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

Gilles Pagès, Abass Sagna

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

GILLES PAGÈS * ABASS SAGNA † ‡

Abstract

We take advantage of recent (see [31, 49]) and new results on optimal quantization theory to improve the quadratic optimal quantization error bounds for backward stochastic differential equations (BSDE) and nonlinear filtering problems. For both problems, a first improvement relies on a Pythagoras like Theorem for quantized conditional expectation. While allowing for some locally Lipschitz functions conditional densities in nonlinear filtering, the analysis of the error brings into playing a new robustness result about optimal quantizers, the so-called distortion mismatch property: L^r-quadratic optimal quantizers of size N behave in L^s in term of mean error at the same rate $N^{-\frac{1}{d}}$, $0 < s < r + d$.

1 Introduction

In this work we propose improved error bounds for quantization based numerical schemes introduced in [4] and [47] 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 6). For both problems, we also improve the error bounds themselves by using a Pythagoras like theorem for the approximation of conditional expectations introduced in [49] (see also [46]). 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 [52]. 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 ξ:

$$Y_t = \xi + \int_t^T f(s, Y_s, Z_s)ds - \int_t^T Z_s dW_s, \qquad 0 \le t \le T, \tag{1}$$

where W is a q-dimensional brownian motion. We mean by a solution a pair $(Y_t, Z_t)_{t \le T}$ (valued in $\mathbb{R}^d \times \mathbb{R}^{d\times q}$) of square integrable progressively measurable (with respect to the augmented Brownian filtration $(\mathcal{F}_t)_{t\ge 0}$) and satisfying Equation (1). Extensions of these existence and uniqueness results

*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.

have been investigated in more general situations (less regular drivers f (see [23] for driver having a little regularity in time, called rough path driver, [1, 32] for locally Lipschitz driver, [38] for quadratic BSDEs and [39] for super-linear quadratic BSDEs), randomized horizon (see [51]), introduction of Poisson random measure component subject to constraints on the jump component (see [37, 36]), extension to second order BSDEs (see [44])).

Since the pioneering work [25] 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 [24] for a class of forward-backward stochastic differential equations, based a four step scheme developed later on in [42].

In [57], 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, 16, 17, 22]). Note in fact that in [17], 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 decoupled Markovian BSDE,

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

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

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

In this case, the approximating methods of the solution of the BSDE are written (for a given time discretization at instants $t_0 = 0, \ldots, t_n = T$) as a functional of the paths of $(X_{t_k})_{k=0,\ldots,n}$ 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 consider an *explicit* time discretization scheme where the conditioning is performed *inside* the driver f (see also [35]). It is recursively defined in a backward way as:

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

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

with

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

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

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

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 $n \geq 1$,

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

where $\tilde{Z} = \tilde{Z}^{(n)}$ comes from the martingale representation of $\sum_{k=1}^n \tilde{Y}_{t_k} - \mathbb{E}(\tilde{Y}_{t_k}|\mathcal{F}_{t_{k-1}})$.

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, 28]). 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, 19, 34]). New approaches have been proposed recently: a combination of Picard iterates and a decomposition in Wiener chaos (see [12]), a "forward" approach in connection with the semi-linear PDE associated to the BSDE (see [33]), an analytic approach in [29].

In this paper, we go back to the *optimal quantization tree* approach originally introduced in [5] (in fact for *Reflected* BSDEs) and developed in [4, 3, 6]. 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 $(t_k)_{k=0,\ldots,n}$) with random variables having a finite support. However, we consider a different quantization tree (or quantized scheme) defined recursively by mimicking (4)-(5) as follows:

$$\begin{aligned} \hat{Y}_{t_n} &= h(\hat{X}_{t_n}) && (7)\\ \hat{Y}_{t_k} &= \hat{\mathbb{E}}_k(\hat{Y}_{t_{k+1}}) + \Delta_n f\big(t_k, \hat{X}_{t_k}, \hat{\mathbb{E}}_k(\hat{Y}_{t_{k+1}}), \hat{\zeta}_{t_k}\big) && (8) \end{aligned}$$

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

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 Γ_k, $i.e.$, $\hat{X}_{t_k} = q_k(\bar{X}_{t_k})$, where $q_k : \mathbb{R}^d \to \Gamma_k$ are Borel functions, $k = 0, \ldots, n$. This is an *explicit inner* scheme in the sense that the conditioning is performed inside the driver f in contrast with what is usually done in the literature (where implicit or *outer* explicit schemes are in force). This scheme, though quite natural seems not to have been extensively analyzed (see however [35]), is well designed to establish our improved rates with quite satisfactory numerical performances. Our objective here is two-fold: first include the Z term in the driver and to dramatically improve the error bounds in [3, 6], especially its dependence in the size $n+1$ of the time discretization mesh.

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 $k = 0, \ldots, n$, 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(a) that, at every step k of the procedure,

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

for positive real constants $\widetilde{K}_i$ depending on t_i and T and on the regularity of the coefficients of b, σ and the driver f which remain bonded as $n \uparrow +\infty$. The presence of the squared quadratic norms on both sides of (9) improves the control of the time discretization effect, compared with [4, 5] in which error bounds of the form $\|\tilde{Y}_{t_k} - \hat{Y}_{t_k}\|_p \le \sum_{i=k}^{n} K_i \|\bar{X}_{t_i} - \hat{X}_{t_i}\|_p$ are established for $p \in [1,+\infty)$. In fact, we switch from a global error (at $t = 0$) of order $n \times \max_{0\le k\le n} \|X_{t_k} - \hat{X}_{t_k}\|_2$ to $\sqrt{n} \times \max_{0\le k\le n} \|X_{t_k} - \hat{X}_{t_k}\|_2$. As for the (root of the) integrated mean squared hedging error in Theorem 3.2(b), we switch from $n^{\frac{3}{2}} \times \max_{0\le k\le n} \|X_{t_k} - \hat{X}_{t_k}\|_2$ to $n \times \max_{0\le k\le n} \|X_{t_k} - \hat{X}_{t_k}\|_2$. This theoretical improvement confirms the results of numerical experiments first carried out in [6] though it was in a less favourable framework (with reflection). They were then extensively investigated in [14, 13] (for American options) to devise a a Romberg extrapolation combining two time discretization steps which dramatically improves the performances of such schemes.

We notice here that other quantization based discretization schemes have been devised, especially for *Forward-Backward* BSDE (see [21]) where the diffusion and the BSDE are fully coupled (including the Z in the driver) where the grids Γ_k are the trace of $\delta\mathbb{Z}^d$ ($\delta > 0$) on an (expanding compact as t_k grows). In contrast the Brownian increments are replaces by optimal quantization of the $\mathcal{N}(0; I_d)$-distribution. But the obtained resulting error bound for the scheme are not of the improved form (9). A multistep approach based on two reference ODEs from the computation of conditional expectation has been developed in a similar framework (coupled and uncoupled) in [58].

In the second part of the paper, we first propose (Section 4) a short background on optimal vector quantization, enriched by a new result, namely Theorem 4.3, which essentially solves the-called *distortion mismatch* problem. By distortion mismatch we mean the robustness of optimal quantization grids. An optimal (quadratic) grid Γ_N for the distribution of a random vector X is such that $\|X - \mathrm{Proj}_{\Gamma_N}(X)\|_2 = e_{N,2}(X) := \inf\big\{\|X - q(X)\|_2,\, q : \mathbb{R}^d \to \mathbb{R}^d,\, \mathrm{card}(\Gamma_N) \le N\big\}$ where Proj_{Γ_N} denotes a (Borel) nearest neighbor projection on Γ_N. It exists for every size (or level) $N \ge 1$ as soon as $X \in L^2$ and it follows from Zador's Theorem that $e_{N,2}(X) \sim c(X)N^{-\frac{1}{d}}$ as $N \to +\infty$ (see Section 4 for details). The distortion mismatch property established in Theorem 4.3 states that, for every $s \in (0, d+2)$, $\overline{\lim}_N N^{\frac{1}{d}} \|X - \mathrm{Proj}_{\Gamma_N}(X)\|_s < +\infty$. This result holds whenever $X \in L^s$ with a distribution satisfying mild additional property. This theorem extends first results established in [31] for various classes of absolutely continuous distributions. This robustness property is the key of the second kind of improvement proposed in this paper, this time for quantization based schemes for non-linear filtering investigated in the third part. In Section 5 we propose numerical illustrations using optimal quantization based schemes for various types of BSDEs which confirm that the improved rates established in the first part are the true ones.

In this third part of the paper (Section 6), we consider a (discrete time) nonlinear filtering problem and improve (in the quadratic setting) the results obtained in [47]. Firstly, we relax the Lipschitz assumption made on the conditional densities then we provide new improved error bounds for the quantization based scheme introduced in [47] to numerically solve a discrete filter by optimal quantization.

In fact, we consider a discrete time nonlinear filtering problem where the signal process $(X_k)_{k\ge 0}$ is an $\mathbb{R}^d$-valued discrete time Markov process and the observation process $(Y_k)_{k\ge 0}$ is an $\mathbb{R}^q$-valued random vector, both defined on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$. The distribution μ of X_0 is given, as well as the transition probabilities $P_k(x, dx') = \mathbb{P}(X_k \in dx' | X_{k-1} = x)$ of the process $(X_k)_{k\ge 0}$. We also suppose that the process $(X_k, Y_k)_{k\ge 0}$ is a Markov chain and that for every $k \ge 1$, 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 $Y := (Y_0, \ldots, Y_n) = (y_0, \ldots, y_n)$, for $n \ge 1$, we aim at computing the conditional distribution $\Pi_{y,n}$ of X_n given $Y = (y_0, \ldots, y_n)$. It is well-known that for any bounded and measurable function f, $\Pi_{y,n} f$ is given by the celebrated Kallianpur-Striebel formula (see e.g. [47])

$$\Pi_{y,n} f = \frac{\pi_{y,n} f}{\pi_{y,n} \mathbf{1}} \tag{10}$$

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 $H_{y,k}$, $k = 1, \ldots, n$, by

$$H_{y,k} f(x) = \mathbb{E}(f(X_k) g_k(x, y_{k-1}, X_k, y_k) | X_{k-1} = x) \tag{11}$$

for every bounded or non-negative Borel function $f : \mathbb{R}^d \to \mathbb{R}$ 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:

$$\pi_{y,k} f = \pi_{y,k-1} H_{y,k} f, \qquad k = 1, \ldots, n, \tag{12}$$

with $\pi_{y,0} = H_{y,0}$. A useful formulation, especially to establish error bound for the quantization based approximate filter is its *backward* counterpart defined by setting

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

where $u_{y,-1}$ is the final value of the backward recursion:

$$u_{y,n}(f)(x) = f(x), \quad u_{y,k-1}(f) = H_{y,k} u_{y,k}(f), \quad k = 0, \ldots, n. \tag{13}$$

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 formulas (12) or (13). However these kernels have no closed formula in general so that we have to approximate them. Optimal quantization based algorithms for non linear filtering has been introduced in [47] (see also [54, 15, 47, 56] for further developments and contributions). It turned out to be an efficient alternative approach to particle methods (we refer e.g. to [20] and the references therein which rely on Monte Carlo simulation of interacting particles) with owing to its tractability. For a survey and comparisons between optimal quantization and particle methods, we refer to [56]).

The quantization based approximate filter is designed as follows: denoting for every $k = 0, \ldots, n$ by $\hat{X}_k$ a quantization of X_k at level N_k by the grid $\Gamma_k = \{x_k^1, \ldots, x_k^{N_k}\}$, we will formally replace X_k in (12) or (13) by $\hat{X}_k$. As a consequence the (optimally) quantized approximation $\hat{\pi}_{y,n}$ of $\pi_{y,n}$ is defined simply by the quantized counterpart of the Kallianpur-Striebel formula: we introduce every bounded or non-negative Borel function $f : \mathbb{R}^d \to \mathbb{R}$ the family of quantized transition kernels $\hat{H}_{y,k}$, $k = 0, \ldots, n$, by $\hat{H}_{y,0} f(x) = \mathbb{E}(f(\hat{X}_0))$ and

$$\begin{aligned}
\hat{H}_{y,k} f(x_{k-1}^i) &= \mathbb{E}\big(f(\hat{X}_k) g_k(x_{k-1}^i, y_{k-1}, \hat{X}_k, y_k) | X_{k-1} = x_{k-1}^i\big), \quad k = 1, \ldots, n. && (14)\\
&= \frac{1}{2} \sum_{j=1}^{N_k} \hat{H}_{y,k}^{ij} f(x_k^j), \; i = 1, \ldots, N_{k-1} && (15)\\
\text{with} \qquad \hat{H}_{y,k}^{ij} &= g_k(x_{k-1}^i, y_{k-1}; x_k^j, y_k)\, \hat{p}_k^{ij} && (16)\\
\text{and} \qquad \hat{p}_k^{ij} &= \mathbb{P}(\widehat{X}_k = x_k^j \,|\, \widehat{X}_{k-1} = x_{k-1}^i) \quad i = 1, \ldots, N_{k-1}, j = 1, \ldots, N_k. && (17)
\end{aligned}$$

Then set

$$\hat{\pi}_{y,k} = \hat{\pi}_{y,k-1} \hat{H}_{y,k}, \quad k = 1, \ldots, n, \quad \text{ and } \quad \hat{\pi}_{y,0} = \hat{H}_{y,0} \tag{18}$$

or, equivalently,

$$\hat{\pi}_{y,k} = \sum_{i=1}^{N_k} \hat{\pi}^i_{y,k} \delta_{x^i_k} \quad \text{with} \quad \hat{\pi}^i_{y,k} = \sum_{j=1}^{N_{k-1}} \hat{\pi}^j_{y,k-1} \hat{H}^{ij}_{y,k},\ k = 1, \ldots, n$$

and $\hat{\pi}_0 = \sum_{i=0}^{N_0} \hat{p}^i_0 \delta_{x^i_0}$ with $\hat{p}^i_0 = \mathbb{P}(\hat{X}_0 = x^j_0)$, $i = 1, \ldots, N_0$. As a final step, we approximate the normalized filter $\Pi_{y,n}$ by $\hat{\Pi}_{y,n}$ given by

$$\hat{\Pi}_{y,n} f = \frac{\hat{\pi}_{y,n} f}{\hat{\pi}_{y,n} \mathbf{1}} = \sum_{i=1}^{N_n} \hat{\Pi}^i_{y,n} f(x^i_n) \quad \text{with} \quad \hat{\Pi}^i_{y,n} = \frac{\hat{\pi}^i_{y,n}}{\sum_{j=1}^{N_n} \hat{\pi}^j_{y,n}},\ i = 1, \ldots, N_n.$$

One shows (see [47]) that the un-normalized quantized filter may also be computed by the following *backward* induction formula, defined by

$$\hat{\pi}_{y,n} f = \hat{u}_{y,-1}(f)$$

where $\hat{u}_{y,-1}$ is the final value of the backward recursion:

$$\hat{u}_{y,n}(f) = f \ \text{ on } \Gamma_n, \quad \hat{u}_{y,k-1}(f) = \hat{H}_{y,k} \hat{u}_{y,k}(f) \ \text{ on } \Gamma_k, \quad k = 0, \ldots, n. \tag{19}$$

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 [47] where it has been shown that, for every bounded Borel 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 sum of the L^r-quantization errors $\|X_k - \hat{X}_k\|_r$, $k = 0, \ldots, n$. 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 similarly to what we did for BSDEs inducing a similar improvement for dependence in n of the error bounds ($i.e.$ the time discretization step $1/n$ if $(X_k)_{k \geq 0}$ is a discretization step of a diffusion). Once again, this confirms numerical evidences observed in [47, 6]. Secondly, we show that these improved error bounds hold under local Lipschitz continuity assumptions on the conditional density functions g_k (instead of Lipschitz conditions in [47]). The distortion mismatch property established in Theorem 4.3 is the key of this extension.

The paper is divided into three 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 the error analysis for the time discretization and the quantization scheme. In the second part, some results about optimal quantization are recalled in Section 4 and a new distortion mismatch theorem is established about the robustness of L^r-optimal quantization in L^s for $s \in (r, r + d)$. Some numerical tests confirm and illustrate these improved error bonds in Section 5. The final part, Section 6, is devoted to the nonlinear filtering problem analysis when estimating the nonlinear filter by optimal quantization with new improved error bounds obtained under less than stringent – local – Lipschiz assumptions than in the existing literature.

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

• For every $f : \mathbb{R}^d \to \mathbb{R}$, set $\|f\|_\infty = \sup_{x \in \mathbb{R}^d} |f(x)|$ and $[f]_{\rm Lip} = \sup_{x \neq y} \frac{|f(x) - f(y)|}{|x-y|} \leq +\infty$.

• If $A \in \mathcal{M}(d, q)$ we define the Fröbenius norm of A by $\|A\| = \sqrt{{\rm Tr}(AA^*)}$.

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:

$$X_t = x + \int_0^t b(s, X_s)ds + \int_0^t \sigma(s, X_s)dW_s, \tag{20}$$

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

$$Y_t = h(X_T) + \int_t^T f(s, X_s, Y_s, Z_s)ds - \int_t^T Z_s \cdot dW_s, \quad t \in [0,T], \tag{21}$$

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

$$(Lip_f) \;\equiv\; |f(t,x,y,z) - f(t,x',y',z')| \;\leq\; [f]_{\text{Lip}}(|x-x'| + |y-y'| + |z-z'|). \tag{22}$$

Under the previous assumptions on b, σ, h, f, the BSDE (21) has a unique $\mathbb{R}\times\mathbb{R}^q$-valued, $\mathcal{F}_t$-adapted solution (Y,Z) satisfying (see [52], see also [43])

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

Let us consider now $(\bar{X}_{t_k})_{k=0,\ldots,n}$ the discrete time Euler scheme with step $\Delta_n = \frac{T}{n}$ of the diffusion process $(X_t)_{t\in[0,T]}$:

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

where $t_k := \frac{kT}{n}$, $k = 0,\ldots,n$ and its continuous time counterpart, sometimes called *genuine Euler scheme* (we drop the dependence in n when no ambiguity) defined as an Itô process by

$$d\bar{X}_t = b(\underline{t}, \bar{X}_{\underline{t}})dt + \sigma(\underline{t}, \bar{X}_{\underline{t}})dW_t, \quad \bar{X}_0 = x. \tag{23}$$

where $\underline{t} = \frac{kT}{n}$ when $t \in [t_k, t_{k+1})$. In particular $(\bar{X}_t)_{t\in[0,T]}$ is an $\mathcal{F}_t$-adapted Itô process satisfying under the above assumptions made on b and σ (see $e.g.$ [11]) :

$$\forall\, p \in (0,+\infty), \qquad \Big\|\sup_{t\in[0,T]} |X_t|\Big\|_p + \sup_{n\geq 1}\Big\|\sup_{t\in[0,T]} |\bar{X}^n_t|\Big\|_p \leq C_{b,\sigma,p,T}\big(1+|x|\big)$$

and

$$\forall\, p\in (0,+\infty),\; \forall\, n\geq 1, \qquad \Big\|\sup_{t\in[0,T]} |X_t - \bar{X}^n_t|\Big\|_p \leq C'_{b,\sigma,p,T}\sqrt{\Delta_n}\big(1+|x|\big)$$

for a positive constant $C_{b,\sigma,p,T}$.

As a consequence, general existence-uniqueness results for BSDEs entail (see [53]) the existence of a unique solution $(\bar{Y}, \bar{Z})$ to the Markovian BSDE having the genuine Euler scheme $\bar{X}$ instead of X as a forward process. Then, we can apply the classical comparison result (Proposition 2.1 from [25]) 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 $C^{(i)}_{b,\sigma,f,T} > 0$, $i = 1,2$, such that

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

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 a second time discretization on the Euler scheme based BSDE, only involving discrete instants t_k, $k = 0, \dots, n$.

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

$$\begin{aligned}
\tilde{Y}_{t_n} &= h(\bar{X}_{t_n}) && (24)\\
\tilde{Y}_{t_k} &= \mathbb{E}(\tilde{Y}_{t_{k+1}}|\mathcal{F}_{t_k}) + \Delta_n f\big(t_k, \bar{X}_{t_k}, \mathbb{E}(\tilde{Y}_{t_{k+1}}|\mathcal{F}_{t_k}), \tilde{\zeta}_{t_k}\big) && (25)\\
\tilde{\zeta}_{t_k} &= \frac{1}{\Delta_n}\mathbb{E}\big(\tilde{Y}_{t_{k+1}}(W_{t_{k+1}} - W_{t_k})|\mathcal{F}_{t_k}\big),\ k = 0, \dots, n-1. && (26)
\end{aligned}$$

It slightly differs from the other explicit schemes analyzed in the literature to our knowledge, since the conditioning is applied directly to $\tilde{Y}_{t_{k+1}}$ *inside the driver function* rather than outside. Note that in many situations, one uses the following more symmetric alternative formula

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

which is clearly quite natural when thinking of a hedging term as a derivative ($e.g.$ computed in a binomial tree). It also has virtues in terms of variance reduction (see $e.g.$ [6]). 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 $\sup_{t\in[0,T]} |\bar{X}_t| \in L^2(\mathbb{P})$.

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 $(\bar{Y}, \bar{Z})$. 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

$$M_T = \frac{1}{2}\sum_{k=1}^{n} \tilde{Y}_{t_k} - \mathbb{E}\big(\tilde{Y}_{t_k}\,|\,\mathcal{F}_{t_{k-1}}\big).$$

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

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

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

$$\tilde{Y}_t = \tilde{Y}_{t_k} - (t - t_k)f\big(t_k, \bar{X}_{t_k}, \mathbb{E}(\tilde{Y}_{t_{k+1}}|\mathcal{F}_{t_k}), \tilde{\zeta}_{t_k}\big) + \int_{t_k}^{t} \tilde{Z}_s dW_s, \quad t \in [t_k, t_{k+1}]. \tag{27}$$

3 Error analysis

3.1 The time discretization error

We provide in the theorem below the quadratic error bound for the inside explicit time discretization scheme $(\tilde{Y}, \tilde{Z})$ defined by (24)-(41) and (27). The result is postponed to an Appendix for self-completeness. Like for most results of this type, the proof follows the lines of that devised for comparison theorems in [25]. In particular, though slightly more technical at some places, it is close to its counterpart for the standard outer explicit scheme originally established in [3] (in L^p for reflected BSDEs, but without Z on the driver) or in [57] (in the quadratic case, see also [27] for an extension error bounds in L^p or [10] for implicit scheme).

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

$$\forall t \geq 0, \quad |f(t,x,y,z)| \leq C(f)(1 + |x| + |y| + |z|). \tag{28}$$

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

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

(b) *Assume that the functions* b, σ, 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* $\sigma\sigma^\star$ *is uniformly elliptic. Then*

$$\int_0^T \mathbb{E}|Z_s - Z_{\underline{s}}|^2 ds \leq C^{(1)}_{b,\sigma,f,T} \Delta_n, \tag{29}$$

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

$$\max_{k=0,\ldots,n} \mathbb{E}|Y_{t_k} - \tilde{Y}_{t_k}|^2 + \int_0^T \mathbb{E}|Z_t - \tilde{Z}_t|^2 dt \leq \tilde{C}_{b,\sigma,f,T} \Delta_n. \tag{30}$$

NOTATIONS (CHANGE OF). The previous schemes (24)-(25) 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

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

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 $k = 0, \ldots, n$. This leads to the following result.

Theorem 3.2. *Assume that the drift* b *and the diffusion coefficient* σ *of the diffusion* $(X_t)_{t\in[0,T]}$ *defined by* (20) *are Lipschitz continuous, that the driver function* f *satisfies* (Lip_f) *(Assumption(22)) and that the function* h *is* $[h]_{\mathrm{Lip}}$*-Lipschitz continuous.*

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

$$\big\|\tilde{Y}_k - \hat{Y}_k\big\|_2^2 \le \sum_{i=k}^{n} e^{(1+[f]_{\rm Lip})t_i} K_i(b,\sigma,T,f)\big\|\bar{X}_i - \hat{X}_i\big\|_2^2, \tag{31}$$

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

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

with $\kappa_0 = C_{b,\sigma,T} + [f]_{\rm Lip}\Big(1 + \dfrac{[f]_{\rm Lip}}{2}\Big)$, $\kappa_1 = \dfrac{[f]_{\rm Lip}}{\kappa_0} + [h]_{\rm Lip}$,

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

$$C_{b,\sigma,T} = [b]_{\rm Lip} + \frac{1}{2}([\sigma]_{\rm Lip}^2 + T[b]_{\rm Lip}^2). \tag{32}$$

Furthermore, if $n \ge n_0$, one can take $C_{b,\sigma,T} = [b]_{\rm Lip} + \dfrac{1}{2}\left([\sigma]_{\rm Lip}^2 + \dfrac{T}{n_0}[b]_{\rm Lip}^2\right)$.

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

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

The proof is divided in two main steps: in the first one we establish the propagation of the Lipschitz property through the functions y_k and z_k involved in the Markov representation (24)-(25) 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 Lipchitz properties of the flow of the SDE like in [35]). As a second step, we introduce the quantization based scheme which is the counterpart of (24) and (25) for which we establish a backward recursive inequality satisfied by $\|\tilde{Y}_k - \hat{Y}_k\|_2^2$.

Remark 3.1. (*About the relationship between the temporal and the spacial partitions*) Owing to the non-asymptotic bound for the quantization (see Theorem 4.1 further), we deduce from the upper bound of Equation (31) that there exists some constants c_i, $i = 1, \ldots, n$ (only depending on the coefficients b and σ of the diffusion X) such that for every $k = 1, \ldots, n$,

$$\big\|\tilde{Y}_k - \hat{Y}_k\big\|_2^2 \le \sum_{i=k}^{n} c_i N_i^{-2/d}. \tag{33}$$

So, a natural question is to determine how to dispatch optimally the sizes $N_1, \cdots, N_n$ (for a fixed mesh of length n, given that X_0 is deterministic and, as such, perfectly quantized with $N_0 = 1$) of the quantization grids under the total "budget" constraint $N_1 + \cdots + N_n \le N$ of elementary quantizers (with $N \ge n$ and $N_k \ge 1$, for every $k = 1, \ldots, n$). This amounts (at least at time $k = 0$) to solving the constrained minimization problem

$$\min_{N_1+\cdots+N_n \le N} \sum_{i=1}^{n} c_i N_i^{-2/d},$$

whose solution reads $N_i = \left\lfloor \dfrac{c_i^{\frac{d}{d+2}}}{\sum_{k=1}^{n} c_k^{\frac{d}{d+2}}} N \right\rfloor \vee 1$, $i = 1, \ldots, n$. Coming back to (33), and using the Hölder inequality (to get the second inequality below) yields

$$\big\|\tilde{Y}_0 - \hat{Y}_0\big\|_2 \le N^{-1/d}\Big(\sum_{i=1}^{n} c_i^{\frac{d}{d+2}}\Big)^{1/2+1/d} \le \Big(\frac{n}{N}\Big)^{1/d}\Big(\sum_{i=1}^{n} c_i\Big)^{\frac{1}{2}} \le \Big[\max_{i=1,\ldots,n} c_i^{\frac{1}{2}}\Big]\frac{n^{1/2+1/d}}{N^{\frac{1}{d}}}. \tag{34}$$

Notice that for the standard ("non-improved) error bounds (see the introduction), the same optimal allocation procedure would yield (starting from $\|\tilde{Y}_0 - \hat{Y}_0\|_2 \le \sum_{i=0}^n c_i' N_i^{-1/d}$),

$$\|\tilde{Y}_0 - \hat{Y}_0\|_2 \le \left(\frac{n}{N}\right)^{1/d} \sum_{i=1}^n c_i' \le \Big[\max_{i=1,\dots,n} c_i'\Big] \frac{n^{1+1/d}}{N^{\frac{1}{d}}}$$

which emphasizes the improvement of the error bound as concerns the dependence in the time mesh size n.

3.2.1 First step toward the proof of Theorem 3.2: Lipschitz operators

As a first step we introduce several operators which appear naturally when representing Y_k. We will show that these operators propagate Lipschitz continuity. It is a classical step when establishing *a priori error bounds* going back to [4, 3], see also more recently [28] (Proposition 3.4). However we do not skip it since it emphasizes the technical specificities induced by our choice of an inner explicit scheme.

To be more precise, we set for every $k \in \{0, \dots, n-1\}$ and every Borel function $g : \mathbb{R}^d \to \mathbb{R}$ with polynomial growth

$$\begin{aligned} \mathcal{E}_k(x, u) &= x + \Delta_n b(t_k, x) + \sqrt{\Delta_n}\sigma(t_k, x)u, \ x \in \mathbb{R}^d, \ u \in \mathbb{R}^q && (35)\\ P_{k+1}g(x) &= \mathbb{E}\, g\big(\mathcal{E}_k(x, \varepsilon)\big) \quad \text{where } \varepsilon \sim \mathcal{N}(0; I_q) && (36)\\ Q_{k+1}g(x) &= \frac{1}{\sqrt{\Delta_n}} \mathbb{E}\Big(g\big(\mathcal{E}_k(x, \varepsilon)\big)\varepsilon\Big). && (37) \end{aligned}$$

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

$$\mathbb{E}_k g(\bar{X}_{k+1}) = P_{k+1}g(\bar{X}_k) \quad \text{and} \quad \mathbb{E}_k\big(g(\bar{X}_{k+1})(W_{t^n_{k+1}} - W_{t^n_k})\big) = \Delta_n Q_{k+1}g(\bar{X}_k).$$

Note that the process $(\bar{X}_k)_{0 \le k \le n}$ is an $(\mathcal{F}_k)_{0 \le k \le n}$-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 $[P_{k+1}]_{\rm Lip} := \sup_{f,\, [f]_{\rm Lip} \le 1} [P_{k+1}f]_{\rm Lip}$ is finite. More precisely, it satisfies:*

$$[P_{k+1}]_{\rm Lip} \le e^{\Delta_n C_{b,\sigma,T}} \tag{38}$$

where $C_{b,\sigma,T}$ is given by (32) *(see also the comment that follows).*

Proof. We have for every $x, x' \in \mathbb{R}^d$, and for every Lipschitz continuous function g

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

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

$$\begin{aligned} \mathbb{E}|\mathcal{E}_k(x, \varepsilon) - \mathcal{E}_k(x', \varepsilon)|^2 &\le \big(1 + \Delta_n(2[b(t^n_k, .)]_{\rm Lip} + [\sigma(t^n_k, .)]^2_{\rm Lip}) + \Delta_n^2 [b(t^n_k, .)]^2_{\rm Lip}\big)|x - x'|^2\\ &\le \big(1 + \Delta_n(2[b]_{\rm Lip} + [\sigma]^2_{\rm Lip}) + \Delta_n^2 [b]^2_{\rm Lip}\big)|x - x'|^2\\ &\le (1 + \Delta_n C_{b,\sigma,T})^2 |x - x'|^2\\ &\le e^{2\Delta_n C_{b,\sigma,T}} |x - x'|^2 \end{aligned}$$

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

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

$$y_n = h, \quad y_k = P_{k+1}y_{k+1} + \Delta_n f_k\big(\,.\,, P_{k+1}y_{k+1}, Q_{k+1}y_{k+1}\big),\ k=0,\ldots,n-1,$$

satisfies $\tilde{Y}_k = y_k(\bar{X}_k)$ for every $k\in\{0,\ldots,n\}$. Moreover, $\tilde{\zeta}_k = \frac{z_k(\bar{X}_k)}{\Delta_n}$ where, for every $k\in\{0,\ldots,n-1\}$,

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

(b) Furthermore, assume that the function h is $[h]_{\rm Lip}$-Lipschitz continuous and that the function $f(t,x,y,z)$ is $[f]_{\rm Lip}$-Lipschitz continuous in (x,y,z), uniformly in $t\in[0,T]$. Then, for every $k\in\{0,\ldots,n\}$, the function y_k is $[y_k]_{\rm Lip}$-Lipschitz continuous and there exists real constants $\kappa_0 = C_{b,\sigma,T} + [f]_{\rm Lip}(1+\frac{1}{2}[f]_{\rm Lip})$, and $\kappa_1 = \dfrac{[f]_{\rm Lip}}{\kappa_0} + [h]_{\rm Lip}$ (where $C_{b,\sigma,T}$ is given by (32)), such that

$$[y_k]_{\rm Lip} \le \frac{\Delta_n}{e^{\kappa_0\Delta_n}-1}\big(e^{\kappa_0(T-t_k^n)}-1\big)[f]_{\rm Lip} + e^{\kappa_0(T-t_k^n)}[h]_{\rm Lip}e^{\kappa_0(T-t_k^n)}\kappa_1.$$

In particular, $\sup_{n\ge1}\max_{k=0,\ldots,n}[y_k]_{\rm Lip} \le e^{\kappa_0 T}\kappa_1 < +\infty$. Moreover the functions z_k are Lipschitz too and

$$[z_k]_{\rm Lip} \le \sqrt{q}\,e^{\Delta_n C_{b,\sigma,T}}\kappa_1 e^{\kappa_0(T-t_{k+1}^n)},\ k=0,\ldots,n-1.$$

Proof. (a) We proceed by a backward induction using (24) and (25), relying on the fact that $(\bar{X}_k)_{k=0,\ldots,n}$ 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 (25) and the Markov property, we get

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

One shows likewise that $\zeta_k = Q_{k+1}(y_{k+1})(\bar{X}_k) = \frac{z_k}{\Delta_n}(\bar{X}_k)$, $k=0,\ldots,n-1$.

(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 $[h]_{\rm Lip}$-Lipschitz. Suppose that y_{k+1} is $[y_{k+1}]_{\rm Lip}$-Lipschitz continuous. Then, for every $x,x'\in\mathbb{R}^d$, we can write

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

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

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

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 $|A_{x,x'}|, |B_{x,x'}|, |C_{x,x'}| \le [f]_{\rm Lip}$. Now, taking advantage of the linearity of expectation, we get

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

Then Schwarz's Inequality yields

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

Now,

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

by Lemma 3.3. On the other hand, using that $|B_{x,x'}|$, $|C_{x,x'}| \leq [f]_{\rm Lip}$ and $\mathbb{E}(\varepsilon)=0$,

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

Finally, owing to the definition of κ_0, we get

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

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

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

The conclusion follows by induction. As for the functions z_k, we get for every $k=0,\ldots,n-1$,

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

Hence, using that $\varepsilon \sim \mathcal{N}(0;I_q)$,

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

□

3.2.2 Second step of the proof of Theorem 3.2

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

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

where $\hat{\mathbb{E}}_k=\mathbb{E}(\cdot\,|\,\hat{X}_k)$. 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 $k\in\{0,\ldots,n\}$, $\sigma(\hat{X}_k)\subset\sigma(\bar{X}_k)$, we have

$$\tilde{Y}_k-\hat{Y}_k=\tilde{Y}_k-\hat{\mathbb{E}}_k(\tilde{Y}_k)+\hat{\mathbb{E}}_k(\tilde{Y}_k-\hat{Y}_k) \tag{39}$$

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

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

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

$$\|\tilde{Y}_k - \hat{\mathbb{E}}_k(\tilde{Y}_k)\|_2^2 = \|y_k(\bar{X}_k) - \hat{\mathbb{E}}_k(y_k(\bar{X}_k))\|_2^2 \le \|y_k(\bar{X}_k) - y_k(\hat{X}_k)\|_2^2 \le [y_k]_{\rm Lip}^2\|\bar{X}_k - \hat{X}_k\|_2^2.$$

Let us consider now the last term of the equality (39). We have,

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

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

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

As

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

we deduce that

$$\begin{aligned}
\hat{\mathbb{E}}_k(\tilde{Y}_k - \hat{Y}_k) &= \hat{\mathbb{E}}_k\big[\big(\tilde{Y}_{k+1} - \hat{Y}_{k+1}\big)\big(1 + \Delta_n\hat{B}_k + \hat{C}_k\Delta W_{t_{k+1}}\big)\big] \\
&\quad + \Delta_n\big(f_k(\bar{X}_k, \mathbb{E}_k(\tilde{Y}_{k+1}), \tilde{\zeta}_k) - f_k(\hat{X}_k, \hat{\mathbb{E}}_k(\tilde{Y}_{k+1}), \hat{\mathbb{E}}_k(\tilde{\zeta}_k))\big).
\end{aligned} \tag{40}$$

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

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

First, from the very definition of conditional expectation operator $\hat{\mathbb{E}}_k$ as the best quadratic approximation by a Borel function of $\hat{X}_k$, we derive that

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

On the other hand, starting from $\zeta_k = \mathbb{E}_k(\tilde{Y}_{k+1}\Delta W_{t_{k+1}}) = \frac{z_k(\bar{X}_k)}{\Delta_n}$, $k = 0, \ldots, n-1$ (see Proposition 3.4(a), we get, using again the above characterization of the conditional expectation operator $\hat{\mathbb{E}}_k$,

$$\begin{aligned}
\|\tilde{\zeta}_k - \hat{\mathbb{E}}_k\tilde{\zeta}_k\|_2^2 &= \frac{1}{\Delta_n^2}\|\mathbb{E}_k(\tilde{Y}_{k+1}\Delta W_{t_{k+1}}) - \hat{\mathbb{E}}_k\big(\mathbb{E}_k(\tilde{Y}_{k+1}\Delta W_{t_{k+1}})\big)\|_2^2 \\
&= \frac{1}{\Delta_n}\|z_k(\bar{X}_k) - \hat{\mathbb{E}}_k(z_k(\bar{X}_k))\|_2^2 \\
&\le \frac{1}{\Delta_n}\|z_k(\bar{X}_k) - z_k(\hat{X}_k)\|_2^2 \le \frac{1}{\Delta_n}[z_k]_{\rm Lip}^2\|\bar{X}_k - \hat{X}_k\|_2^2.
\end{aligned} \tag{41}$$

Finally, using the upper-bound for $[z_k]_{\rm Lip}$ established in Proposition 3.4(b), we deduce that

$$\|f_k(\bar X_k, \mathbb{E}_k(\tilde Y_{k+1}), \tilde\zeta_k) - f_k(\hat X_k, \hat{\mathbb{E}}_k(\tilde Y_{k+1}), \hat{\mathbb{E}}_k(\tilde\zeta_k))\|_2 \le \Big(C_{1,k}(b,\sigma,T,f) + \frac{C_{2,k}(b,\sigma,T,f)}{\Delta_n}\Big)^{\frac12}\|\bar X_k - \hat X_k\|_2 \tag{42}$$

since, owing to (38) and (39), we have

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

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

$$C_{2,k}(b,\sigma,T,f) = q\kappa_1^2[f]_{\rm Lip}^2 e^{2\Delta_n C_{b,\sigma,T}+2\kappa_0(T-t_{k+1})} \ \text{ and } \ C_{1,k}(b,\sigma,T,f) = [f]_{\rm Lip}^2 + \frac{C_{2,k}(b,\sigma,T,f)}{q}. \tag{43}$$

To complete the proof, it suffices to control the remaining terms in Equation (40). Using the (conditional) Schwarz's inequality yields

$$\Big|\hat{\mathbb{E}}_k\big[\big(\tilde Y_{k+1}-\hat Y_{k+1}\big)\big(1-\Delta_n\hat B_k-\hat C_k\Delta W_{t_{k+1}}\big)\big]\Big| \le \big[\hat{\mathbb{E}}_k(\tilde Y_{k+1}-\hat Y_{k+1})^2\big]^{\frac12}\big[\hat{\mathbb{E}}_k(1-\Delta_n\hat B_k-\hat C_k\Delta W_{t_{k+1}})^2\big]^{\frac12}.$$

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

$$\begin{aligned}
\hat{\mathbb{E}}_k\big[(1-\Delta_n\hat B_k-\hat C_k\Delta W_{t_{k+1}})^2\big] &= (1-\Delta_n\hat B_k)^2 + \hat C_k^2\hat{\mathbb{E}}_k((\Delta W_{t_{k+1}})^2)\\
&= (1-\Delta_n\hat B_k)^2 + \hat C_k^2\Delta_n\\
&\le (1+\Delta_n[f]_{\rm Lip})^2 + (\Delta_n[f]_{\rm Lip})^2 \le e^{2\Delta_n[f]_{\rm Lip}}.
\end{aligned}$$

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

$$\Big\|\hat{\mathbb{E}}_k\big[\big(\tilde Y_{k+1}-\hat Y_{k+1}\big)\big(1-\Delta_n\hat B_k-\hat C_k\Delta W_{t_{k+1}}\big)\big]\Big\|_2 \le e^{\Delta_n[f]_{\rm Lip}}\|\tilde Y_{k+1}-\hat Y_{k+1}\|_2. \tag{44}$$

Using Schwarz's Inequality for the L^2-norm, we derive from (39), (40), (42) and (44) that

$$\begin{aligned}
\|\tilde Y_k - \hat Y_k\|_2^2 &= \|\tilde Y_k - \hat{\mathbb{E}}_k(\tilde Y_k)\|_2^2 + \|\hat{\mathbb{E}}_k(\tilde Y_k - \hat Y_k)\|_2^2 \qquad (45)\\
&\le [y_k]_{\rm Lip}^2\|\bar X_k-\hat X_k\|_2^2 + \Big(e^{\Delta_n[f]_{\rm Lip}}\|\tilde Y_{k+1}-\hat Y_{k+1}\|_2 + \Delta_n\|f_k(\bar X_k,\mathbb{E}_k(\tilde Y_{k+1}),\tilde\zeta_k)\\
&\quad -f_k(\hat X_k,\hat{\mathbb{E}}_k(\tilde Y_{k+1}),\hat{\mathbb{E}}_k(\tilde\zeta_k))\|_2\Big)^2\\
&\le [y_k]_{\rm Lip}^2\|\bar X_k-\hat X_k\|_2^2 + \Big(e^{\Delta_n[f]_{\rm Lip}}\|\tilde Y_{k+1}-\hat Y_{k+1}\|_2\\
&\quad +\Delta_n\Big(C_{1,k}(b,\sigma,T,f)+\frac{C_{2,k}(b,\sigma,T,f)}{\Delta_n}\Big)^{\frac12}\|\bar X_k-\hat X_k\|_2\Big)^2.
\end{aligned}$$

Relying on the classical identity

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

we derive that

$$\begin{aligned}
&\Big(e^{\Delta_n[f]_{\rm Lip}}\|\tilde Y_{k+1}-\hat Y_{k+1}\|_2 + \Delta_n\Big(C_{1,k}(b,\sigma,T,f)+\frac{C_{2,k}(b,\sigma,T,f)}{\Delta_n}\Big)^{\frac12}\|\bar X_k-\hat X_k\|_2\Big)^2\\
&\le e^{\Delta_n[f]_{\rm Lip}}(1+\Delta_n)\|\tilde Y_{k+1}-\hat Y_{k+1}\|_2^2\\
&\quad +\Big(1+\frac{1}{\Delta_n}\Big)\Delta_n^2\left(C_{1,k}(b,\sigma,T,f)+\frac{C_{2,k}(b,\sigma,T,f)}{\Delta_n}\right)\|\bar X_k-\hat X_k\|_2^2\\
&\le e^{\Delta_n(1+[f]_{\rm Lip})}\|\tilde Y_{k+1}-\hat Y_{k+1}\|_2^2 + \Big(1+\Delta_n\Big)\Big(C_{1,k}(b,\sigma,T,f)\Delta_n + C_{2,k}(b,\sigma,T,f)\Big)\|\bar X_k-\hat X_k\|_2^2.
\end{aligned}$$

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

$$\|\tilde{Y}_k - \hat{Y}_k\|_2^2 \;\leq\; e^{\Delta_n(1+[f]_{\rm Lip})}\|\tilde{Y}_{k+1} - \hat{Y}_{k+1}\|_2^2 + K_k(b,\sigma,T,f)\|\bar{X}_k - \hat{X}_k\|_2^2. \tag{46}$$

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

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

where

$$\begin{aligned}\widetilde{K}_k(b,\sigma,T,f) &:= [y_{k+1}]_{\rm Lip}^2 + \Big(1+\frac{T}{n}\Big)\Big(C_{1,k}(b,\sigma,T,f)\frac{T}{n_0} + C_{2,k}(b,\sigma,T,f)\Big),\ k=0,\ldots,n-1,\\ &\leq K_k(b,\sigma,T,f)\end{aligned}$$

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

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

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

$$\tilde{\zeta}_k - \hat{\zeta}_k = \big(\tilde{\zeta}_k - \hat{\mathbb{E}}_k(\tilde{\zeta}_k)\big) \overset{\perp}{+} \big(\hat{\mathbb{E}}_k(\tilde{\zeta}_k) - \hat{\zeta}_k\big)$$

where $\overset{\perp}{+}$ means that both random variables are L^2-orthogonal. We know from (41) in the the proof of claim (a) that

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

On the other hand, as $\sigma(\hat{X}_k) \subset \sigma(\bar{X}_k) \subset \mathcal{F}_k$, 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

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

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

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

which in turn implies that

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

so that finally

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

Remark 3.2. 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 (45) 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 $\hat{\zeta}_k$ terms

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

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

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

$$\hat{z}_k(x_i^k) = \frac{1}{2} \sum_{j=1}^{N_{k+1}} \hat{y}_{k+1}(x_j^{k+1})\, \pi_{ij}^{W,k}$$

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

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

These vector valued "weights" appear as new companion parameters (as well as the original weights π_{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 $k \in \{0, \ldots, n-1\}$ and for every $i \in \{1, \ldots, N_k\}$,

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

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

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

4 Background and new results on optimal vector quantization

It is important to have in mind that all what precedes holds true for any quantizations $\hat{X}_{t_k}$ of the Euler scheme $\bar{X}_{t_k}$ $i.e.$ for any sequence of the form $\hat{X}_{t_k} = \pi_k(\bar{X}_{t_k})$. 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 $\|\bar{X}_{t_k} - \hat{X}_{t_k}\|_2$ 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 the1940's motivated by Signal transmission and processing. These techniques have been imported in Numerical Probability, originally for numerical integration by cubature formulas in the 1990's (see [45] or [18]).

4.1 Short background

Let $X : (\Omega, \mathcal{A}, \mathbb{P}) \to \mathbb{R}^d$ 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 $\pi : \mathbb{R}^d \to \mathbb{R}$ 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 $|\,.\,|$ 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 $\|\,.\,\|_r^r$. As a consequence, finding the best approximation of X in the earlier described sense amounts to find the solution to the following minimization problem:

$$e_{N,r}(X) = \inf\Big\{\|X - \pi(X)\|_r, v : \mathbb{R}^d \to \Gamma,\, \Gamma \subset \mathbb{R}^d,\ \mathrm{card}(\Gamma) \le N\Big\},$$

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

$$|\xi - \pi(\xi)| \ge \mathrm{dist}(\xi, \Gamma) = \min_{1 \le i \le N} |\xi - x_i|.$$

Equality holds if and only if π is a Borel nearest neighbor projection π_Γ defined by

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

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

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

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

$$e_{N,r}(X) = \inf\big\{\|X - \hat{X}^\Gamma\|_r, \Gamma \subset \mathbb{R}^d,\ \mathrm{card}(\Gamma) \le N\big\}. \tag{47}$$

Note that for every level $N \ge 1$, the infimum in (47) is in fact a minimum $i.e.$ is attained at least at one grid (or codebook) Γ_N at least (see $e.g.$ [30] or [45]). Any such grid or any of the resulting Borel nearest neighbor projections is called an *L^r-optimal N-quantizer*.

One shows that if card(supp($\mathbb{P}_X$)) $\ge N$ then any optimal N-quantizer is of full size N. Furthermore (see again [30] or [45]), the L^r-mean quantization error $e_{N,r}(X)$ (sometimes denoted by $e_r(\Gamma_N, X)$) 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, $|\,.\,|$ *temporarily* denotes *any* norm on $\mathbb{R}^d$.

Theorem 4.1. Zador's Theorem (a) *Sharp asymptotic rate (see [30]): Let X be an $\mathbb{R}^d$-valued random vector such that $X \in L^{r+\delta}(\mathbb{P})$ for some real number $\delta > 0$ and let $\mathbb{P}_X = \varphi.\lambda_d + P_s$ denote the canonical Lebesgue decomposition of $\mathbb{P}_X$ where P_s its singular part of $\mathbb{P}_X$. Then*

$$\lim_{N \to +\infty} N^{r/d}(e_{N,r}(P))^r = J_{r,d}\, \|\varphi\|_{\frac{d}{d+r}} \in [0, +\infty) \tag{48}$$

$$\textit{with} \quad \|\varphi\|_{\frac{d}{d+r}} = \left(\int_{\mathbb{R}^d} \varphi^{\frac{d}{d+r}} d\lambda_d\right)^{\frac{d+r}{d}} \quad \textit{and} \quad J_{r,d,|.|} = \inf_{N \ge 1} N^{r/d} e^r_{N,r}(U([0,1]^d)) \in (0, +\infty) \tag{49}$$

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

(b) *Non-asymptotic bound (see [30, 40]). 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*

$$\forall\, N \ge 1, \quad e_{N,r}(X) \le C_{r,r',d}\, \sigma_{r'}(X) N^{-\frac{1}{d}}$$

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

Numerical aspects (few words about) From the numerical probability viewpoint, finding an optimal N-quantizer Γ 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, $|\,.\,|$ 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:

$$\forall\, x = (x_1, \ldots, x_N) \in (\mathbb{R}^d)^N, \quad D^X_{N,2}(x) \;=\; \int_{\mathbb{R}^d} \min_{1\leq i\leq N} |\xi - x_i|^2 d\mathbb{P}_X(\xi). \tag{50}$$

To any N-tuple $x = (x_1, \ldots, x_N) \in (\mathbb{R}^d)^N$, we associate its grid of values $\Gamma^x = \{x_1, \ldots, x_N\}$, so that $D^X_{N,2}(x) = \|X - \hat{X}^{\Gamma^x}\|_2^2$. 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 (see Theorem 4.2 in [30]). (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:*

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

Its gradient is given by

$$\nabla D^X_{N,2}(x) = 2\Big(\int_{C_i(\Gamma^x)} (x_i - \xi) d\mathbb{P}_X(\xi)\Big)_{i=1,\ldots,N}. \tag{51}$$

(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 card(supp$(\mathbb{P}_X)$) $\geq N$, when $x \in$* $\mathrm{argmin} D_{N,2}$

The result is a consequence of the interchange of the differentiation and the integral leading to (51) when formally differentiating (50) (see [30, 45]). Consequently, any N-tuple $x \in \mathrm{argmin} D_{N,2}$ satisfies

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

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

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

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. [26, 50]) or recursive stochastic gradient descent like the Competitive Learning Vector Quantization (CLVQ) algorithm (see [26, 45] or [48]) 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`. For more details about numerical methods we refer to the recent survey [46] and the references therein.

4.2 Distortion mismatch: L^s-robustness of L^r-optimal quantizers

The distortion mismatch problem is the following: when does an L^r-optimal sequence of quantizers $(\Gamma_N)_{N\geq 1}$ for a random variable X remain L^s-rate optimal for some $s > r$ (if $X \in L^s$) ? Or in more mathematical terms, if $X \in L^s$, $s > r$, when do we have for such a sequence of L^r-optimal quantizers

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

This problem has obvious applications in numerical probability since, for algorithmic reasons, one usually has access to optimal quadratic quantizers (see $e.g.$ the website `www.quantize.maths-fi.com`) whereas they are currently used in a non quadratic framework. What will be done in Section 6 for nonlinear filtering is precisely to take advantage of this result to strongly relax some growth assumptions on the conditional densities involved in the Kallianpur-Striebel formula.

The distortion mismatch problem was first addressed in [31] for various classes of distributions on $\mathbb{R}^d$, in particular for distributions having a *radial density* satisfying (an almost necessary) moment assumption of order higher than s. In the theorem below we extend this result to *all random vectors satisfying this moment condition*.

Theorem 4.3 (L^r-L^s-distortion mismatch)**.** *Let $X : (\Omega, \mathcal{A}, \mathbb{P}) \to \mathbb{R}^d$ 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 φ. Let $(\Gamma_N)_{N\geq 1}$ be an L^r-optimal sequence of grids and let $s \in (r, r + d)$. If*

$$X \in L^{\frac{sd}{d+r-s}+\delta}(\mathbb{P}) \tag{52}$$

for some $\delta > 0$, then

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

Note that, as expected, $\frac{sd}{d+r-s} > s$ so the preservation of the rate optimality for $s > r$ requires more than L^s- integrability.

We will say that X *has an* (r, s)*-distribution* if the moment condition (52) is satisfied. Also remark that the function $s \mapsto \frac{sd}{d+r-s}$ is non-decreasing, so that if X has an (r, s)-distribution, then it has an (r, s') distribution for any $s' \leq s$.

Proof. 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 $d(\Gamma_N, \text{supp}(P)^c) \to 0$ as $N \to +\infty$ 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 Γ_N, $N \geq 1$, allow to apply to the micro-macro inequality (see Equation (3.2) in the proof of Theorem 2 in [31]), namely : for every real constant $c \in (0, \frac{1}{2})$ and every $y \in \mathbb{R}^d$,

$$e_r(\Gamma_N, P)^r - e_r(\Gamma_{N+1}, P)^r \geq \big((1-c)^r - c^r\big) P\big(B\big(y; cd(y, \Gamma_N)\big)\big) d(y, \Gamma_N)^r. \tag{53}$$

Let ν be an auxiliary Borel probability measure on $\mathbb{R}^d$ to be specified further on. Set $C(r) = (1-c)^r - c^r$. Integrating the above inequality with respect to $\nu(dy)$ yields, owing to Fubini's Theorem ,

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

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

$$\big\{(\xi,y)\,:\,|y-\xi|\leq \frac{c}{c+1}d(\xi,\Gamma_N)\big\} \subset \Big\{(\xi,y)\,:\,|y-\xi|\leq cd(y,\Gamma_N),\; d(y,\Gamma_n)\geq \frac{1}{c+1}d(\xi,\Gamma_N)\Big\}$$

and, still by Fubini's Theorem,

$$e_r(\Gamma_N, P)^r - e_r(\Gamma_{N+1}, P)^r \geq \frac{C(r)}{(1+c)^r}\int \nu\Big(B\big(\xi; cd(y,\Gamma_N)\big)\Big)d(\xi,\Gamma_N)^r P(d\xi). \tag{54}$$

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

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

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

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

and $V_d = \lambda_d\big(B(0;1)\big)$. 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

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

Let $s\in[r,r+d)$. It follows from Equation (54) 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 $\delta > 0$,

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

As a consequence

$$e_r(\Gamma_N, P)^r - e_r(\Gamma_{N+1}, P)^r \geq C_{X,r,s,\varepsilon,\delta}\, e_s(\Gamma_N, X)^{r+d} \tag{55}$$

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

STEP 3 (*Upper-bound for the quantization error increments*): Since the distribution of X is absolutely continuous X ($i.e.$ admits a density), one derives following the lines of the proof of Theorem 2 in [31] this upper-bound for the increments of the L^r-quantization error: there exists of a real constant $\kappa_{X,r} > 0$ such that

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

Combining this inequality with (55) yields

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

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

Remarks. • If φ is radial, more precisely if $\varphi = \tilde{\varphi}(|x|_0)$ where $\tilde{\varphi} : \mathbb{R}_+ \to \mathbb{R}_+$ is bounded and non-increasing on $[R_0, +\infty)$ and $|\,.\,|_0$ denotes any norm on $\mathbb{R}^d$ the above result holds true even if $\delta = 0$ (see [41]).

• This result is in close to optimality for the following reason. It has been established in [31] (Theorem 1) that if $X \in L^{r+}(\mathbb{P})$, and if $(\Gamma_n)_{N\ge1}$ is a sequence of L^r-asymptotically optimal quantization grids, then

$$\varliminf_{N} N^{\frac{1}{d}} e_s(\Gamma_n, X) \ge J^{\frac{1}{s}}_{r,d,|.|} \left[\int_{\mathbb{R}^d} \varphi^{\frac{d}{r+d}} d\lambda_d\right]^{\frac{1}{d}} \left[\int_{\mathbb{R}^d} \varphi^{\frac{d+r-s}{r+d}} d\lambda_d\right]^{\frac{1}{s}}$$

where $J^{\frac{1}{s}}_{r,d,|.|}$ is given by (49). Since $X \in L^{r+}(\mathbb{P})$, $\int_{\mathbb{R}^d} \varphi^{\frac{d}{r+d}} d\lambda_d < +\infty$ by an elementary application of the inverse Minkowski inequality (see Equation (2.11) from [31]). On the other hand,

$$\int_{\mathbb{R}^d} \varphi^{\frac{d+r-s}{r+d}} d\lambda_d = +\infty \Longrightarrow X \notin L^{\frac{ds}{d+r-s}}(\mathbb{P}).$$

5 Numerical experiments for the BSDE scheme

To illustrate empirically the improved theoretical rate obtained in the previous section, we deal here with two toy examples: 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. Note that our aim is not to make an extensive numerical test with a complete description (or a complexity analysis) of several used algorithms for the optimal grid search. These subjects have extensively been considered on the past and we refer for example to [48] for more details.

Numerical tests are performed using our quantized BSDE algorithm. At each discretization instant t_k, we associate a quantization grid $\Gamma_k = \{x_i^k,\, i = 1,\ldots,N_k\}$ of size N_k, possibly not optimal *a priori*, and $\hat{X}_k = \mathrm{Proj}_{\Gamma_k}(\bar{X}_k)$ the resulting Voronoi quantization of $\bar{X}_{t_k}$. Then, we set for every $k = 0,\ldots,n-1$, $i = 1,\ldots,N_k$, $j = 1,\ldots,N_{k+1}$, the *transition weights* (or probabilities)

$$p_{ij}^k = \mathbb{P}(\hat{X}_{k+1} = x_j^{k+1} |\, \hat{X}_k = x_i^k),\ k = 0,\ldots,n-1.$$

and, for $k = 0,\ldots,n$, $i = 1,\ldots,N_k$, the *marginal weights* $p_i^k = \mathbb{P}(\hat{X}_k = x_i^k)$, $k = 0,\ldots,n$.

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

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

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

$$\hat{\alpha}_k(x_i^k) = \sum_{j=1}^{N_{k+1}} \hat{y}_{k+1}(x_j^{k+1})\, p_{ij}^k \quad \text{and} \quad \hat{\beta}_k(x_i^k) = \frac{1}{\Delta_n} \sum_{j=1}^{N_{k+1}} \hat{y}_{k+1}(x_j^{k+1})\, \pi_{ij}^{W,k}, \tag{56}$$

with

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

We use a time discretization mesh of length $n = 20$ for the first example and of length $n = 10$ for all dimensions in the second example. In both examples below, the quantizers $\hat{X}_k$, $k = 1, \ldots, n$ (with $\hat{X}_0 = X_0$) are computed from a scaling of the optimal grid of $\mathcal{N}(0, I_d)$ Gaussian distributions available on the website devoted to quantization `www.quantize.maths-fi.com`. The transition probabilities are approximated using a Monte Carlo simulation of size 10^7 for all examples (keep in mind that we may have the same precision with a smaller size of Monte Carlo trials but our aim is not to optimize these sizes of the trials). For simplicity reasons, we use a uniform dispatching across teh time lauers for the quantizers by assigning the same grid size N_k to all $\hat{X}_k$'s at every discretization step t_k, $k = 1, \ldots, n$.

5.1 Bid-ask spread for interest rate

Let us consider a model with two interest rates introduced in [8]: a borrowing rate R and a lending rate $r \leq R$ where the stock price $(X_t)_{t\in[0,T]}$ evolves following the Black-Scholes dynamics

$$dX_t = \mu X_t dt + \sigma X_t dW_t, \; X_0 = x_0 > 0.$$

Let φ_t be the amount of assets held at time t. Then, the dynamics of the replicating portfolio is given by

$$Y_t = Y_T + \int_t^T f(Y_s, Z_s) ds - \int_t^T Z_s dW_s \tag{57}$$

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

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

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 (57) has no analytical solution. We refer to the reference prices given in [7, 55] where (Y_0, Z_0) is approximated by $(2.96, 0.55)$. We put $n = 20$ and, for every $k = 1, \ldots, n$, the grid sizes $N_k = \bar{N} = \frac{N}{n}$ is constant (keep in mind that $N = N_1 + \ldots + N_n$). The quantizers $\hat{X}_{t_k}$ have been obtained by using dilatations of optimal Gaussian quantization grids that we subsitute to W_{t_k} into the formula $X_{t_k} = x_0 e^{(\mu - \frac{\sigma^2}{2})t + \sigma W_{t_k}}$.

The numerical convergence rate of the error $\bar{N} \mapsto |Y_0 - \hat{Y}_0^{\bar{N}}|$, $\bar{N} = 5\ell$, $\ell = 1, \ldots, 30$, is depicted in Figure 1, including a polynomial regression which emphasizes the empirical order of the convergence rate, namely $\bar{N}^{-1}$.

5.2 Multidimensional example

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

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

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

$$Y_t = \frac{e_t}{1+e_t}, \qquad Z_t = \frac{e_t}{(1+e_t)^2}, \tag{58}$$

with

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

For the numerical experiments, we put the (regular) time discretization mesh to $n = 10$. We choose $t = 0.5$, $d = 2, 3, 4, 5$, so that $Y_0 = 0.5$ and $Z_0^i = 0.24$, for every $i = 1, \ldots, d$. We depict in Figures 2 and 3, the rates of convergence of $|\hat{Y}_0^{\bar{N}} - 0.5|$ towards 0, for the (constant) layer grid sizes $N_k = \bar{N} = \frac{N}{n} = 5, \ldots, 150$. The graphics in Figures 2 and 3 confirm a rate of convergence of order $N^{-1/d}$.

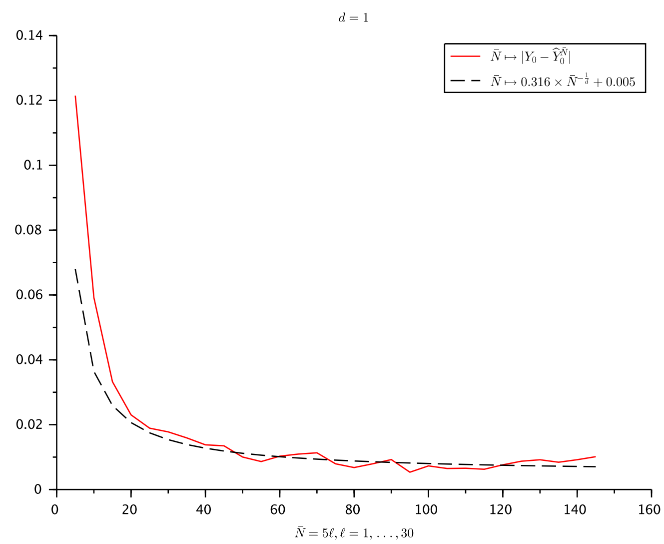

Figure 1: Convergence rate of the quantization error for the Bid-ask spread in the Black-Scholes model. Abscissa axis: the size $\bar{N} = 5\ell, \ell = 1, \ldots, 30$ of the quantization. Ordinate axis: The error $|Y_0 - \hat{Y}_0^{\bar{N}}|$ and the graph $\bar{N} \mapsto \hat{a}/\bar{N} + \hat{b}$, where $\hat{a}$ and $\hat{b}$ are the regression coefficients.

6 Nonlinear filtering problem

We consider in this section the discrete time nonlinear filtering model and the quantization based numerical scheme presented in the introduction. Our aim is two-fold: improving the error bounds like for BSDE on the one hand and relaxing the Lipschitz continuity on the conditional densities g_k (in favor of a local Lipschitz continuity assumption with polynomial growth). In particular, these new error bounds confirm the results obtained in the survey [56] devoted to a comparison between quantization and particle based numerical methods for non-linear filtering.

6.1 Error analysis

Let us first recall the assumptions made in [47] on the conditional transition density functions g_k and the Markov transitions P_k:

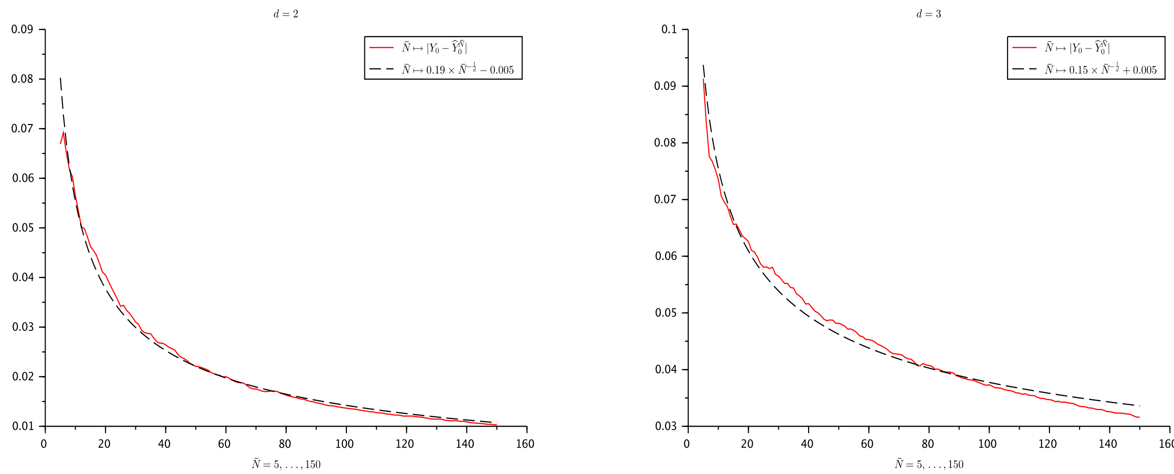

Figure 2: Convergence rate of the quantization error for the multidimensional example. Abscissa axis: the size $\bar{N} = 5, \ldots, 150$ of the quantization. Ordinate axis: The error $|Y_0 - \hat{Y}_0^{\bar{N}}|$ and the graph $\bar{N} \mapsto \hat{a}\bar{N}^{-1/d} + \hat{b}$, where $\hat{a}$ and $\hat{b}$ are the regression coefficients. The left hand side graphic corresponds to the dimension $d = 2$ and the right hand side to $d = 3$.

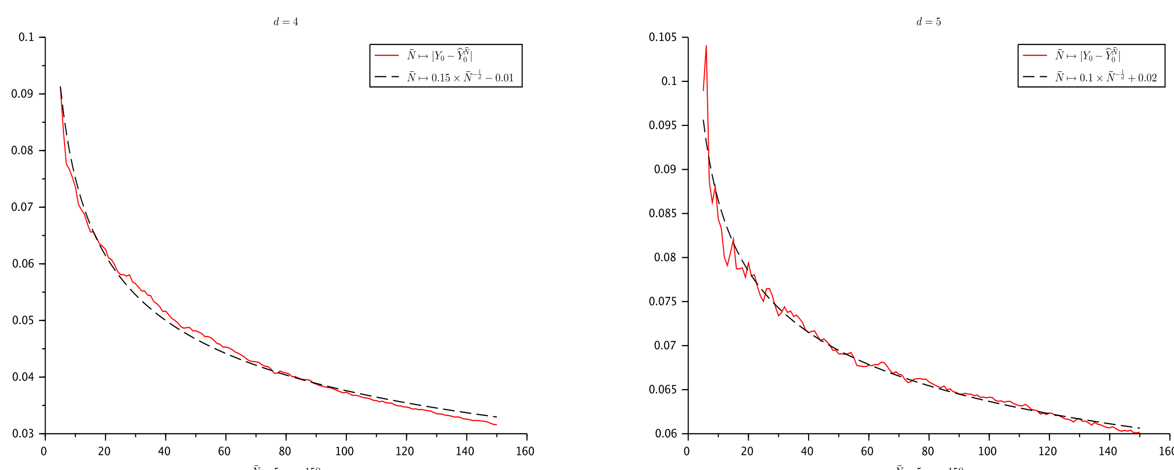

Figure 3: Convergence rate of the quantization error for the multidimensional example. Abscissa axis: the size $\bar{N} = 5, \ldots, 150$ of the quantization. Ordinate axis: The error $|Y_0 - \hat{Y}_0^{\bar{N}}|$ and the graph $\bar{N} \mapsto \hat{a}\bar{N}^{-1/d} + \hat{b}$, where $\hat{a}$ and $\hat{b}$ are the regression coefficients. The left hand side graphic corresponds to the dimension $d = 4$ and the right hand side to $d = 5$.

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

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

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

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

(ii) For every $k = 1, \ldots, n$, 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

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

We will relax these Lipschitz assumptions into controlled Lipschitz assumptions. Let us consider, for a fixed *non-negative* function $\theta : \mathbb{R}^d \mapsto \mathbb{R}_+$ satisfying,

$(\mathscr{I}_\theta) \equiv \qquad \forall\, k \in \{1, \ldots, n\}, \quad \mathbb{E}(\theta(X_k)) < +\infty.$

We make the following θ-local Lipschitz continuity assumption (which is weaker than $(\mathscr{H}_0)$) on the growth of the conditional transition density functions g_k:

$(\mathscr{H}_{\rm Liploc}) \equiv$ There exists $[g_k^1]_{\rm Liploc}, [g_k^2]_{\rm Liploc} : \mathbb{R}^q \times \mathbb{R}^q \mapsto \mathbb{R}_+$ such that, for every $k \in \{1, \ldots, n\}$,

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

A standard situation is the sometimes called $Li(1, \alpha)$ framework when the g_k satisfy $(\mathscr{H}_{\rm Liploc})$ with the function $\theta : x \mapsto \theta(x) = |x|^\alpha$ for an $\alpha \ge 0$, namely

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

$$\begin{aligned}|g_k(x, y, x', y') - g_k(\hat{x}, y, \hat{x}', y')| \;\le\;& [g_k^1]_{\rm pol}(y, y')(1 + |x|^\alpha + |\hat{x}|^\alpha + |x'|^\alpha + |\hat{x}'|^\alpha)|x - \hat{x}|\\ &+ [g_k^2]_{\rm pol}(y, y')(1 + |x|^\alpha + |\hat{x}|^\alpha + |x'|^\alpha + |\hat{x}'|^\alpha)|x' - \hat{x}'|.\end{aligned}$$

When $\theta \equiv 0$, this framework coincides with the Lipschitz one. To simplify some statement we will introduce

$$[g^i]_{\rm Liploc}(y, y') = \max_{k=1,\ldots,n} [g_k^i]_{\rm Liploc}(y, y'), \quad i = 1, 2. \tag{59}$$

Then we ask the transitions $P_k(x, dy)$ to propagate this θ-local Lipschitz property as a counterpart of $(\mathscr{A}_1)$. Let $f : \mathbb{R}^d \to \mathbb{R}$ be θ-locally Lipschitz with a local Lipschitz coefficient $[f]_{\rm Liploc}$ defined by

$$[f]_{\rm Liploc} = \sup_{x \neq x'} \frac{|f(x) - f(x')|}{\big(1 + \theta(x) + \theta(x')\big)\,|x - x'|} < +\infty. \tag{60}$$

$$(\mathscr{A}_{1,\rm loc}) \equiv [P]_{\rm Liploc} = \max_{k=1,\ldots,n} [P_k]_{\rm Liploc} < +\infty \;\text{ where }\; [P_k]_{\rm Liploc} = \frac{1}{2} \sup_{[f]_{\rm Liploc} \le 1} [P_k f]_{\rm Liploc}.$$

The following classical lemma is borrowed (and straghtforwardly adapted) from [47] (Lemma 3.1).

Lemma 6.1. *Let μ_y and ϑ_y be two families of finite and positive measure on a measurable space $(E, \mathcal{E})$. Suppose that there exists two symmetric functions R and S defined on the set of positive finite measures such that for every bounded θ-Lipschitz function f,*

$$\left|\int f d\mu_y - \int f d\vartheta_y\right|^2 \le \|f\|_\infty^2 R(\mu_y, \vartheta_y) + [f]_{\rm Liploc}^2 S(\mu_y, \vartheta_y). \tag{61}$$

Then,

$$\left|\frac{\int f d\mu_y}{\mu_y(E)}-\frac{\int f d\vartheta_y}{\vartheta_y(E)}\right| \le 4\frac{\|f\|_\infty^2 R(\mu_y,\vartheta_y)+\frac{1}{2}[f]^2_{\rm Liploc}S(\mu_y,\vartheta_y)}{\mu_y(E)\vee\vartheta_y(E)}. \tag{62}$$

In Theorem 6.3 below we will consider Assumption $(\mathscr{H}_{\rm Liploc})$ in place of Assumption $(\mathscr{H}_0)$ (considered [47]) to derive an error bound. This less stringent assumption is compensated by taking advantage of the *distortion mismatch* property established in Theorem 4.3. More precisely, we will 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 $N^{-\frac{1}{d}}$.

The following lemma provides a control of the θ-local Lipschitz coefficients of the functions $u_{y,k}(f)$ defined recursively by (13) (we drop the in the subscript y and f to alleviate notations in what follows).

Proposition 6.2. (a) *Assume that* $(\mathscr{I}_\theta)$, $(\mathscr{H}_{\rm Liploc})$ *and* $(\mathscr{A}_{1,\rm loc})$ *hold and that, for every* $k=1,\ldots,n$,

$$\mathbb{E}\big(\theta(X_k)|X_{k-1}=x\big)\le C_{\theta,X}(1+\theta(x)). \tag{63}$$

Let $f:\mathbb{R}^d\to\mathbb{R}^d$ *be* θ*-locally Lipschitz function. Then, the functions* u_k *defined by* (13) *satisfy*

$$[u_k]_{\rm Liploc} \;\le\; K_g^{n-k}\left[\kappa_{g,X}\frac{[P]^{n-k}_{\rm Liploc}-1}{[P]_{\rm Liploc}-1}\|f\|_\infty+[P]^{n-k}_{\rm Liploc}[f]_{\rm Liploc}\right] \text{ (Convention: } \tfrac{1^n-1}{1-1}=n\text{)} \tag{64}$$

where $\kappa_{g,X}=2C_{\theta,X}[g^1]_{\rm Liploc}+[P]_{\rm Liploc}[g^2]_{\rm Liploc}$ *and* $\|u\|_\infty\le(K_g)^n\|f\|_\infty$.
(b) *Let* $(X_k)_{k=0,\ldots,n}$ *be the Markov chain defined as an iterated random map of the form*

$$X_k=F_k(X_{k-1},\varepsilon_k),\qquad k=1,\ldots,n \tag{65}$$

where $(\varepsilon_k)_{k=1,\ldots,n}$ *is an i.i.d sequence of random variables independent of* X_0.

(i) *If there exists some* $p\in(0,+\infty)$ *such that*

$$\theta(X_0)\in L^p \quad \text{and} \quad \big\|\theta(F_k(x,\varepsilon_1))\big\|_p\le C'_{\theta,X}(1+\theta(x)) \tag{66}$$

then $\max_{k=0,\ldots,n}\|\theta(X_k)\|_p<+\infty$.
In particular, for $p=1$*, the chain satisfies the integrability assumption* $(\mathscr{I}(\theta))$ *and* (63).

(ii) *If* $F_k(0,\varepsilon_1),\theta(X_0)\in L^2$ *and, for every* $k\in\{1,\ldots,n\}$, $x,\,x'\in\mathbb{R}^d$,

$$\big\|\theta(F_k(x,\varepsilon_1))\big\|_2\le C'_{\theta,X}(1+\theta(x)) \quad\text{and}\quad \|F_k(x,\varepsilon_1)-F_k(x',\varepsilon_1)\|_2\le[F_k]_{2,\rm Lip}|x-x'|, \tag{67}$$

then both $(\mathscr{I}(\theta))$ *and* $(\mathscr{H}_{\rm Liploc})$ *are satisfied. To be more precise*

$$\forall\, k=1,\ldots,n,\quad [P_k]_{\rm Liploc}\le\max(1,C'_{\theta,X})[F_k]_{2,\rm Lip}.$$

Proof. (a) By the Markov property, we have for every $k=0,\ldots,n-1$ and every $x\in\mathbb{R}^d$,

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

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

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

since $\|u_n\|_\infty=\|f\|_\infty$. Let $k\in\{0,\ldots,n-1\}$; for every $x,x'\in\mathbb{R}^d$,

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

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

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

so that

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

Finally, collecting these inequalities, we deduce from Assumption (63) that, for every $k = 0, \ldots, n-1$,

$$\begin{aligned}[u_k]_{\mathrm{Liploc}} \;&\leq\; \big(2C_{\theta,X}[g^1_{k+1}]_{\mathrm{Liploc}} + [P_{k+1}]_{\mathrm{Liploc}}[g^2_{k+1}]_{\mathrm{Liploc}}\big)\|u_{k+1}\|_\infty + K_g[P_{k+1}]_{\mathrm{Liploc}}[u_{k+1}]_{\mathrm{Liploc}}\\ &\leq\; \kappa_{g,X}\|u_{k+1}\|_\infty + K_g[P]_{\mathrm{Liploc}}[u_{k+1}]_{\mathrm{Liploc}}.\end{aligned}$$

The conclusion follows by induction (discrete time Gronwall lemma) having in mind that $u_n = f$.

(b) Claim (i) is obvious. As for claim (ii), let f be a θ-locally Lipschitz with constant $[f]_{\mathrm{Liploc}}$. Then, for every $x, x' \in \mathbb{R}^d$ and every $k = 1, \ldots, n$,

$$\begin{aligned}|P_kf(x) \;-\; P_kf(x')| &= |\mathbb{E}f(F_k(x,\varepsilon_1)) - \mathbb{E}f(F_k(x',\varepsilon_1))|\\ &\leq\; [f]_{\mathrm{Liploc}}\mathbb{E}\Big(\big|F_k(x,\varepsilon_1) - \mathbb{E}F_k(x',\varepsilon_1)\big|\big(1+\theta(F_k(x,\varepsilon_1)) + \theta(F_k(x',\varepsilon_1))\big)\Big)\\ &\leq\; [f]_{\mathrm{Liploc}}\big\|F_k(x,\varepsilon_1) - F_k(x',\varepsilon_1)\big\|_2\big(1+\|\theta(F_k(x,\varepsilon_1))\|_2 + \|\theta(F_k(x,\varepsilon_1))\|_2\big)\\ &\leq\; [f]_{\mathrm{Liploc}}[F_k]_{2,\mathrm{Lip}}|x-x'|\big(1+C'_{\theta,X}\theta(x) + C'_{\theta,X}\theta(x')\big)\end{aligned}$$

where we used Schwarz' Inequality in the third line, and (67) and (66) in the last line. We deduce that $[P_k]_{\mathrm{Liploc}} \leq (1 \vee C'_{\theta,X})$, for every $k = 1, \ldots, n$. □

Notice that assumptions (63) and (66) hold when θ is a polynomial convex function and when $(X_k)_{0\leq k\leq n}$ is the Euler scheme (with step $\frac{T}{n}$ and horizon T) associated with a stochastic differential equation of the form (20). In the latter case, the transition operator $P_{k+1} = P$, $k = 0, \ldots, n-1$, is time homogenous. We suppose in the sequel that (P_k) is time homogenous.

Theorem 6.3. *Let $(\mathscr{H}_{\mathrm{Liploc}})$ holds and assume that $(\mathscr{A}_{1,loc})$ is fulfilled, as well as assumptions of Proposition 6.2. Suppose that for every $k = 0, \ldots, n$, X_k has a $(2, 2+\nu_k)$-distribution for some $\nu_k \in (0,d)$, and set $\bar\nu_n = \min_{k=0,\ldots,n}\nu_k/2$. Then for every $\nu \in (0, \bar\nu_n)$,*

$$|\Pi_{y,n}f - \hat\Pi_{y,n}f|^2 \leq \frac{4(M_{n,\nu}K^n_g)^2}{\phi^2_n(y) \vee \hat\phi^2_n(y)}\sum_{k=0}^n B^n_k(f,y,\alpha)\|X_k - \hat X_k\|^2_{2(1+\nu)} \tag{68}$$

where

$$\phi_n(y) = \pi_{y,n}\mathbf{1} \quad \textit{and} \quad \hat\phi_n(y) = \hat\pi_{y,n}\mathbf{1}$$

and

$$B^n_k(f,y) := 2^k\left(2\kappa^2_{g,X}\left(\frac{[P]^{n-k}_{\mathrm{Liploc}} - 1}{[P]_{\mathrm{Liploc}} - 1}\right)^2 + \frac{2[g^1]^2_{\mathrm{Liploc}} + [g^2]^2_{\mathrm{Liploc}}}{K^2_g} + [P]^{2(n-k)}_{\mathrm{Liploc}}\right)$$

with $\kappa_{g,X} = 2C_{\theta,X}[g^1]_{\mathrm{Liploc}} + [P]_{\mathrm{Liploc}}[g^2]_{\mathrm{Liploc}}$ and

$$M_{n,\nu} := 1 + \max_{k=0,\ldots,n-1}\Big(\|\theta(X_k)\|_{2(1+\frac{1}{\nu})} + \|\theta(\hat X_k)\|_{2(1+\frac{1}{\nu})} + \|\theta(X_{k+1})\|_{2(1+\frac{1}{\nu})} + \|\theta(\hat X_{k+1})\|_{2(1+\frac{1}{\nu})}\Big).$$

Let us make few remarks about the assumptions of the theorem before dealing with the proof.

Remark 6.1. (a) If θ is convex and if all $\hat{X}_k$ are quadratic optimal quantizer, then it is a stationary $i.e.$ satisfies $\hat{X}_k = \mathbb{E}(X_k \,|\, \hat{X}_k)$ so that, for every $k = 0, \ldots, n$, we have, owing to the convexity of $\theta^{2(1+\frac{1}{\nu})}$ and Jensen's Inequality,

$$\|\theta(\hat{X}_k)\|_{2(1+\frac{1}{\nu})} \leq \|\theta(X_k)\|_{2(1+\frac{1}{\nu})} < +\infty.$$

(b) If $(X_k)_{k=0,\ldots,n}$ is a Markov chain of iterated random maps

$$X_k = F_k(X_{k-1}, \varepsilon_k), \quad k = 1, \ldots, n,$$

under the assumptions of Proposition 6.2. Assume it satisfy the θ-local Lipschitz assumption with a function $\theta(y) \geq C|y|^a$ for some real constants $C, a > 0$. Then, if $a > \frac{1}{2}\frac{(2+\nu)d}{d-\nu}$ for some $\nu \in (0, d)$, and the distributions of X_k are absolutely continuous distribution. Then, all X_k have a $(2, 2+\nu)$-distribution.

Proof. Like in [47], the proof relies on the backward formulas (13) and (19) involving the functions $u_{y,k}(f)$ and their quantized counterpart $\hat{u}_{y,k}(f)$ whose final values u_{-1} and $\hat{u}_{-1}$ define the un-normalized filter $\pi_{y,n}(f()$ (applied to the function f) and its quantized counterpart, respectively.

Following the lines of the proof of Theorem 3.1 in [47], one shows by a backward induction taking advantage of the Markov property that the functions $u_k : \mathbb{R}^d \to \mathbb{R}$, $k = 0, \ldots, n$ defined recursively by (13) satisfy $u_n = f$ and

$$u_k(X_k) = \mathbb{E}_k\big(\varphi_{k+1}(X_k, X_{k+1}, X_{k+1})\big) = \mathbb{E}\big(\varphi_{k+1}(X_k, X_{k+1}, X_{k+1}) \,|\, X_k\big),\ k = 0, \ldots, n-1. \tag{69}$$

where

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

Finally $u_1 = \mathbb{E}(u_0(X_0)) = \pi_{y,n}(f)$ (un-normalized filter applied to f). One shows likewise that the functions $\hat{u}_k$ defined by (19) satisfy $\hat{u}_n = f$ (on the grid Γ_n) and

$$\hat{u}_k(\hat{X}_k) = \hat{\mathbb{E}}_k\big(\hat{u}_{k+1}(\hat{X}_{k+1})g_{k+1}(\hat{X}_k, \hat{X}_{k+1})\big),\ k = 0, \ldots, n-1, \tag{70}$$

so that finally $\hat{u}_{-1}(f) = \mathbb{E}\hat{u}_0(\hat{X}_0)) = \hat{\pi}_{y,n}(f)$ (quantized un-normalized filter). One shows like for the functions u_k in Proposition 6.2 that $\|\hat{u}_k\|_\infty \leq K_g^{n-k}\|f\|_\infty$. Now, using the definition of conditional expectation $\hat{\mathbb{E}}_k$ as an orthogonal projector (hence an L^2-contraction as well), we have

$$\begin{aligned}
\|u_k(X_k) - \hat{u}_k(\hat{X}_k)\|_2^2 &= \|u_k(X_k) - \hat{\mathbb{E}}_k(u_k(X_k))\|_2^2 + \|\hat{\mathbb{E}}_k(u_k(X_k)) - \hat{u}_k(\hat{X}_k)\|_2^2 \\
&\leq \|u_k(X_k) - \hat{\mathbb{E}}_k(u_k(X_k))\|_2^2 \\
&\quad + \|\varphi_{k+1}(X_k, X_{k+1}, X_{k+1}) - \hat{u}_{k+1}(\hat{X}_{k+1})g_{k+1}(\hat{X}_k, \hat{X}_{k+1})\|_2^2
\end{aligned} \tag{71}$$

where we used in the second line the chaining rule for conditional expectation to show that $\hat{\mathbb{E}}_k(u_k(X_k)) = \hat{\mathbb{E}}_k(\varphi_{k+1}(X_k, X_{k+1}, X_{k+1}))$ (the σ-field $\sigma(\widehat{X}_k) \subset \mathcal{F}_k$) and the contraction property of $\hat{\mathbb{E}}_k$.

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

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

Let $\nu \in (0, \bar{\nu}_n)$, so that for every $k = 0, \ldots, n$, $2(1+\nu) \leq 2 + \nu_k$. Hölder's inequality with conjugate exponents $p_\nu = 1 + \nu$ and $q_\nu = 1 + \frac{1}{\nu}$ gives

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

Lest us deal now with the second term on the right hand side of (71) and set for convenience

$$\Delta_k := \varphi_{k+1}(X_k, X_{k+1}, X_{k+1}) - \hat{u}_{k+1}(\hat{X}_{k+1}) g_{k+1}(\hat{X}_k, \hat{X}_{k+1}).$$

By the triangular inequality and the boundedness of g_{k+1}, we get

$$\begin{aligned}|\Delta_k| &\le \big|(u_{k+1}(X_{k+1}) - \hat{u}_{k+1}(\hat{X}_{k+1}))g_{k+1}(X_k, X_{k+1})\big| \\ &\quad + \big|\hat{u}_{k+1}(\hat{X}_{k+1}))(g_{k+1}(X_k, X_{k+1}) - g_{k+1}(\hat{X}_k, \hat{X}_{k+1}))\big| \\ &\le K_g\big|u_{k+1}(X_{k+1}) - \hat{u}_{k+1}(\hat{X}_{k+1})\big| + \|\hat{u}_{k+1}\|_\infty \big|g_{k+1}(X_k, X_{k+1}) - g_{k+1}(\hat{X}_k, \hat{X}_{k+1})\big|.\end{aligned}$$

so that

$$\|\Delta_k\|_2^2 \le 2K_g^2 \|u_{k+1}(X_{k+1}) - \hat{u}_{k+1}\hat{X}_{k+1}\|_2^2 + 2\|\hat{u}_{k+1}\|_\infty^2 \|g_{k+1}(X_k, X_{k+1}) - g_{k+1}(\hat{X}_k, \hat{X}_{k+1})\|_2^2.$$

It follows from $(\mathscr{H}_{\rm Liploc})$, Schwarz's and Minkowski inequalities that

$$\begin{aligned}\|g_{k+1}(X_k, X_{k+1}) &- g_{k+1}(\hat{X}_k, \hat{X}_{k+1})\|_2^2 \\ &\le [g^2_{k+1}]^2_{\rm Liploc}\mathbb{E}\Big[\big(1+\theta(X_k)+\theta(\hat{X}_k)+\theta(X_{k+1})+\theta(\hat{X}_{k+1})\big)^2 |X_{k+1} - \hat{X}_{k+1}|^2\Big] \\ &\quad + [g^1_{k+1}]^2_{\rm Liploc}\mathbb{E}\Big[\big(1+\theta(X_k)+\theta(\hat{X}_k)+\theta(X_{k+1})+\theta(\hat{X}_{k+1})\big)^2 |X_k - \hat{X}_k|^2\Big]. \\ &\le \big(M_{k,\nu}\big)^2\Big([g^2_{k+1}]^2_{\rm Liploc}\|X_{k+1} - \hat{X}_{k+1}\|^2_{2(1+\nu)} + [g^1_{k+1}]^2_{\rm Liploc}\|X_k - \hat{X}_k\|^2_{2(1+\nu)}\Big)\end{aligned}$$

where $M_{k,\nu} := 1 + \|\theta(X_k)\|_{2q_\nu} + \|\theta(\hat{X}_k)\|_{2q_\nu} + \|\theta(X_{k+1})\|_{2q_\nu} + \|\theta(\hat{X}_{k+1})\|_{2q_\nu}$.

Plugging these bounds in (71), we finally get that, for every $k = 0, \ldots, n-1$,

$$\|u_k(X_k) - \hat{u}_k(\hat{X}_k)\|_2^2 \le \widetilde{K}\|u_{k+1}(X_{k+1}) - \hat{u}_{k+1}(\hat{X}_{k+1})\|_2^2 + \alpha_k\|X_k - \hat{X}_k\|^2_{2(1+\nu)} + \beta_{k+1}\|X_{k+1} - \hat{X}_{k+1}\|^2_{2(1+\nu)}$$

where $\widetilde{K} = 2(K_g)^2$

$$\alpha_k := (M_{k,\nu})^2\big([u_k]^2_{\rm Liploc} + 2\|\hat{u}_{k+1}\|^2_\infty [g^1_{k+1}]^2_{\rm Liploc}\big), \quad 0 \le k \le n$$

and

$$\beta_k := 2\left(M_{k,\nu}\|\hat{u}_k\|_\infty [g^2_k]_{\rm Liploc}\right)^2, \quad 1 \le k \le n,$$

(we set $u_{n+1} = 0$ by convention so that $\alpha_n := \big([f]_{\rm Liploc} M_{n,\nu})^2\big)$. It follows by induction that,

$$\|u_k(X_k) - \hat{u}_k(\hat{X}_k)\|_2^2 \le \frac{1}{\widetilde{K}^k}\sum_{\ell=k}^n C_{\ell,n}(f,y)\|X_\ell - \hat{X}_\ell\|^2_{2(1+\nu)},\ k = 0, \ldots, n,$$

where, using the upper-bound for $[u_\ell]_{\rm Liploc}$ given by (64) (and the definition of $\kappa_{g,X}$ that follows),

$$\begin{aligned}C_{\ell,n}(f,y) &:= \widetilde{K}^{\ell-1}(\alpha_\ell\widetilde{K} + \beta_\ell),\ \ell = 0, \ldots, n \\ &= 2^\ell (M_{\ell,\nu})^2\left((K_g)^{2\ell}[u_\ell]^2_{\rm Liploc} + (K_g)^{2(n-1)}\|f\|^2_\infty(2[g^1_{\ell+1}]^2_{\rm Liploc} + [g^2_\ell]^2_{\rm Liploc})\right) \\ &\le 2^{\ell+1}(M_{\ell,\nu}(K_g)^n)^2\left[\left(\kappa^2_{g,X}\left(\frac{[P]^{n-\ell}_{\rm Liploc} - 1}{[P]_{\rm Liploc} - 1}\right)^2 + \frac{2[g^1]^2_{\rm Liploc} + [g^2]^2_{\rm Liploc}}{2(K_g)^2}\right)\|f\|^2_\infty \right. \\ &\quad \left. + [P]^{2(n-\ell)}_{\rm Liploc}[f]^2_{\rm Liploc}\right]\end{aligned}$$

(we also used the elementary inequality $ab \le \frac{1}{2}(a^2 + b^2)$, $a, b \ge 0$ in the third line). Finally

$$\begin{aligned}|\pi_{y,n}f - \hat{\pi}_{y,n}f|^2 &= |\mathbb{E}\,u_0(X_0) - \mathbb{E}\,\hat{u}_0(\hat{X}_0)|^2 \\ &\le \|u_0(X_0) - \hat{u}_0(\hat{X}_0)\|_2^2 \\ &\le \big((K_g)^n M_{n,\nu}\big)^2\big(R_{y,n}\|f\|^2_\infty + S_{y,n}[f]^2_{\rm Liploc}\big)\end{aligned}$$

where

$$R_{y,n} = \sum_{\ell=0}^{n} \left[2^{\ell+1} \left(\kappa_{g,X}^2 \left(\frac{[P]_{\text{Liploc}}^{n-\ell} - 1}{[P]_{\text{Liploc}} - 1} \right)^2 + \frac{2[g^1]_{\text{Liploc}}^2 + [g^2]_{\text{Liploc}}^2}{2(K_g)^2} \right) \right] \|X_\ell - \hat{X}_\ell\|_{2(1+\nu)}^2$$

and

$$S_{y,n} = \sum_{\ell=0}^{n} 2^{\ell+1} [P]_{\text{Liploc}}^{2(n-\ell)} \|X_\ell - \hat{X}_\ell\|_{2(1+\nu)}^2.$$

We conclude by Lemma 6.1. □

The previous theorem emphasizes the usefulness of the distortion mismatch result: it allows to switch from Lipschitz continuous assumptions on the functions g_k into *local Lipschitz* assumptions.

Remark 6.2. Note that, if we consider Assumption $(\mathscr{H}_0)$ instead of Assumption $(\mathscr{H}_{\text{Liploc}})$ in Theorem 6.3 *we still improve the upper bound established in Theorem 3.1 of [47] since* this amounts to setting $\theta \equiv 0$ and replace everywhere the "$[.]_{\text{Liploc}}$" coefficients by $[.]_{\text{Lip}}$. Then, like for BSDEs, the squared global error appear as the (weighted) cumulated sum of the squared quantization errors.

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] P. Briand and C. Labart. Simulation of BSDEs by Wiener chaos expansion. *Ann. Appl. Probab.*, 24(3):1129–1171, 2014.

[13] A.-L. Bronstein, G. Pagès, and J. Portès. Multi-asset American options and parallel quantization. *Methodol. Comput. Appl. Probab.*, 15(3):547–561, 2013.

[14] Anne Laure Bronstein, Gilles Pagès, and Benedikt Wilbertz. How to speed up the quantization tree algorithm with an application to swing options. *Quant. Finance*, 10(9):995–1007, 2010.

[15] 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.

[16] J. F. Chassagneux and D. Crisan. Runge kutta schemes for BSDEs. *Annals of Applied Probability*, 24(2):2679–720, 2014.

[17] J. F. Chassagneux and A. Richou. Numerical simulation of quadratic BSDEs. *Ann. Appl. Probab.*, 26(1):262–304, 2016.

[18] E.V. Chernaya. On the optimization of weighted cubature formulae on certain classes of continuous functions. *East J. Aprox.*, 1:47–60, 1995.

[19] 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.

[20] P. Del Moral and P. Jacod, J.and Protter. The Monte-Carlo method for filtering with discrete-time observations. *Probab. Theory Related Fields*, 120(3):346–368, 2001.

[21] F. Delarue and S. Menozzi. A Foward-Backward stochastic algorithm for quasi-linear PDEs. *The Annals of Applied Probability*, 16(1):140–184, 2006.

[22] F. Delarue and S. Menozzi. An interpolated stochastic algorithm for quasi-linear PDEs. *Math. Comp.*, 77(26):125–158, 2008.

[23] J. Diehl and P. Friz. Backward stochastic differential equations with rough drivers. *The Annals of Probability*, 40(4):1715–1758, 2012.

[24] J. Jr. Douglas, J. Ma, and P. Protter. Numerical methods for forward-backward stochastic differential equations. *Ann. Appl. Probab.*, 6:940–968, 1996.

[25] N. El Karoui, S. Peng, and M. C. Quenez. Backward stochastic differential equations in finance. *Mathematical Finance*, 7(1):1–71, 1997.

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

[27] E. Gobet and C. Labart. Error expansion for the discretization of backward stochastic differential equations. *Stochastic Processes and their Applications*, 117:803–829, 2007.

[28] 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.

[29] E. Gobet and S. Pagliarani. Analytical approximation of BSDEs with nonsmooth driver. *SIAM Journal on Financial Mathematics*, 6(1):919–958, 2015.

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

[31] S. Graf, H. Luschgy, and G. Pagès. Distortion mismatch in the quantization of probability measures. *ESAIM Probab. Stat.*, 12:127–153, 2008.

[32] S. Hamadène. Equations différentielles stochastiques rétrogrades: le cas localement lipschitz. *Ann. Inst. Henri Poincaré*, 32(5):645–659, 1996.

[33] P. Henry-Labordère, X. Tan, and N. Touzi. A numerical algorithm for a class of BSDEs via the branching process. *Stochastic Process. Appl.*, 124(2):1112–1140, 2014.

[34] 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.

[35] C. Illand. *Contrôle stochastique par quantification et application à la finance*. PhD thesis, Université Pierre et Marie Curie, 2012.

[36] I. Kharroubi and R. Elie. Adding constraints to BSDEs with jumps: an alternative to multidimensional reflections. *ESAIM, Probability and Statistics*, 18:233–250, 2014.

[37] 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.

[38] M. Kobylanski. Backward stochastic differential equations and partial differential equations with quadratic growth. *Annals of Probability*, 28:558–602, 2000.

[39] J.-P. Lepeltier and J. San Martin. On the existence of BSDE with superlinear quadratic coefficient. *Stochastics and Stochastic Report*, 63:227–240, 1998.

[40] 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.

[41] H. Luschgy and G. Pagès. *Marginal and Functional Quantization of Stochastic Processes*. book, in progress.

[42] 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.

[43] J. Ma and J. Yong. *Forward-Backward Stochastic Differential Equations and their Applications*. Lect. Notes in Math. 1702. Springer., 2007.

[44] H. Meter Soner, N. Touzi, and J. Zhang. Wellposedness of second order backward SDEs. *Probab. Theory Relat. Fields*, 153:149–190, 2012.

[45] G. Pagès. A space vector quantization method for numerical integration. *J. Computational and Applied Mathematics*, 89:1–38, 1998.

[46] G. Pagès. Introduction to optimal vector quantization and its applications to numerics. *ESAIM Proc. & Survey*, 48:29–79, 2015.

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

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

[49] 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.

[50] G. Pagès and J. Yu. Pointwise convergence of the Lloyd algorithm in higher dimension. To appear in *SIAM J. on Control and Optimization,* Preprint LPMA *1604*, 2013.

[51] E. Pardoux. BSDEs, weak convergence and homogenization of semilinear PDEs. *Nonlinear Analysis, Differential Equations and Control, NATO Science Series*, 528:503–549, 1999.

[52] E. Pardoux and S. G. Peng. Adapted solutions of backward stochactic differential equation. *Systems and Control Letters*, 14:55–61, 1990.

[53] E. Pardoux and S. G. Peng. Backward stochastic differential equations and quasilinear parabolic partial differential equations. *Lecture Notes in CIS*, 176:200–217, 1992.

[54] H. Pham, W. Runggaldier, and A. Sellami. Approximation by quantization of the filter process and applications to optimal stopping problems under partial observation. *Monte Carlo Methods and Applications*, 11:893–932, 2005.

[55] 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.

[56] 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.

[57] J. Zhang. A numerical scheme for bsdes. *Ann. Appl. Probab.*, 14(1):459–488, 2004.

[58] W. Zhao, Y. Fu, and T. Zhou. New kind of high-order multi-step schemes for forward backward stochastic differential equations. *SIAM J. Sci. Comput.*, 36(4):A1731–A1751, 2014.

A Proof of Theorem 3.1

STEP 1. Temporarily set for convenience $\bar{s} = t_k$ for $s \in [t_k, t_{k+1})$. Applying Ito's formula we have

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

Then, using assumption (28) we have

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

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

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

After choosing α and θ such that $\theta C(f) - \alpha < 0$, we take the expectation in both sizes of the previous inequality and use the fact that $\mathbb{E}|\mathbb{E}_{\underline{s}}(\tilde{Y}_{\bar{s}})|^2 \le \mathbb{E}|\tilde{Y}_{\bar{s}}|^2$ (owing to conditional Jensen inequality) to get

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

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

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

On the other hand, we have

$$
\tilde{\zeta}_{t_\ell} = \frac{1}{\Delta_n}\mathbb{E}_\ell\int_{t_\ell}^{t_{\ell+1}}\tilde{Z}_s ds, \quad \text{so that by Jensen's inequality,} \quad |\tilde{\zeta}_{t_\ell}|^2 \le \frac{1}{\Delta_n}\mathbb{E}_\ell\int_{t_\ell}^{t_{\ell+1}}|\tilde{Z}_s|^2 ds. \tag{72}
$$

It follows that

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

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

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

Now, let us choose θ 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 $e^{\alpha\Delta_n} \to 1$ as n goes to infinity we may consequently choose $\theta \in \big(C(f)(e^{\alpha\Delta_n} \vee T), \frac{\alpha}{C(f)}\big)$, for every $n \geq n_0 \in \mathbb{N}$. Setting

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

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

$$
e^{\alpha t_k}\mathbb{E}|\tilde{Y}_{t_k}|^2 + \big(1 - C^{(1,3)}\big)\int_{t_k}^{T} e^{\alpha s}\mathbb{E}|\tilde{Z}_s|^2 ds \leq C^{(1,1)} + \Delta C^{(1,2)}\sum_{\ell=k+1}^{n} e^{\alpha t_\ell}\mathbb{E}|\tilde{Y}_{t_\ell}|^2. \tag{73}
$$

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

$$
e^{\alpha t_k}\mathbb{E}|\tilde{Y}_{t_k}|^2 \leq C^{(1,1)} + \Delta_n C^{(1,2)}\sum_{\ell=k+1}^{n} e^{\alpha t_\ell}\mathbb{E}|\tilde{Y}_{t_\ell}|^2, \quad \forall\, k \in \{0,\ldots,n-1\}. \tag{74}
$$

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

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

Now, setting $k = 0$ in (73) 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 (72)), 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

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

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

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

Then, using the assumptions (28) yield

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

Now, thanks to the previous step we know that

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

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

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

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

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

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

$$\begin{aligned}
e^{\alpha t}|Y_t - \tilde{Y}_t|^2 &\leq [f]_{\mathrm{Lip}}\int_t^T e^{\alpha s}\Big(\theta|Y_s-\tilde{Y}_s|^2 + \frac{1}{\theta}\big(\Delta_n^2 + |X_s-\bar{X}_{\underline{s}}|^2 + |Y_s - \mathbb{E}_{\underline{s}}(\tilde{Y}_{\bar{s}})|^2 + |Z_s - \tilde{\zeta}_{\underline{s}}|^2\big)\Big)ds \\
&\quad -\alpha\int_t^T e^{\alpha s}|Y_s-\tilde{Y}_s|^2 ds - \int_t^T e^{\alpha s}|Z_s - \tilde{Z}_s|^2 ds + 2\int_t^T e^{\alpha s}(Z_s-\tilde{Z}_s)dW_s. \qquad (75)
\end{aligned}$$

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

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

for some positive real constants $C^{(3,1)}, C^{(3,2)}$ and $C^{(3,3)}$. Then, taking the expectation in (75) 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

$$\begin{aligned}
&\mathbb{E}\Big(e^{\alpha t}|Y_t-\tilde{Y}_t|^2 + \int_t^T e^{\alpha s}|Z_s-\tilde{Z}_s|^2 ds\Big) \leq \Big(-\alpha + [f]_{\mathrm{Lip}}\big(\theta + \frac{2}{\theta}\big)\Big)\int_t^T e^{\alpha s}\mathbb{E}|Y_s - \tilde{Y}_s|^2 ds \\
&+ \frac{[f]_{\mathrm{Lip}}}{\theta}\left(\frac{e^{\alpha T}-e^{\alpha t}}{\alpha}\big(\Delta_n^2 + C^{(3,3)}\Delta_n\big) + 2\int_t^T e^{\alpha s}\mathbb{E}|\tilde{Y}_s - \mathbb{E}_{\underline{s}}(\tilde{Y}_{\bar{s}})|^2 ds + \int_t^T e^{\alpha s}\mathbb{E}|Z_s - \tilde{\zeta}_{\underline{s}}|^2 ds\right). \ (76)
\end{aligned}$$

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

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

and

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

where $a \underline{\in} \mathcal{F}_{t_k}$ 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

$$\begin{aligned}
\int_{t_k}^{t_{k+1}} e^{\alpha s} \mathbb{E}_k |Z_s - \tilde{\zeta}_{\underline{s}}|^2 ds \;&\leq\; 2 \int_{t_k}^{t_{k+1}} e^{\alpha s} \mathbb{E}_k |Z_s - \zeta_{\underline{s}}|^2 ds + 2 \int_{t_k}^{t_{k+1}} e^{\alpha s} \mathbb{E}_k |\zeta_{\underline{s}} - \tilde{\zeta}_{\underline{s}}|^2 ds \qquad (77)\\
&\leq\; 2 \int_{t_k}^{t_{k+1}} e^{\alpha s} \mathbb{E}_k |Z_s - Z_{\underline{s}}|^2 ds + \frac{2}{\Delta_n^2} \int_{t_k}^{t_{k+1}} e^{\alpha s} \mathbb{E}_k \Big| \mathbb{E}_k \int_{t_k}^{t_{k+1}} (Z_u - \tilde{Z}_u) du \Big|^2 ds.
\end{aligned}$$

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

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

Consequently, taking the expectation in (77) leads to

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

Coming back to Inequality (76) and setting $\alpha = \alpha(\theta) = [f]_{\text{Lip}}\big(\theta + \frac{2}{\theta}\big)$ yields

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

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

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

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

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

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

$$2\frac{[f]_{\rm Lip}}{\theta}e^{\alpha\Delta_n} \le \frac{1}{2}\exp\Big([f]_{\rm Lip}\big([f]_{\rm Lip}^2+\frac{1}{2}\big)\Delta_n\Big)$$

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

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

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

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

Now, setting $k=0$ yields likewise

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

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:

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

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

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

(where $\sigma^\star$ stands for the transpose of σ). We know (see e.g. [21]) 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 [21]), 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

$$\forall\, s,t,\in[0,T], \forall\, x\in\mathbb{R}^d, \quad |\nabla_x u(t,x)-\nabla_x u(s,x)| \le C^{(4,1)}|t-s|^{1/2}. \tag{78}$$

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

$$\mathbb{E}|X_t-X_s|^2 \le C^{(4,2)}|t-s|. \tag{79}$$

Then, using (78), the assumptions on σ 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 $\partial^2_{x^2}\phi$ exist and are uniformly bounded (in particular $\nabla_x u$ is Lipschitz continuous in x), we have for every $t\in[0,T]$,

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

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

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