned}$$

Owing to Step 2, we have for every  $t \in [0, T]$ ,  $\mathbb{E} |\tilde{Y}_t - \tilde{Y}_{\underline{t}}|^2 \leq C_{b,\sigma,f,T}(t - \underline{t})$  with  $C_{b,\sigma,f,T} > 0$  so that, using the conditional Jensen inequality we get

$$\begin{aligned}
\mathbb{E} |\tilde{Y}_s - \mathbb{E}_{\underline{s}}(\tilde{Y}_s)|^2 &\leq 2\mathbb{E} |\tilde{Y}_s - \tilde{Y}_{\underline{s}}|^2 + 2\mathbb{E} |\mathbb{E}_{\underline{s}}(\tilde{Y}_s - \tilde{Y}_{\underline{s}})|^2 \\
&\leq 2\mathbb{E} |\tilde{Y}_s - \tilde{Y}_{\underline{s}}|^2 + 2\mathbb{E} |\tilde{Y}_{\underline{s}} - \tilde{Y}_{\underline{s}}|^2 \\
&\leq 4C_{b,\sigma,f,T} \Delta_n.
\end{aligned}$$

As a consequence, using that  $\frac{e^{\alpha \Delta_n} - 1}{\alpha \Delta_n} \leq e^{\alpha \Delta_n}$ , we have

$$\begin{aligned}
\mathbb{E} \left( e^{\alpha t} |Y_t - \tilde{Y}_t|^2 + \int_t^T e^{\alpha s} |Z_s - \tilde{Z}_s|^2 ds \right) &\leq \frac{[f]_{\text{Lip}}}{\theta} \left( \Delta_n C'_{b,\sigma,f,T} + 2e^{\alpha \Delta_n} \int_t^T e^{\alpha s} \mathbb{E} |Z_u - \tilde{Z}_u|^2 du \right. \\
&\quad \left. + 2 \int_t^T e^{\alpha s} \mathbb{E} |Z_s - Z_{\underline{s}}|^2 ds \right).
\end{aligned}$$

Let  $\theta \in [4[f]_{\text{Lip}}, 6[f]_{\text{Lip}}]$ . Then

$$2 \frac{[f]_{\text{Lip}}}{\theta} e^{\alpha \Delta_n} \leq \frac{1}{2} \exp \left( [f]_{\text{Lip}} ([f]_{\text{Lip}}^2 + \frac{1}{2}) \Delta_n \right)$$

so that, for large enough  $n$ , say  $n \geq n_0$ ,  $2 \frac{[f]_{\text{Lip}}}{\theta} e^{\alpha \Delta_n} \leq \frac{3}{4}$  since  $\Delta_n \rightarrow 0$ . It follows that

$$\mathbb{E} \left( e^{\alpha t} |Y_t - \tilde{Y}_t|^2 + \frac{1}{4} \int_t^T e^{\alpha s} |Z_s - \tilde{Z}_s|^2 ds \right) \leq C^{(3,4)} \left( \Delta_n + \int_t^T e^{\alpha s} \mathbb{E} |Z_s - Z_{\underline{s}}|^2 ds + \int_t^T e^{\alpha s} \mathbb{E} |Z_s - \tilde{Z}_s|^2 ds \right).$$

In particular, for every  $k = 0, \dots, n$ , as  $\underline{t}_k = t_k$ ,

$$\begin{aligned}
\mathbb{E} |Y_{t_k} - \tilde{Y}_{t_k}|^2 &\leq C^{(3,4)} \left( \Delta_n + \int_{t_k}^T e^{\alpha s} \mathbb{E} |Z_s - Z_{\underline{s}}|^2 ds \right) \\
&\leq C^{(3,5)} e^{\alpha T} \left( \Delta_n + \int_0^T \mathbb{E} |Z_s - Z_{\underline{s}}|^2 ds \right).
\end{aligned}$$

Now, setting  $k = 0$  yields likewise

$$\mathbb{E} \left( \int_0^T e^{\alpha s} |Z_s - \tilde{Z}_s|^2 ds \right) \leq C^{(3,6)} \left( \Delta_n + e^{\alpha T} \int_0^T \mathbb{E} |Z_s - Z_{\underline{s}}|^2 ds \right),$$

which completes the proof since one can always satisfy this inequality for  $n = 1, \dots, n_0$ , by increasing the constant  $C^{(3,6)}$ .

STEP 4. Let us consider the following PDE:

$$\frac{\partial u}{\partial t}(t, x) + \mathcal{L}u(t, x) + f(t, x, u, \nabla_x u(t, x)\sigma(t, x)) = 0, \quad u(T, x) = h(T, x), \quad x \in \mathbb{R}^d,$$

where  $\mathcal{L}$  is the second order differential operator defined by

$$\mathcal{L} = \frac{1}{2} \sum_{i,j} [\sigma \sigma^*(t, x)]_{ij} \partial_{x_i x_j}^2 + \sum_i b_i(t, x) \partial_{x_i},$$

(where  $\sigma^*$  stands for the transpose of  $\sigma$ ). We know (see e.g. [17]) that under our assumptions, the solution  $u$  of this PDE satisfies

$$Y_t = u(t, X_t) \quad \text{and} \quad Z_t = \nabla_x u(t, X_t) \sigma(t, X_t).$$

Furthermore (see again [17]), under the hypothesis made on the coefficients of the Forward-Backward SDEs, there is a positive real constant  $C^{(4,1)}$  which depends only on  $T$ , such that

$$\forall s, t, \in [0, T], \forall x \in \mathbb{R}^d, \quad |\nabla_x u(t, x) - \nabla_x u(s, x)| \leq C^{(4,1)} |t - s|^{1/2}. \quad (70)$$

We also know that there is a constant  $C^{(4,2)} > 0$  such that for every  $s, t \in [0, T]$ ,

$$\mathbb{E} |X_t - X_s|^2 \leq C^{(4,2)} |t - s|. \quad (71)$$

Then, using (70), the assumptions on  $\sigma$  and the fact that  $u$  belongs to the set  $C_b^{1,2}$  of continuously differentiable functions  $\phi(t, x)$  which partial derivatives  $\partial_t \phi$ ,  $\partial_x \phi$  and  $\partial_{x^2}^2 \phi$  exist and are uniformly bounded (in particular  $\nabla_x u$  is Lipschitz continuous in  $x$ ), we have for every  $t \in [0, T]$ ,

$$\begin{aligned} \mathbb{E} |Z - Z_{\underline{t}}|^2 &\leq 2\mathbb{E} |\nabla_x u \sigma(t, X_t) - \nabla_x u(\underline{t}, X_{\underline{t}}) \sigma(t, X_t)|^2 + 2\mathbb{E} |\nabla_x u(\underline{t}, X_{\underline{t}}) \sigma(t, X_t) - \nabla_x u \sigma(\underline{t}, X_{\underline{t}})|^2 \\ &\leq C(\mathbb{E} |\nabla_x u(t, X_t) - \nabla_x u(\underline{t}, X_{\underline{t}})|^2 + \mathbb{E} |\sigma(t, X_t) - \sigma(\underline{t}, X_{\underline{t}})|^2) \\ &\leq C((t - \underline{t}) + \mathbb{E} |X_t - X_{\underline{t}}|^2) \\ &\leq C(t - \underline{t}), \end{aligned}$$

for some real positive constant  $C > 0$  which may change from line to line in the previous inequalities. It follows that

$$\int_0^T \mathbb{E} |Z - Z_{\underline{t}}|^2 dt \leq C \int_0^T (t - \underline{t}) dt = C \frac{(t_n - t_{n-1})^2}{2} \leq C \Delta_n^2.$$

□

## B Distortion mismatch: Proof of Theorem 4.3

STEP 1 (*Control of the distance to the quantizers*): Let  $(\Gamma_N)_{N \geq 1}$  be a sequence of  $L^r$ -optimal quantizers. it is clear that, for every  $\xi \in \mathbb{R}^d$ ,

$$d(\xi, \Gamma_N) \leq |\xi| + d(0, \Gamma_N).$$

The sequence  $(d(0, \Gamma_N))_{N \geq 1}$  is bounded since  $d(\Gamma_N, \text{supp}(P)^c) \rightarrow 0$  as  $N \rightarrow +\infty$  and  $d(0, \text{supp}(P)^c) < +\infty$ . Then there exists a real constant  $A_X \geq 0$  such that for every  $\xi \in \mathbb{R}^d$ ,

$$d(\xi, \Gamma_N) \leq |\xi| + A_X.$$

STEP 2 (*Micro-macro inequality*): The optimality of the grids  $\Gamma_N$ ,  $N \geq 1$ , allow to apply to the micro-macro inequality (see Equation (3.2) in the proof of Theorem 2 in [24]), namely : for every real constant  $c \in (0, \frac{1}{2})$  and every  $y \in \mathbb{R}$ ,

$$e_r(\Gamma_N, P)^r - e_r(\Gamma_{N+1}, P)^r \geq ((1-c)^r - c^r) P(B(y; cd(y, \Gamma_N))) d(y, \Gamma_N)^r. \quad (72)$$

Let  $\nu$  be an auxiliary Borel probability measure on  $\mathbb{R}^d$  to be specifies further on. Integrating the above inequality with respect to  $\nu(dy)$  yields, we derive, owing to Fubini's Theorem,

$$\begin{aligned} e_r(\Gamma_N, P)^r - e_r(\Gamma_{N+1}, P)^r &\geq ((1-c)^r - c^r) \int \int (B(y; cd(y, \Gamma_N))) d(y, \Gamma_N)^r p(d\xi) \nu(dy) \\ &= ((1-c)^r - c^r) \int \int \mathbf{1}_{\{|y-\xi| \leq cd(y, \Gamma_N)\}} d(y, \Gamma_N)^r \nu(dy) P(d\xi) \\ &\geq ((1-c)^r - c^r) \int \int \mathbf{1}_{\{|y-\xi| \leq cd(y, \Gamma_N), d(y, \Gamma_N) \geq \frac{1}{c+1} d(\xi, \Gamma_N)\}} d(y, \Gamma_N)^r \nu(dy) P(d\xi). \end{aligned}$$

Now using that  $\xi \mapsto d(\xi, \Gamma_n)$  is Lipschitz continuous with coefficient 1, one derives that

$$\{(\xi, y) : |y - \xi| \leq \frac{c}{c+1} d(\xi, \Gamma_N)\} \subset \left\{ (\xi, y) : |y - \xi| \leq cd(y, \Gamma_N), d(y, \Gamma_N) \geq \frac{1}{c+1} d(\xi, \Gamma_N) \right\}$$

and, still by Fubini's Theorem,

$$e_r(\Gamma_N, P)^r - e_r(\Gamma_{N+1}, P)^r \geq \frac{(1-c)^r - c^r}{(1+c)^r} \int \nu(B(\xi; cd(y, \Gamma_N))) d(\xi, \Gamma_N)^r P(d\xi). \quad (73)$$

Let  $\varepsilon \in (0, 1/2)$ . We set  $\nu = f_{\varepsilon, \delta} \cdot \lambda_d$  where  $f_{\varepsilon, \delta}$  is a probability density given by

$$f_{\varepsilon, \delta}(\xi) = \frac{\kappa_{\varepsilon, \delta}}{(|x| + 1 + \varepsilon)^{d+\delta}} \quad \text{with } \delta > 0.$$

( $|\cdot|$  is any norm). The density  $f_{\varepsilon, \delta}$  shares the following property on balls: let  $\xi \in \mathbb{R}^d$  and  $t \in \mathbb{R}_+$ . If  $t \leq \varepsilon(|\xi| + 1)$ , then

$$\nu(B(\xi, t)) \geq g_{\varepsilon, \delta}(\xi) t^d \quad \text{with} \quad g_{\varepsilon, \delta}(\xi) = \frac{1}{(1+\varepsilon)^{d+\delta}} \frac{\kappa_{\varepsilon, \delta}}{(|\xi| + 1)^{d+\delta}} V_d$$

and  $V_d = \lambda_d(B(0; 1))$ . Now let  $c = c(\varepsilon) \in (0, 1)$  such that  $\frac{c}{c+1} = \varepsilon(A_X^{-1} \wedge 1)$ . As  $d(\xi, \Gamma_N) \leq |\xi| + A_X$ , this in turn implies that  $\frac{c}{c+1} d(\xi, \Gamma_N) \leq \varepsilon(|\xi| + 1)$ . As a consequence

$$e_r(\Gamma_N, P)^r - e_r(\Gamma_{N+1}, P)^r \geq \frac{(1-c)^r - c^r}{(c+1)^r} \int g_{\varepsilon, \delta}(\xi) d(\xi, \Gamma_N)^{r+d} P(d\xi).$$

Let  $s \in [r, r+d)$ . It follows from Equation (73) and the inverse Minkowski inequality applied with  $p = \frac{s}{r+d} \in (0, 1)$  and  $q = -\frac{s}{d+r-s} \in (-\infty, 0)$  that

$$\int g_{\varepsilon, \delta}(\xi) d(\xi, \Gamma_N)^{r+d} P(d\xi) \geq \left[ \int_{\mathbb{R}^d} d(\xi, \Gamma_N)^s P(d\xi) \right]^{\frac{r+d}{s}} \left[ \int g_{\varepsilon, \delta, a}(\xi)^{-\frac{s}{d+r-s}} P(d\xi) \right]^{-\frac{d+r-s}{s}}.$$

It follows from the assumption made on  $X$  (or  $P$ ) that, for small enough  $\delta > 0$ ,

$$\left[ \int_{\mathbb{R}^d} g_{\varepsilon, \delta}^{-\frac{s}{d+r-s}}(\xi) P(d\xi) \right]^{-\frac{d+r-s}{s}} = \frac{\kappa_{\varepsilon, \delta} V_d}{(1+\varepsilon)^{d+\delta}} \left[ \mathbb{E} \left[ (1+|X|)^{\frac{(d+\delta)s}{d+r-s}} \right] \right]^{-\frac{d+r-s}{s}} < +\infty$$

As a consequence

$$e_r(\Gamma_N, P)^r - e_r(\Gamma_{N+1}, P)^r \geq C_{X,r,s,\varepsilon,\delta} e_s(\Gamma_N, X)^{r+d} \quad (74)$$

where  $C_{X,r,s,\varepsilon,\delta} = \frac{(1-c)^r - c^r}{(1+c)^r (1+\varepsilon)^{d+\delta}} \kappa_{\varepsilon, \delta} \|1 + |X|\|_{\frac{(d+\delta)s}{d+r-s}}^{-(d+\delta)}$ .

STEP 3 (*Upper-bound for the quantization error increments*): One follows the lines of the proof of Theorem 2 in [24] to derive this upper-bound for the increments of the  $L^r$ -quantization error, namely there exists of a real constant  $\kappa_{X,r}$  such that

$$e_r(\Gamma_N, P)^r - e_r(\Gamma_{N+1}, P)^r \leq \kappa_{X,r} N^{-1-\frac{r}{d}}.$$

Combining the last two inequalities yields

$$\left[ e_s(\Gamma_N, X)^s \right]^{\frac{r+d}{s}} \leq \tilde{C}_{X,r,s,\varepsilon,\delta} N^{-\frac{r+d}{d}}$$

where  $\tilde{C}_{X,r,s,\varepsilon,\delta} = \frac{\kappa_{X,r}}{C_{X,r,s,\varepsilon,\delta}}$ . This completes the proof by considering the  $(d+r)^{th}$  root of the inequality.  $\square$