The Annals of Applied Probability
2011, Vol. 21, No. 5, 1933–1964
DOI: 10.1214/10-AAP744

NUMERICAL SIMULATION OF BSDES WITH DRIVERS OF QUADRATIC GROWTH

BY ADRIEN RICHOU

Université Rennes 1

This article deals with the numerical resolution of Markovian backward stochastic differential equations (BSDEs) with drivers of quadratic growth with respect to z and bounded terminal conditions. We first show some bound estimates on the process Z and we specify the Zhang's path regularity theorem. Then we give a new time discretization scheme with a nonuniform time net for such BSDEs and we obtain an explicit convergence rate for this scheme.

1. Introduction. Since the early nineties, there has been an increasing interest for backward stochastic differential equations (BSDEs). These equations have a wide range of applications in stochastic control, in finance or in partial differential equation theory. A particular class of BSDE has been studied for a few years: BSDEs with drivers of quadratic growth with respect to the variable z. This class arises, for example, in the context of utility optimization problems with exponential utility functions or alternatively in questions related to risk minimization for the entropic risk measure (see, e.g., [13]). Many papers deal with existence and uniqueness of solution for such BSDEs; we refer the reader to [17, 18] when the terminal condition is bounded and [3, 4, 9] for the unbounded case. Our concern is rather related to the simulation of BSDEs and more precisely time discretization of BSDEs coupled with a forward stochastic differential equation (SDE). Actually, the design of efficient algorithms which are able to solve BSDEs in any reasonable dimension has been intensely studied since the first work of Chevance [6] (see, e.g., [1, 11, 19]). But in all these works, the driver of the BSDE is a Lipschitz function with respect to z and this assumption plays a key role in their proofs. In a recent paper, Cheridito and Stadje [5] studied approximation of BSDEs by backward stochastic difference equations which are based on random walks instead of Brownian motions. They obtain a convergence result when the driver has a subquadratic growth with respect to z and they give an example where this approximation does not converge when the driver has a quadratic growth. To the best of our knowledge, the only work where the time approximation of a BSDE with a quadratic growth with respect to z is studied is the one of Imkeller and Reis [14]. Notice that, when the driver has a specific form (roughly speaking, the driver

Received February 2010; revised September 2010.
MSC2010 subject classifications. Primary 60H35; secondary 65C30, 60H10.
Key words and phrases. BSDEs, driver of quadratic growth, time discretization scheme.

is a sum of a quadratic term $z \mapsto C|z|^2$ and a function that has a linear growth with respect to z), it is possible to get around the problem by using an exponential transformation method (see [15]) or by using results on fully coupled forward–backward differential equations (see [7]).

To explain the ideas of this paper, let us introduce (X, Y, Z) the solution to the forward–backward system

$$X_t = x + \int_0^t b(s, X_s)\,ds + \int_0^t \sigma(s)\,dW_s,$$
$$Y_t = g(X_T) + \int_t^T f(s, X_s, Y_s, Z_s)\,ds - \int_t^T Z_s\,dW_s,$$

where g is bounded, f is locally Lipschitz and has a quadratic growth with respect to z. A well-known result is that when g is a Lipschitz function with Lipschitz constant K_g, then the process Z is bounded by $C(K_g + 1)$ (see Theorem 3.1). So, in this case, the driver of the BSDE is a Lipschitz function with respect to z and we are able to use standard results about discretization of BSDEs. Because of the above observation, this paper will focus on the case that the terminal function g is not Lipschitz. To obtain our main results, we will assume that g is an α-Hölder function but it is also possible to adapt our methods when g is not α-Hölder; for example, Remark 4.13 deals with the case of an indicator function of a smooth domain. Let us notice that the time approximation of BSDEs with an irregular terminal function has already been studied by Gobet and Makhlouf [12] when the generator is a Lipschitz function with respect to z.

In light of previous observation, a simple idea is to do an approximation of (Y, Z) by the solution (Y^N, Z^N) to the BSDE

$$Y_t^N = g_N(X_T) + \int_t^T f(s, X_s, Y_s^N, Z_s^N)\,ds - \int_t^T Z_s^N\,dW_s,$$

where g_N is a Lipschitz approximation of g. Thanks to bounded mean oscillation martingale (BMO martingale in the sequel) tools, we have an error estimate for this approximation (see, e.g., [2, 14] or Proposition 4.2). For example, if g is α-Hölder, we are able to obtain the error bound $C K_{g_N}^{-\alpha/(1-\alpha)}$ (see Proposition 4.11). Moreover, we can have an error estimate for the time discretization of the approximated BSDE thanks to any numerical scheme for BSDEs with Lipschitz driver. But this error estimate depends on K_{g_N}; roughly speaking, this error is $Ce^{CK_{g_N}^2} n^{-1}$ with n the number of discretization times. The exponential term results from the use of Gronwall's inequality. Finally, when g is α-Hölder and $K_{g_N} = N$, the global error bound is

$$C\left(\frac{1}{N^{\alpha/(1-\alpha)}} + \frac{e^{CN^2}}{n}\right). \tag{1}$$

So, when N increases, n^{-1} will have to become small very quickly and the speed of convergence turns out to be bad; if we take $N = (\frac{C}{\varepsilon}\log n)^{1/2}$ with $0 < \varepsilon < 1$, then

the global error bound becomes $C_\varepsilon(\log n)^{-\alpha/(2(1-\alpha))}$. The same drawback appears in the work of Imkeller and Reis [14]. Indeed, their idea is to do an approximation of (Y, Z) by the solution (Y^N, Z^N) to the truncated BSDE

$$Y_t^N = g(X_T) + \int_t^T f(s, X_s, Y_s^N, h_N(Z_s^N))\,ds - \int_t^T Z_s^N\,dW_s,$$

where $h_N : \mathbb{R}^{1\times d} \to \mathbb{R}^{1\times d}$ is a smooth modification of the projection on the open Euclidean ball of radius N about 0. Thanks to several statements concerning the path regularity and stochastic smoothness of the solution processes, the authors show that for any $\beta \geq 1$, the approximation error is lower than $C_\beta N^{-\beta}$. So they obtain the global error bound

$$C_\beta\left(\frac{1}{N^\beta} + \frac{e^{CN^2}}{n}\right) \tag{2}$$

and, consequently, the speed of convergence also turns out to be bad; if we take $N = (\frac{C}{\varepsilon}\log n)^{1/2}$ with $0 < \varepsilon < 1$, then the global error bound becomes $C_{\beta,\varepsilon}(\log n)^{-\beta/2}$.

Another idea is to use an estimate of Z that does not depend on K_g. So we extend a result of [8] which shows

$$|Z_t| \leq M_1 + \frac{M_2}{(T-t)^{1/2}}, \qquad 0 \leq t < T. \tag{3}$$

Let us notice that this type of estimation is well known in the case of drivers with linear growth as a consequence of the Bismut–Elworthy formula (see, e.g., [10]). But in our case, we do not need to suppose that σ is invertible. Then, thanks to this estimation, we know that when $t < T$, $f(t, \cdot, \cdot, \cdot)$ is a Lipschitz function with respect to z and the Lipschitz constant depends on t. So we are able to modify the classical uniform time net to obtain a convergence speed for a modified time discretization scheme for our BSDE; the idea is to put more discretization points near the final time T than near 0. Roughly speaking, our discretization grid is equal to

$$t_k := T\left(1 - \left(\frac{\varepsilon}{T}\right)^{k/n}\right), \qquad 0 \leq k \leq n,$$

with ε a parameter. But due to technical reasons we need to apply this modified time discretization scheme to the approximated BSDE

$$Y_t^{N,\varepsilon} = g_N(X_T) + \int_t^T f^\varepsilon(s, X_s, Y_s^{N,\varepsilon}, Z_s^{N,\varepsilon})\,ds - \int_t^T Z_s^{N,\varepsilon}\,dW_s$$

with

$$f^\varepsilon(s, x, y, z) := \mathbb{1}_{s\leq T-\varepsilon} f(s, x, y, z) + \mathbb{1}_{s>T-\varepsilon} f(s, x, y, 0).$$

Thanks to the estimate (3), we obtain a speed convergence for the time discretization scheme of this approximated BSDE (see Theorem 4.9). Moreover, BMO tools

give us again an estimate of the approximation error (see Proposition 4.2). Finally, if we suppose that g is α-Hölder, we prove that we can choose properly N and ε to obtain the global error estimate $Cn^{-2\alpha/((2-\alpha)(2+K)-2+2\alpha)}$ (see Theorem 4.14) where $K > 0$ depends on constant M_2 defined in equation (3) and constants related to f. Let us notice that such a speed of convergence where constants related to f, g, b and σ appear in the power of n is unusual. Even if we have an error far better than (1) or (2), this result is not very interesting in practice because the speed of convergence strongly depends on K. But, when b is bounded, we prove that we can take M_2 as small as we want in (3). Finally, we obtain a global error estimate lower than $C_\eta n^{-(\alpha-\eta)}$ for all $\eta > 0$ (see Theorem 4.17).

To conclude, it could be interesting to do some comparisons between our work and the article of Gobet and Makhlouf [12]. We already explain that this paper studies the time approximation of Lipschitz BSDEs with irregular terminal functions. These authors show that the error of approximation is lower than $C_\eta n^{-\alpha}$ when g is an α-Hölder function and the discretization grid is uniform. So, our better speed of convergence is very close to their result. Nevertheless, they also show that it is possible to obtain the classical speed of convergence, that is to say Cn^{-1}, when we use the nonuniform grid given by

$$t_k := T - T\left(1 - \frac{k}{n}\right)^{1/\beta}, \qquad 0 \le k \le n,$$

with $\beta < \alpha$. It is interesting to notice that we both use nonuniform time discretization points but their grid is different than our grid; the accumulation speed of discretization points near the terminal time T is not the same; it is faster in our case.

The paper is organized as follows. In the introductory Section 2 we recall some of the well-known results concerning SDEs and BSDEs. In Section 3 we establish some estimates concerning the process Z; we show a first uniform bound for Z, then a time dependent bound and finally we specify the classical path regularity theorem. In Section 4 we define a modified time discretization scheme for BSDEs with a nonuniform time net and we obtain an explicit error bound.

2. Preliminaries.

2.1. *Notation*. Throughout this paper, $(W_t)_{t\ge 0}$ will denote a d-dimensional Brownian motion, defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. For $t \ge 0$, let $\mathcal{F}_t$ denote the σ-algebra $\sigma(W_s; 0 \le s \le t)$, augmented with the $\mathbb{P}$-null sets of $\mathcal{F}$. The Euclidean norm on $\mathbb{R}^d$ will be denoted by $|\cdot|$. The operator norm induced by $|\cdot|$ on the space of linear operator is also denoted by $|\cdot|$. For $p \ge 2$, $m \in \mathbb{N}$, we denote further:

(1) $\mathcal{S}^p(\mathbb{R}^m)$ or $\mathcal{S}^p$ when no confusion is possible, the space of all adapted processes $(Y_t)_{t\in[0,T]}$ with values in $\mathbb{R}^m$ normed by

$$\|Y\|_{\mathcal{S}^p} = \mathbb{E}\Big[\Big(\sup_{t\in[0,T]} |Y_t|\Big)^p\Big]^{1/p};$$

$\mathcal{S}^\infty(\mathbb{R}^m)$ or $\mathcal{S}^\infty$, the space of bounded measurable processes;

(2) $\mathcal{M}^p(\mathbb{R}^m)$ or $\mathcal{M}^p$, the space of all progressively measurable processes $(Z_t)_{t\in[0,T]}$ with values in $\mathbb{R}^m$ normed by

$$\|Z\|_{\mathcal{M}^p} = \mathbb{E}\bigg[\bigg(\int_0^T |Z_s|^2\,ds\bigg)^{p/2}\bigg]^{1/p}.$$

In the following we keep the same notation C for all finite, nonnegative constants that appear in our computations; they may depend on known parameters deriving from assumptions and on T but not on any of the approximation and discretization parameters. In the same spirit, we keep the same notation η for all finite, positive constants that we can take as small as we want independently of the approximation and discretization parameters.

2.2. *Some results on BMO martingales.* In our work, the space of BMO martingales play a key role for the a priori estimates needed in our analysis of BSDEs. We refer the reader to [16] for the theory of BMO martingales and we just recall the properties that we will use in the sequel. Let $\Phi_t = \int_0^t \phi_s\,dW_s$, $t\in[0,T]$, be a real square integrable martingale with respect to the Brownian filtration. Then Φ is a BMO martingale if

$$\|\Phi\|_{\mathrm{BMO}} = \sup_{\tau\in[0,T]} \mathbb{E}[\langle\Phi\rangle_T - \langle\Phi\rangle_\tau|\mathcal{F}_\tau]^{1/2} = \sup_{\tau\in[0,T]} \mathbb{E}\bigg[\int_\tau^T \phi_s^2\,ds\big|\mathcal{F}_\tau\bigg]^{1/2} < +\infty,$$

where the supremum is taken over all stopping times in $[0,T]$; $\langle\Phi\rangle$ denotes the quadratic variation of Φ. In our case, the very important feature of BMO martingales is the following lemma.

LEMMA 2.1. *Let Φ be a BMO martingale. Then we have*:

(1) *The stochastic exponential*

$$\mathcal{E}(\Phi)_t = \mathcal{E}_t = \exp\bigg(\int_0^t \phi_s\,dW_s - \frac{1}{2}\int_0^t |\phi_s|^2\,ds\bigg), \qquad 0\le t\le T,$$

is a uniformly integrable martingale.

(2) *Thanks to the reverse Hölder inequality, there exists $p>1$ such that $\mathcal{E}_T \in L^p$. The maximal p with this property can be expressed in terms of the BMO norm of Φ.*

(3) $\forall n\in\mathbb{N}^*$, $\mathbb{E}[(\int_0^T |\phi_s|^2\,ds)^n] \le n!\|\Phi\|_{\mathrm{BMO}}^{2n}$.

2.3. *The backward–forward system.* Given functions b,σ,g and f, for $x\in\mathbb{R}^d$ we will deal with the solution (X,Y,Z) to the following system of (decoupled) backward–forward stochastic differential equations: for $t\in[0,T]$,

$$X_t = x + \int_0^t b(s,X_s)\,ds + \int_0^t \sigma(s)\,dW_s, \tag{4}$$

$$Y_t = g(X_T) + \int_t^T f(s,X_s,Y_s,Z_s)\,ds - \int_t^T Z_s\,dW_s. \tag{5}$$

For the functions that appear in the above system of equations we give some general assumptions.

(HX0). $b:[0,T]\times\mathbb{R}^d\to\mathbb{R}^d$, $\sigma:[0,T]\to\mathbb{R}^{d\times d}$ are measurable functions. There exist four positive constants M_b, K_b, M_σ and K_σ such that $\forall t,t'\in[0,T]$, $\forall x,x'\in\mathbb{R}^d$,

$$\begin{aligned}|b(t,x)|&\le M_b(1+|x|),\\ |b(t,x)-b(t',x')|&\le K_b(|x-x'|+|t-t'|^{1/2}),\\ |\sigma(t)|&\le M_\sigma,\\ |\sigma(t)-\sigma(t')|&\le K_\sigma|t-t'|.\end{aligned}$$

(HY0). $f:[0,T]\times\mathbb{R}^d\times\mathbb{R}\times\mathbb{R}^{1\times d}\to\mathbb{R}$, $g:\mathbb{R}^d\to\mathbb{R}$ are measurable functions. There exist five positive constants M_f, $K_{f,x}$, $K_{f,y}$, $K_{f,z}$ and M_g such that $\forall t\in[0,T]$, $\forall x,x'\in\mathbb{R}^d$, $\forall y,y'\in\mathbb{R}$, $\forall z,z'\in\mathbb{R}^{1\times d}$,

$$\begin{aligned}|f(t,x,y,z)|&\le M_f(1+|y|+|z|^2),\\ |f(t,x,y,z)-f(t,x',y',z')|&\le K_{f,x}|x-x'|+K_{f,y}|y-y'|\\ &\quad+\big(K_{f,z}+L_{f,z}(|z|+|z'|)\big)|z-z'|,\\ |g(x)|&\le M_g.\end{aligned}$$

We next recall some results on BSDEs with quadratic growth. For their original version and their proof we refer to [2, 17] and [14].

THEOREM 2.2. *Under* (HX0), (HY0), *the system* (4)–(5) *has a unique solution* $(X,Y,Z)\in\mathcal{S}^2\times\mathcal{S}^\infty\times\mathcal{M}^2$. *The martingale* $Z*W$ *belongs to the space of BMO martingales and* $\|Z*W\|_{\mathrm{BMO}}$ *only depends on* T, M_g *and* M_f. *Moreover, there exists* $r>1$ *such that* $\mathcal{E}(Z*W)\in L^r$.

3. Some useful estimates of Z.

3.1. *A first bound for* Z.

THEOREM 3.1. *Suppose that* (HX0), (HY0) *hold and that* g *is Lipschitz with Lipschitz constant* K_g. *Then, there exists a version of* Z *such that*, $\forall t\in[0,T]$,

$$|Z_t|\le e^{(2K_b+K_{f,y})T}M_\sigma(K_g+TK_{f,x}).$$

PROOF. First, we suppose that b, g and f are differentiable with respect to x, y and z. Then (X,Y,Z) is differentiable with respect to x and $(\nabla X,\nabla Y,\nabla Z)$ is

the solution of

$$\nabla X_t = I_d + \int_0^t \nabla b(s, X_s)\nabla X_s\, ds, \tag{6}$$

$$\begin{aligned}\nabla Y_t = {} & \nabla g(X_T)\nabla X_T - \int_t^T \nabla Z_s\, dW_s \\ & + \int_t^T \nabla_x f(s, X_s, Y_s, Z_s)\nabla X_s + \nabla_y f(s, X_s, Y_s, Z_s)\nabla Y_s\, ds \\ & + \int_t^T \nabla_z f(s, X_s, Y_s, Z_s)\nabla Z_s\, ds,\end{aligned} \tag{7}$$

where $\nabla X_t = (\partial X_t^i/\partial x^j)_{1\le i,j\le d}$, $\nabla Y_t = {}^t(\partial Y_t/\partial x^j)_{1\le j\le d} \in \mathbb{R}^{1\times d}$, $\nabla Z_t = (\partial Z_t^i/\partial x^j)_{1\le i,j\le d}$ and $\int_t^T \nabla Z_s\, dW_s$ means

$$\sum_{1\le i\le d} \int_t^T (\nabla Z_s)^i\, dW_s^i$$

with $(\nabla Z)^i$ denoting the ith line of the $d\times d$ matrix process ∇Z. Thanks to usual transformations on the BSDE we obtain

$$\begin{aligned}& e^{\int_0^t \nabla_y f(s,X_s,Y_s,Z_s)\, ds}\nabla Y_t \\ & \quad = e^{\int_0^T \nabla_y f(s,X_s,Y_s,Z_s)\, ds}\nabla g(X_T)\nabla X_T \\ & \qquad - \int_t^T e^{\int_0^s \nabla_y f(u,X_u,Y_u,Z_u)\, du}\nabla Z_s\, d\tilde{W}_s \\ & \qquad + \int_t^T e^{\int_0^s \nabla_y f(u,X_u,Y_u,Z_u)\, du}\nabla_x f(s, X_s, Y_s, Z_s)\nabla X_s\, ds\end{aligned}$$

with $d\tilde{W}_s = dW_s - \nabla_z f(s, X_s, Y_s, Z_s)\, ds$. We have

$$\begin{aligned}& \left\| \int_0^\cdot \nabla_z f(s, X_s, Y_s, Z_s)\, dW_s \right\|_{\mathrm{BMO}}^2 \\ & \quad = \sup_{\tau\in[0,T]} \mathbb{E}\left[\int_\tau^T |\nabla_z f(s, X_s, Y_s, Z_s)|^2\, ds | \mathcal{F}_\tau \right] \\ & \quad \le C\left(1 + \sup_{\tau\in[0,T]} \mathbb{E}\left[\int_\tau^T |Z_s|^2\, ds | \mathcal{F}_\tau \right]\right) \\ & \quad = C(1 + \|Z * W\|_{\mathrm{BMO}}^2).\end{aligned}$$

Since $Z * W$ belongs to the space of BMO martingales,

$$\left\| \int_0^\cdot \nabla_z f(s, X_s, Y_s, Z_s)\, dW_s \right\|_{\mathrm{BMO}} < +\infty.$$

Lemma 2.1 gives us that $\mathcal{E}(\int_0^\cdot \nabla_z f(s, X_s, Y_s, Z_s)\, dW_s)_t$ is a uniformly integrable martingale so we are able to apply Girsanov's theorem: there exists a probability $\mathbb{Q}$ under which $(\tilde{W})_{t\in[0,T]}$ is a Brownian motion. Then,

$$\begin{aligned} &e^{\int_0^t \nabla_y f(s,X_s,Y_s,Z_s)\, ds} \nabla Y_t \\ &\quad = \mathbb{E}^{\mathbb{Q}}\Big[e^{\int_0^T \nabla_y f(s,X_s,Y_s,Z_s)\, ds} \nabla g(X_T) \nabla X_T \\ &\qquad + \int_t^T e^{\int_0^s \nabla_y f(u,X_u,Y_u,Z_u)\, du} \nabla_x f(s, X_s, Y_s, Z_s) \nabla X_s \, ds \big| \mathcal{F}_t \Big] \end{aligned}$$

and

$$|\nabla Y_t| \le e^{(K_b + K_{f,y})T} (K_g + T K_{f,x}), \tag{8}$$

because $|\nabla X_t| \le e^{K_b T}$. Moreover, thanks to the Malliavin calculus, it is classical to show that a version of $(Z_t)_{t\in[0,T]}$ is given by $(\nabla Y_t (\nabla X_t)^{-1} \sigma(t))_{t\in[0,T]}$. So we obtain

$$|Z_t| \le e^{K_b T} M_\sigma |\nabla Y_t| \le e^{(2K_b + K_{f,y})T} M_\sigma (K_g + T K_{f,x}) \qquad \text{a.s.},$$

because $|\nabla X_t^{-1}| \le e^{K_b T}$.

When b, g and f are not differentiable, we can also prove the result by a standard approximation and stability results for BSDEs with linear growth. □

REMARK 3.2. Thanks to Theorem 3.1, the generator f becomes a Lipschitz function with respect to z, so we are able to use standard results about time discretization of BSDEs. In this case, we obtain that the error of approximation is lower than Cn^{-1} with n the number of discretization times (see, e.g., [1, 11]). Let us notice that, in all studies about discretization of BSDEs, we do not care about the constant in the error bound; we only consider the asymptotic speed of convergence. But, with a practical point of view, the constant could play an important role, particularly for small n. In our case, the generator f may be viewed as Lipschitz in z with a Lipschitz constant $Ce^{(2K_b + K_{f,y})T}$. So, if we apply the standard result, the generic constant in the rate of convergence will be in the order of $Ce^{Ce^{2(2K_b + K_{f,y})T}}$. This is, of course, not desirable because it blows up when K_b, $K_{f,y}$ or T increase. We think that it could be interesting to see if we are able to observe such a phenomena with numerical simulation.

3.2. *A time dependent estimate of Z.* We will introduce two alternative assumptions.

(HX1). b is differentiable with respect to x and σ is differentiable with respect to t. There exists $\lambda \in \mathbb{R}^+$ such that $\forall \eta \in \mathbb{R}^d$

$$|{}^t\eta \sigma(s)[{}^t\sigma(s)^t \nabla b(s,x) - {}^t\sigma'(s)]\eta| \le \lambda |{}^t\eta \sigma(s)|^2. \tag{9}$$

(HX1′). σ is invertible and $\forall t \in [0, T]$, $|\sigma(t)^{-1}| \leq M_{\sigma^{-1}}$.

Example. Assumption (HX1) is verified when, $\forall s \in [0, T]$, $\nabla b(s, \cdot)$ commutes with $\sigma(s)$ and $\exists A : [0, T] \to \mathbb{R}^{d\times d}$ bounded such that $\sigma'(t) = \sigma(t)A(t)$.

THEOREM 3.3. *Suppose that* (HX0), (HY0) *hold and that* (HX1) *or* (HX1′) *holds. Moreover, suppose that g is lower* (*or upper*) *semi-continuous. Then there exists a version of Z and there exist two constants* $C, C' \in \mathbb{R}^+$ *that depend only in* T, M_g, M_f, $K_{f,x}$, $K_{f,y}$, $K_{f,z}$ *and* $L_{f,z}$ *such that*, $\forall t \in [0, T[$,

$$|Z_t| \leq C + C'(T - t)^{-1/2}.$$

PROOF. In a first time, we will suppose that (HX1) holds and that f, g are differentiable with respect to x, y and z. Then (Y, Z) is differentiable with respect to x and $(\nabla Y, \nabla Z)$ is the solution of the BSDE

$$\begin{aligned}\nabla Y_t = {} & \nabla g(X_T)\nabla X_T - \int_t^T \nabla Z_s \, dW_s \\ & + \int_t^T \nabla_x f(s, X_s, Y_s, Z_s)\nabla X_s + \nabla_y f(s, X_s, Y_s, Z_s)\nabla Y_s \, ds \\ & + \int_t^T \nabla_z f(s, X_s, Y_s, Z_s)\nabla Z_s \, ds.\end{aligned}$$

Thanks to usual transformations we obtain

$$\begin{aligned}& e^{\int_0^t \nabla_y f(s,X_s,Y_s,Z_s)\,ds}\nabla Y_t \\ & \quad + \int_0^t e^{\int_0^s \nabla_y f(u,X_u,Y_u,Z_u)\,du}\nabla_x f(s, X_s, Y_s, Z_s)\nabla X_s \, ds \\ & \quad = e^{\int_0^T \nabla_y f(s,X_s,Y_s,Z_s)\,ds}\nabla g(X_T)\nabla X_T \\ & \qquad + \int_0^T e^{\int_0^s \nabla_y f(u,X_u,Y_u,Z_u)\,du}\nabla_x f(s, X_s, Y_s, Z_s)\nabla X_s \, ds \\ & \qquad - \int_t^T e^{\int_0^s \nabla_y f(u,X_u,Y_u,Z_u)\,du}\nabla Z_s \, d\tilde{W}_s\end{aligned}$$

with $d\tilde{W}_s = dW_s - \nabla_z f(s, X_s, Y_s, Z_s)\, ds$. We can rewrite it as

$$F_t = F_T - \int_t^T e^{\int_0^s \nabla_y f(u,X_u,Y_u,Z_u)\,du}\nabla Z_s \, d\tilde{W}_s \tag{10}$$

with

$$\begin{aligned}F_t := {} & e^{\int_0^t \nabla_y f(s,X_s,Y_s,Z_s)\,ds}\nabla Y_t \\ & + \int_0^t e^{\int_0^s \nabla_y f(u,X_u,Y_u,Z_u)\,du}\nabla_x f(s, X_s, Y_s, Z_s)\nabla X_s \, ds.\end{aligned}$$

$Z * W$ belongs to the space of BMO martingales so we are able to apply Girsanov's theorem: there exists a probability $\mathbb{Q}$ under which $(\tilde{W})_{t\in[0,T]}$ is a Brownian motion. Thanks to the Malliavin calculus, it is possible to show that $(\nabla Y_t(\nabla X_t)^{-1}\sigma(t))_{t\in[0,T]}$ is a version of Z. Now we define

$$\begin{aligned}
\alpha_t &:= \int_0^t e^{\int_0^s \nabla_y f(u,X_u,Y_u,Z_u)\,du}\nabla_x f(s,X_s,Y_s,Z_s)\nabla X_s\,ds\,(\nabla X_t)^{-1}\sigma(t),\\
\tilde{Z}_t &:= F_t(\nabla X_t)^{-1}\sigma(t) = e^{\int_0^t \nabla_y f(s,X_s,Y_s,Z_s)\,ds}Z_t + \alpha_t \qquad \text{a.s.},\\
\tilde{F}_t &:= e^{\lambda t}F_t(\nabla X_t)^{-1}.
\end{aligned}$$

Since $d\nabla X_t = \nabla b(t,X_t)\nabla X_t\,dt$, then $d(\nabla X_t)^{-1} = -(\nabla X_t)^{-1}\nabla b(t,X_t)\,dt$ and thanks to Itô's formula,

$$d\tilde{Z}_t = dF_t(\nabla X_t)^{-1}\sigma(t) - F_t(\nabla X_t)^{-1}\nabla b(t,X_t)\sigma(t)\,dt + F_t(\nabla X_t)^{-1}\sigma'(t)\,dt$$

and

$$d(e^{\lambda t}\tilde{Z}_t) = \tilde{F}_t\big(\lambda Id - \nabla b(t,X_t)\big)\sigma(t)\,dt + \tilde{F}_t\sigma'(t)\,dt + e^{\lambda t}\,dF_t(\nabla X_t)^{-1}\sigma(t).$$

Finally,

$$\begin{aligned}
d|e^{\lambda t}\tilde{Z}_t|^2 = 2\big[\lambda|\tilde{F}_t\sigma(t)|^2 &- \tilde{F}_t\sigma(t)[{}^t\sigma(t){}^t\nabla b(t,X_t) - {}^t\sigma'(t)]{}^t\tilde{F}_t\big]\,dt\\
&+ d\langle M\rangle_t + dM_t^*
\end{aligned}$$

with $M_t := \int_0^t e^{\lambda s}\,dF_s(\nabla X_s)^{-1}\sigma(s)$ and M_t^* a $\mathbb{Q}$-martingale. Thanks to the assumption (HX1) we are able to conclude that $|e^{\lambda t}\tilde{Z}_t|^2$ is a $\mathbb{Q}$-submartingale. Hence,

$$\begin{aligned}
&\mathbb{E}^{\mathbb{Q}}\bigg[\int_t^T e^{2\lambda s}|\tilde{Z}_s|^2\,ds\big|\mathcal{F}_t\bigg]\\
&\quad\geq e^{2\lambda t}|\tilde{Z}_t|^2(T-t)\\
&\quad\geq e^{2\lambda t}\big|e^{\int_0^t \nabla_y f(s,X_s,Y_s,Z_s)\,ds}Z_t + \alpha_t\big|^2(T-t) \qquad \text{a.s.},
\end{aligned}$$

which implies

$$\begin{aligned}
|Z_t|^2(T-t) &= e^{-2\lambda t}e^{-2\int_0^t \nabla_y f(s,X_s,Y_s,Z_s)\,ds}e^{2\lambda t}\\
&\qquad\times\big|e^{\int_0^t \nabla_y f(s,X_s,Y_s,Z_s)\,ds}Z_t + \alpha_t - \alpha_t\big|^2(T-t)\\
&\leq C\big(e^{2\lambda t}\big|e^{\int_0^t \nabla_y f(s,X_s,Y_s,Z_s)\,ds}Z_t + \alpha_t\big|^2 + 1\big)(T-t)\\
&\leq C\bigg(\mathbb{E}^{\mathbb{Q}}\bigg[\int_t^T e^{2\lambda s}|\tilde{Z}_s|^2\,ds\big|\mathcal{F}_t\bigg] + (T-t)\bigg) \qquad \text{a.s.}
\end{aligned}$$

with C a constant that only depends on T, K_b, M_σ, $K_{f,x}$, $K_{f,y}$ and λ. Moreover, we have, a.s.,

$$\mathbb{E}^{\mathbb{Q}}\Big[\int_t^T e^{2\lambda s}|\tilde{Z}_s|^2\,ds|\mathcal{F}_t\Big] \le C\mathbb{E}^{\mathbb{Q}}\Big[\int_t^T |Z_s|^2 + |\alpha_s|^2\,ds|\mathcal{F}_t\Big] \le C\big(\|Z\|^2_{\mathrm{BMO}(\mathbb{Q})} + (T-t)\big).$$

But $\|Z\|_{\mathrm{BMO}(\mathbb{Q})}$ does not depend on K_g because (Y, Z) is a solution of the following quadratic BSDE:

$$\begin{aligned} Y_t &= g(X_T) + \int_t^T \big(f(s, X_s, Y_s, Z_s) - Z_s\nabla_z f(s, X_s, Y_s, Z_s)\big)\,ds \\ &\quad - \int_t^T Z_s\,d\tilde{W}_s. \end{aligned} \tag{11}$$

Finally, $|Z_t| \le C(1 + (T-t)^{-1/2})$ a.s.

When σ is invertible, the inequality (9) is verified with $\lambda := M_{\sigma^{-1}}(M_\sigma K_b + K_\sigma)$. Since this λ does not depend on ∇b and σ', we can prove the result when $b(t,\cdot)$ and σ are not differentiable by a standard approximation and stability results for BSDEs with linear growth. So, we are allowed to replace assumption (HX1) by (HX1′).

When f is not differentiable and g is only Lipschitz, we can prove the result by a standard approximation and stability results for linear BSDEs. But we notice that our estimation on Z does not depend on K_g. This allows us to weaken the hypothesis on g further; when g is only lower or upper semi-continuous the result stays true. The proof is the same as the proof of Proposition 4.3 in [8]. □

REMARK 3.4. The previous proof gives us a more precise estimation for a version of Z when f is differentiable with respect to z : $\forall t \in [0, T]$,

$$|Z_t| \le C + C'\mathbb{E}^{\mathbb{Q}}\Big[\int_t^T |Z_s|^2\,ds|\mathcal{F}_t\Big]^{1/2}(T-t)^{-1/2}.$$

REMARK 3.5. When assumption (HX1) or (HX1′) is not verified, the process Z may blow up before T. Zhang gives an example of such a phenomenon in dimension 1; we refer the reader to Example 1 in [20].

3.3. *Zhang's path regularity theorem.* Let $0 = t_0 < t_1 < \cdots < t_n = T$ be any given partition of $[0, T]$ and denote δ_n the mesh size of this partition. We define a set of random variables

$$\bar{Z}_{t_i} = \frac{1}{t_{i+1} - t_i}\mathbb{E}\Big[\int_{t_i}^{t_{i+1}} Z_s\,ds|\mathcal{F}_{t_i}\Big] \qquad \forall i \in \{0, \ldots, n-1\}.$$

Then we are able to give a more detailed version at Theorem 3.4.3 in [21].

THEOREM 3.6. *Suppose that* (HX0), (HY0) *hold and g is a Lipschitz function with Lipschitz constant K_g. Then we have*

$$\sum_{i=0}^{n-1} \mathbb{E}\Big[\int_{t_i}^{t_{i+1}} |Z_t - \bar{Z}_{t_i}|^2\, dt\Big] \le C(1 + K_g^2)\delta_n,$$

where C is a positive constant independent of δ_n and K_g.

PROOF. We will follow the proof of Theorem 5.6 in [14]; we just need to specify how the estimate depends on K_g. First, it is not difficult to show that $\bar{Z}_{t_i}$ is the best $\mathcal{F}_{t_i}$-measurable approximation of Z in $\mathcal{M}^2([t_i, t_{i+1}])$, that is,

$$\mathbb{E}\Big[\int_{t_i}^{t_{i+1}} |Z_t - \bar{Z}_{t_i}|^2\, dt\Big] = \inf_{Z_i \in L^2(\Omega, \mathcal{F}_{t_i})} \mathbb{E}\Big[\int_{t_i}^{t_{i+1}} |Z_t - Z_i|^2\, dt\Big].$$

In particular,

$$\mathbb{E}\Big[\int_{t_i}^{t_{i+1}} |Z_t - \bar{Z}_{t_i}|^2\, dt\Big] \le \mathbb{E}\Big[\int_{t_i}^{t_{i+1}} |Z_t - Z_{t_i}|^2\, dt\Big].$$

In the same spirit as previous proofs, we suppose in a first time that b, g and f are differentiable with respect to x, y and z. So,

$$Z_t - Z_{t_i} = \nabla Y_t (\nabla X_t)^{-1}\sigma(t) - \nabla Y_{t_i}(\nabla X_{t_i})^{-1}\sigma(t_i) = I_1 + I_2 + I_3 \qquad \text{a.s.},$$

with $I_1 = \nabla Y_t(\nabla X_t)^{-1}(\sigma(t) - \sigma(t_i))$, $I_2 = \nabla Y_t((\nabla X_t)^{-1} - (\nabla X_{t_i})^{-1})\sigma(t_i)$ and $I_3 = \nabla(Y_t - Y_{t_i})(\nabla X_{t_i})^{-1}\sigma(t_i)$. First, thanks to the estimation (8) we have

$$|I_1|^2 \le |\nabla Y_t|^2 e^{2K_b T} K_\sigma^2 |t_{i+1} - t_i|^2 \le C(1 + K_g^2)\delta_n^2.$$

We obtain the same estimation for $|I_2|$ because

$$|(\nabla X_t)^{-1} - (\nabla X_{t_i})^{-1}| \le \Big|\int_{t_i}^{t} (\nabla X_s)^{-1}\nabla b(s, X_s)\, ds\Big| \le K_b e^{K_b T}|t - t_i|.$$

Last, $|I_3| \le M_\sigma e^{K_b T}|\nabla Y_t - \nabla Y_{t_i}|$. So,

$$\sum_{i=0}^{n-1} \mathbb{E}\Big[\int_{t_i}^{t_{i+1}} |I_3|^2\, dt\Big] \le C\delta_n \sum_{i=0}^{n-1} \mathbb{E}\Big[\operatorname*{ess\,sup}_{t\in[t_i, t_{i+1}]} |\nabla Y_t - \nabla Y_{t_i}|^2\Big].$$

By using the BSDE (7), (HY0), the estimate on ∇X_s and the estimate (8), we have

$$|\nabla Y_t - \nabla Y_{t_i}|^2 \le C\Big(\int_{t_i}^{t} \big(C(1 + K_g) + |\nabla_z f(s, X_s, Y_s, Z_s)||\nabla Z_s|\big)\, ds\Big)^2 + C\Big(\int_{t_i}^{t} \nabla Z_s\, dW_s\Big)^2.$$

The inequalities of Hölder and Burkholder–Davis–Gundy give us

$$\begin{aligned}
&\sum_{i=0}^{n-1}\mathbb{E}\Big[\operatorname*{ess\,sup}_{t\in[t_i,t_{i+1}]}|\nabla Y_t-\nabla Y_{t_i}|^2\Big]\\
&\quad\le C(1+K_g^2)+C\sum_{i=0}^{n-1}\mathbb{E}\bigg(\int_{t_i}^{t_{i+1}}|\nabla_z f(s,X_s,Y_s,Z_s)||\nabla Z_s|\,ds\bigg)^2\\
&\qquad+C\mathbb{E}\bigg(\int_{t_i}^{t_{i+1}}|\nabla Z_s|^2\,ds\bigg)\\
&\quad\le C(1+K_g^2)\\
&\qquad+C\mathbb{E}\bigg[\bigg(\int_0^T|\nabla_z f(s,X_s,Y_s,Z_s)||\nabla Z_s|\,ds\bigg)^2+\int_0^T|\nabla Z_s|^2\,ds\bigg]\\
&\quad\le C(1+K_g^2)\\
&\qquad+C\mathbb{E}\bigg[\bigg(\int_0^T(1+|Z_s|^2)\,ds\bigg)\bigg(\int_0^T|\nabla Z_s|^2\,ds\bigg)+\int_0^T|\nabla Z_s|^2\,ds\bigg]\\
&\quad\le C(1+K_g^2)\\
&\qquad+C\bigg(1+\mathbb{E}\bigg[\bigg(\int_0^T|Z_s|^2\,ds\bigg)^p\bigg]^{1/p}\bigg)\mathbb{E}\bigg[\bigg(\int_0^T|\nabla Z_s|^2\,ds\bigg)^q\bigg]^{1/q}
\end{aligned}$$

for all $p>1$ and $q>1$ such that $1/p+1/q=1$. But, $(\nabla Y,\nabla Z)$ is the solution of BSDE (7) so, from Corollary 9 in [2], there exists q that only depends on $\|Z*W\|_{\mathrm{BMO}}$ such that

$$\mathbb{E}\bigg[\bigg(\int_0^T|\nabla Z_s|^2\,ds\bigg)^q\bigg]^{1/q}\le C(1+K_g^2).$$

Moreover, we can apply Lemma 2.1 to obtain the estimate

$$\mathbb{E}\bigg[\bigg(\int_0^T|Z_s|^2\,ds\bigg)^p\bigg]^{1/p}\le C\|Z\|_{\mathrm{BMO}}^2\le C.$$

Finally,

$$\sum_{i=0}^{n-1}\mathbb{E}\bigg[\int_{t_i}^{t_{i+1}}|I_3|^2\,dt\bigg]\le C(1+K_g^2)\delta_n$$

and

$$\begin{aligned}
\sum_{i=0}^{n-1}\mathbb{E}\bigg[\int_{t_i}^{t_{i+1}}|Z_t-\bar{Z}_{t_i}|^2\,dt\bigg]&\le\sum_{i=0}^{n-1}\mathbb{E}\bigg[\int_{t_i}^{t_{i+1}}(|I_1|^2+|I_2|^2+|I_3|^2)\,dt\bigg]\\
&\le C(1+K_g^2)\delta_n.
\end{aligned}$$

□

4. Convergence of a modified time discretization scheme for the BSDE.

4.1. *An approximation of the quadratic BSDE.* In a first time we will approximate our quadratic BSDE (5) by another one. We set $\varepsilon \in]0, T[$ and $N \in \mathbb{N}$. Let $(Y_t^{N,\varepsilon}, Z_t^{N,\varepsilon})$ be the solution of the BSDE

$$Y_t^{N,\varepsilon} = g_N(X_T) + \int_t^T f^\varepsilon(s, X_s, Y_s^{N,\varepsilon}, Z_s^{N,\varepsilon})\, ds - \int_t^T Z_s^{N,\varepsilon}\, dW_s \tag{12}$$

with

$$f^\varepsilon(s, x, y, z) := \mathbb{1}_{s \le T-\varepsilon} f(s, x, y, z) + \mathbb{1}_{s > T-\varepsilon} f(s, x, y, 0)$$

and g_N a Lipschitz approximation of g with Lipschitz constant N. f^ε verifies assumption (HY0) with the same constants as f. Since g_N is a Lipschitz function, $Z^{N,\varepsilon}$ has a bounded version and the BSDE (12) is a BSDE with a linear growth. Moreover, we can apply Theorem 3.3 to obtain the following proposition.

PROPOSITION 4.1. *Let us assume that* (HX0), (HY0) *and* (HX1) *or* (HX1′) *hold. There exists a version of* $Z^{N,\varepsilon}$ *and there exist three constants* $M_{z,1}, M_{z,2}, M_{z,3} \in \mathbb{R}^+$ *that do not depend on* N *and* ε *such that*, $\forall s \in [0, T]$,

$$|Z_s^{N,\varepsilon}| \le \left(M_{z,1} + \frac{M_{z,2}}{(T-s)^{1/2}}\right) \wedge \big(M_{z,3}(N+1)\big).$$

Thanks to BMO tools we have a stability result for quadratic BSDEs (see [2] and [14]).

PROPOSITION 4.2. *Let us assume that* (HX0) *and* (HY0) *hold. There exists a constant* C *that does not depend on* N *and* ε *such that*

$$\mathbb{E}\Big[\sup_{t\in[0,T]} |Y_t^{N,\varepsilon} - Y_t|^2\Big] + \mathbb{E}\left[\int_0^T |Z_t^{N,\varepsilon} - Z_t|^2\, dt\right] \le C\big(e_1(N) + e_2(N, \varepsilon)\big)$$

with

$$e_1(N) := \mathbb{E}[|g_N(X_T) - g(X_T)|^{2q}]^{1/q},$$

$$e_2(N, \varepsilon) := \mathbb{E}\left[\left(\int_{T-\varepsilon}^T |f(t, X_t, Y_t^{N,\varepsilon}, Z_t^{N,\varepsilon}) - f(t, X_t, Y_t^{N,\varepsilon}, 0)|\, dt\right)^{2q}\right]^{1/q}$$

and q *defined in Theorem* 2.2.

REMARK 4.3. The authors of [14] obtain this result with q^2 instead of q. Nevertheless, we are able to obtain the good result by applying the estimates of [2].

Then, in a second time, we will approximate our modified backward–forward system by a discrete-time one. We will slightly modify the classical discretization by using a nonequidistant net with $2n+1$ discretization times. We define the $n+1$ first discretization times on $[0, T-\varepsilon]$ by

$$t_k = T\left(1 - \left(\frac{\varepsilon}{T}\right)^{k/n}\right)$$

and we use an equidistant net on $[T-\varepsilon, T]$ for the last n discretization times

$$t_k = T - \left(\frac{2n-k}{n}\right)\varepsilon, \qquad n \le k \le 2n.$$

We denote the time step by $(h_k := t_{k+1} - t_k)_{0\le k\le 2n-1}$. We consider $(X^n_{t_k})_{0\le k\le 2n}$ the classical Euler scheme for X given by

$$\begin{aligned} X^n_0 &= x, \\ X^n_{t_{k+1}} &= X^n_{t_k} + h_k b(t_k, X^n_{t_k}) + \sigma(t_k)(W_{t_{k+1}} - W_{t_k}) \end{aligned} \tag{13}$$

for $0 \le k \le 2n-1$. We denote $\rho_s : \mathbb{R}^{1\times d} \to \mathbb{R}^{1\times d}$ the projection on the ball

$$B\left(0, M_{z,1} + \frac{M_{z,2}}{(T-s)^{1/2}}\right)$$

with $M_{z,1}$ and $M_{z,2}$ given by Proposition 4.1. Finally, we denote $(Y^{N,\varepsilon,n}, Z^{N,\varepsilon,n})$ our time approximation of $(Y^{N,\varepsilon}, Z^{N,\varepsilon})$. This couple is obtained by a slight modification of the classical dynamic programming equation

$$Y^{N,\varepsilon,n}_{t_{2n}} = g_N(X^n_{t_{2n}}),$$

$$Z^{N,\varepsilon,n}_{t_k} = \rho_{t_{k+1}}\left(\frac{1}{h_k}\mathbb{E}_{t_k}[Y^{N,\varepsilon,n}_{t_{k+1}}(W_{t_{k+1}} - W_{t_k})]\right), \tag{14}$$

$$Y^{N,\varepsilon,n}_{t_k} = \mathbb{E}_{t_k}[Y^{N,\varepsilon,n}_{t_{k+1}}] + h_k\mathbb{E}_{t_k}[f^\varepsilon(t_k, X^n_{t_k}, Y^{N,\varepsilon,n}_{t_{k+1}}, Z^{N,\varepsilon,n}_{t_k})], \tag{15}$$

where $0 \le k \le 2n-1$ and $\mathbb{E}_{t_k}$ stand for the conditional expectation given $\mathcal{F}_{t_k}$. Let us notice that the classical dynamic programming equation does not use a projection in (14); it is the only difference with our time approximation (see, e.g., [11] for the classical case). This projection comes directly from the estimate of Z in Proposition 4.1. The aim of our work is to study the error of discretization

$$e(N,\varepsilon,n) := \sup_{0\le k\le 2n} \mathbb{E}[|Y^{N,\varepsilon,n}_{t_k} - Y_{t_k}|^2] + \sum_{k=0}^{2n-1} \mathbb{E}\left[\int_{t_k}^{t_{k+1}} |Z^{N,\varepsilon,n}_{t_k} - Z_t|^2\,dt\right].$$

It is easy to see that

$$e(N,\varepsilon,n) \le C\big(e_1(N) + e_2(N,\varepsilon) + e_3(N,\varepsilon,n)\big)$$

with $e_1(N)$ and $e_2(N,\varepsilon)$ defined in Proposition 4.2 and

$$e_3(N,\varepsilon,n) := \sup_{0\le k\le 2n} \mathbb{E}[|Y^{N,\varepsilon,n}_{t_k} - Y^{N,\varepsilon}_{t_k}|^2] + \sum_{k=0}^{2n-1} \mathbb{E}\left[\int_{t_k}^{t_{k+1}} |Z^{N,\varepsilon,n}_{t_k} - Z^{N,\varepsilon}_t|^2\,dt\right].$$

4.2. *Study of the time approximation error* $e_3(N, \varepsilon, n)$. We need an extra assumption.

(HY1). There exists a positive constant $K_{f,t}$ such that $\forall t, t' \in [0, T]$, $\forall x \in \mathbb{R}^d$, $\forall y \in \mathbb{R}$, $\forall z \in \mathbb{R}^{1\times d}$,

$$|f(t, x, y, z) - f(t', x, y, z)| \leq K_{f,t}|t - t'|^{1/2}.$$

Moreover, we set $\varepsilon = Tn^{-a}$ and $N = n^b$, with $a, b \in \mathbb{R}^{+,*}$ two parameters. Before giving our error estimates, we recall two technical lemmas that we will prove in Appendices A and B.

LEMMA 4.4. *For all constant $M > 0$ there exists a constant C that depends only on T, M and a, such that*

$$\prod_{i=0}^{2n-1} (1 + Mh_i) \leq C \qquad \forall n \in \mathbb{N}^*.$$

LEMMA 4.5. *For all constants $M_1 > 0$ and $M_2 > 0$ there exists a constant C that depends only on T, M_1, M_2 and a, such that*

$$\prod_{i=0}^{n-1} \left(1 + M_1 h_i + M_2 \frac{h_i}{T - t_{i+1}}\right) \leq C n^{aM_2}.$$

First, we give a convergence result for the Euler scheme.

PROPOSITION 4.6. *Assume* (HX0) *holds. Then there exists a constant C that does not depend on n, such that*

$$\sup_{0\leq k\leq 2n} \mathbb{E}[|X_{t_k} - X^n_{t_k}|^2] \leq C \frac{\ln n}{n}.$$

PROOF. We just have to copy the classical proof to obtain, thanks to Lemma 4.4,

$$\sup_{0\leq k\leq 2n} \mathbb{E}[|X_{t_k} - X^n_{t_k}|^2] \leq C \sup_{0\leq i\leq 2n-1} h_i = Ch_0.$$

But

$$h_0 = T(1 - n^{-a/n}) \leq C \frac{\ln n}{n},$$

because $(1 - n^{-a/n}) \sim aT\frac{\ln n}{n}$ when $n \to +\infty$, so the proof is complete. □

Now, let us treat the BSDE approximation. In a first time we will study the time approximation error on $[T - \varepsilon, T]$.

PROPOSITION 4.7. *Assume that* (HX0), (HY0) *and* (HY1) *hold. Then there exists a constant C that does not depend on n and such that*

$$\sup_{n\le k\le 2n}\mathbb{E}[|Y_{t_k}^{N,\varepsilon,n}-Y_{t_k}^{N,\varepsilon}|^2]+\sum_{k=n}^{2n-1}\mathbb{E}\left[\int_{t_k}^{t_{k+1}}|Z_{t_k}^{N,\varepsilon,n}-Z_t^{N,\varepsilon}|^2\,dt\right]\le\frac{C\ln n}{n^{1-2b}}.$$

PROOF. The BSDE (12) has a linear growth with respect to z on $[T-\varepsilon,T]$ so we are allowed to apply classical results which give us that

$$\begin{aligned}&\sup_{n\le k\le 2n}\mathbb{E}[|Y_{t_k}^{N,\varepsilon,n}-Y_{t_k}^{N,\varepsilon}|^2]+\sum_{k=n}^{2n-1}\mathbb{E}\left[\int_{t_k}^{t_{k+1}}|Z_{t_k}^{N,\varepsilon,n}-Z_t^{N,\varepsilon}|^2\,dt\right]\\&\quad\le C\left(\mathbb{E}[|g_N(X_T)-g_N(X_T^n)|^2]+\frac{\varepsilon}{n}\right)\end{aligned}$$

by using the fact that g_N is N-Lipschitz and by applying Proposition 4.6. □

REMARK 4.8.

(1) When $a\ge 1-2b$, then $\varepsilon=Tn^{-a}=o(n^{2b-1}\ln n)$. We do not need to have a discretization grid on $[T-\varepsilon,T]$; $n+2$ points of discretization are sufficient on $[0,T]$.

(2) When $a<1-2b$, then it is possible to take only $\lceil n^c\rceil$ discretization points on $[T-\varepsilon,T]$ with $a+c=1-2b$. In this case the error bound becomes

$$\begin{aligned}&\sup_{n\le k\le 2n}\mathbb{E}[|Y_{t_k}^{N,\varepsilon,n}-Y_{t_k}^{N,\varepsilon}|^2]+\sum_{k=n}^{2n-1}\mathbb{E}\left[\int_{t_k}^{t_{k+1}}|Z_{t_k}^{N,\varepsilon,n}-Z_t^{N,\varepsilon}|^2\,dt\right]\\&\quad\le C\left(\frac{\ln n}{n^{1-2b}}+\frac{1}{n^{a+c}}\right)\end{aligned}$$

and the Proposition 4.7 stays true.

Now, let us see what happens on $[0,T-\varepsilon]$.

THEOREM 4.9. *Assume that* (HX0), (HY0), (HY1) *and* (HX1) *or* (HX1′) *hold. Then for all $\eta>0$, there exists a constant C that does not depend on N, ε and n, such that*

$$\sup_{0\le k\le 2n}\mathbb{E}[|Y_{t_k}^{N,\varepsilon,n}-Y_{t_k}^{N,\varepsilon}|^2]+\sum_{k=0}^{2n-1}\mathbb{E}\left[\int_{t_k}^{t_{k+1}}|Z_{t_k}^{N,\varepsilon,n}-Z_t^{N,\varepsilon}|^2\,dt\right]\le\frac{C}{n^{1-2b-Ka}}$$

with $K=4(1+\eta)L_{f,z}^2M_{z,2}^2$.

PROOF. First, we will study the error on Y. From (12) and (15) we get

$$\begin{aligned} Y_{t_k}^{N,\varepsilon} - Y_{t_k}^{N,\varepsilon,n} &= \mathbb{E}_{t_k}[Y_{t_{k+1}}^{N,\varepsilon} - Y_{t_{k+1}}^{N,\varepsilon,n}] \\ &\quad + \mathbb{E}_{t_k}\int_{t_k}^{t_{k+1}} \big(f(s, X_s, Y_s^{N,\varepsilon}, Z_s^{N,\varepsilon}) - f(t_k, X_{t_k}^n, Y_{t_{k+1}}^{N,\varepsilon,n}, Z_{t_k}^{N,\varepsilon,n})\big)\, ds. \end{aligned}$$

We introduce a parameter $\gamma_k > 0$ that will be chosen later. Thanks to Proposition 4.1 and assumption (HY0), f is Lipschitz on $[t_k, t_{k+1}]$ with a Lipschitz constant $K_k := K^1 + \frac{K^2}{(T-t_{k+1})^{1/2}}$ where $K^2 = 2L_{f,z}M_{z,2}$. A combination of Young's inequality $(a+b)^2 \le (1+\gamma_k h_k)a^2 + (1+\frac{1}{\gamma_k h_k})b^2$ and properties of f gives

$$\begin{aligned} &\mathbb{E}|Y_{t_k}^{N,\varepsilon} - Y_{t_k}^{N,\varepsilon,n}|^2 \\ &\quad \le (1+\gamma_k h_k)\mathbb{E}|\mathbb{E}_{t_k}[Y_{t_{k+1}}^{N,\varepsilon} - Y_{t_{k+1}}^{N,\varepsilon,n}]|^2 \\ &\quad\quad + (1+\eta)^{1/3}K_k^2\Big(h_k + \frac{1}{\gamma_k}\Big)\mathbb{E}\int_{t_k}^{t_{k+1}} |Z_s^{N,\varepsilon} - Z_{t_k}^{N,\varepsilon,n}|^2\, ds \\ &\quad\quad + C\Big(h_k + \frac{1}{\gamma_k}\Big)\Big(h_k^2 + \int_{t_k}^{t_{k+1}} \mathbb{E}|X_s - X_{t_k}^n|^2\, ds\Big) \\ &\quad\quad + C\Big(h_k + \frac{1}{\gamma_k}\Big)\Big(\int_{t_k}^{t_{k+1}} \mathbb{E}|Y_s^{N,\varepsilon} - Y_{t_{k+1}}^{N,\varepsilon,n}|^2\, ds\Big). \end{aligned} \tag{16}$$

We define

$$\tilde{Z}_{t_k}^{N,\varepsilon,n} := \frac{1}{h_k}\mathbb{E}_{t_k}[Y_{t_{k+1}}^{N,\varepsilon,n}(W_{t_{k+1}} - W_{t_k})].$$

So, $Z_{t_k}^{N,\varepsilon,n} = \rho_{t_{k+1}}(\tilde{Z}_{t_k}^{N,\varepsilon,n})$. Moreover, Proposition 4.1 implies that $Z_s^{N,\varepsilon} = \rho_{t_{k+1}}(Z_s^{N,\varepsilon})$ and, since $\rho_{t_{k+1}}$ is 1-Lipschitz, we have

$$|Z_s^{N,\varepsilon} - Z_{t_k}^{N,\varepsilon,n}|^2 = |\rho_{t_{k+1}}(Z_s^{N,\varepsilon}) - \rho_{t_{k+1}}(\tilde{Z}_{t_k}^{N,\varepsilon,n})|^2 \le |Z_s^{N,\varepsilon} - \tilde{Z}_{t_k}^{N,\varepsilon,n}|^2. \tag{17}$$

As in Theorem 3.6, we define $\bar{Z}_{t_k}^{N,\varepsilon}$ by

$$\begin{aligned} h_k\bar{Z}_{t_k}^{N,\varepsilon} &:= \mathbb{E}_{t_k}\int_{t_k}^{t_{k+1}} Z_s^{N,\varepsilon}\, ds \\ &= \mathbb{E}_{t_k}\Big(\Big(Y_{t_{k+1}}^{N,\varepsilon} + \int_{t_k}^{t_{k+1}} f(s, X_s, Y_s^{N,\varepsilon}, Z_s^{N,\varepsilon})\, ds\Big)^t (W_{t_{k+1}} - W_{t_k})\Big). \end{aligned}$$

Clearly,

$$\begin{aligned} &\mathbb{E}\int_{t_k}^{t_{k+1}} |Z_s^{N,\varepsilon} - \tilde{Z}_{t_k}^{N,\varepsilon,n}|^2\, ds \\ &\quad = \mathbb{E}\int_{t_k}^{t_{k+1}} |Z_s^{N,\varepsilon} - \bar{Z}_{t_k}^{N,\varepsilon}|^2\, ds + h_k\mathbb{E}|\bar{Z}_{t_k}^{N,\varepsilon} - \tilde{Z}_{t_k}^{N,\varepsilon,n}|^2. \end{aligned} \tag{18}$$

The Cauchy–Schwarz inequality yields

$$\begin{aligned}&\big|\mathbb{E}_{t_k}\big((Y_{t_{k+1}}^{N,\varepsilon}-Y_{t_{k+1}}^{N,\varepsilon,n})^t(W_{t_{k+1}}-W_{t_k})\big)\big|^2\\&\qquad\le h_k\{\mathbb{E}_{t_k}(|Y_{t_{k+1}}^{N,\varepsilon}-Y_{t_{k+1}}^{N,\varepsilon,n}|^2)-|\mathbb{E}_{t_k}(Y_{t_{k+1}}^{N,\varepsilon}-Y_{t_{k+1}}^{N,\varepsilon,n})|^2\}\end{aligned}$$

and consequently

$$\begin{aligned}&h_k\mathbb{E}|\bar{Z}_{t_k}^{N,\varepsilon}-\tilde{Z}_{t_k}^{N,\varepsilon,n}|^2\\&\qquad\le(1+\eta)^{1/3}\mathbb{E}[\mathbb{E}_{t_k}(|Y_{t_{k+1}}^{N,\varepsilon}-Y_{t_{k+1}}^{N,\varepsilon,n}|^2)-|\mathbb{E}_{t_k}(Y_{t_{k+1}}^{N,\varepsilon}-Y_{t_{k+1}}^{N,\varepsilon,n})|^2]\\&\qquad\quad+Ch_k\mathbb{E}\int_{t_k}^{t_{k+1}}|f(s,X_s,Y_s^{N,\varepsilon},Z_s^{N,\varepsilon})|^2\,ds.\end{aligned}\tag{19}$$

Plugging (18) and (19) into (16), we get

$$\begin{aligned}&\mathbb{E}|Y_{t_k}^{N,\varepsilon}-Y_{t_k}^{N,\varepsilon,n}|^2\\&\qquad\le(1+\gamma_kh_k)\mathbb{E}|\mathbb{E}_{t_k}[Y_{t_{k+1}}^{N,\varepsilon}-Y_{t_{k+1}}^{N,\varepsilon,n}]|^2\\&\qquad\quad+(1+\eta)K_k^2\Big(h_k+\frac{1}{\gamma_k}\Big)\mathbb{E}\int_{t_k}^{t_{k+1}}|Z_s^{N,\varepsilon}-\bar{Z}_{t_k}^{N,\varepsilon}|^2\,ds\\&\qquad\quad+C\Big(h_k+\frac{1}{\gamma_k}\Big)\Big(h_k^2+\int_{t_k}^{t_{k+1}}\mathbb{E}|X_s-X_{t_k}^n|^2\,ds\\&\qquad\qquad\qquad+\int_{t_k}^{t_{k+1}}\mathbb{E}|Y_s^{N,\varepsilon}-Y_{t_{k+1}}^{N,\varepsilon,n}|^2\,ds\Big)\\&\qquad\quad+(1+\eta)^{2/3}K_k^2\Big(h_k+\frac{1}{\gamma_k}\Big)\mathbb{E}[\mathbb{E}_{t_k}(|Y_{t_{k+1}}^{N,\varepsilon}-Y_{t_{k+1}}^{N,\varepsilon,n}|^2)\\&\qquad\qquad\qquad-|\mathbb{E}_{t_k}(Y_{t_{k+1}}^{N,\varepsilon}-Y_{t_{k+1}}^{N,\varepsilon,n})|^2]\\&\qquad\quad+CK_k^2\Big(h_k+\frac{1}{\gamma_k}\Big)h_k\mathbb{E}\int_{t_k}^{t_{k+1}}|f(s,X_s,Y_s^{N,\varepsilon},Z_s^{N,\varepsilon})|^2\,ds.\end{aligned}$$

Now write

$$\mathbb{E}|Y_s^{N,\varepsilon}-Y_{t_{k+1}}^{N,\varepsilon,n}|^2\le2\mathbb{E}|Y_s^{N,\varepsilon}-Y_{t_{k+1}}^{N,\varepsilon}|^2+2\mathbb{E}|Y_{t_{k+1}}^{N,\varepsilon}-Y_{t_{k+1}}^{N,\varepsilon,n}|^2,\tag{20}$$

$$\mathbb{E}|X_s-X_{t_k}^n|^2\le2\mathbb{E}|X_s-X_{t_k}|^2+2\mathbb{E}|X_{t_k}-X_{t_k}^n|^2\tag{21}$$

and we obtain

$$\begin{aligned}&\mathbb{E}|Y_{t_k}^{N,\varepsilon}-Y_{t_k}^{N,\varepsilon,n}|^2\\&\qquad\le(1+\gamma_kh_k)\mathbb{E}|\mathbb{E}_{t_k}[Y_{t_{k+1}}^{N,\varepsilon}-Y_{t_{k+1}}^{N,\varepsilon,n}]|^2\\&\qquad\quad+(1+\eta)K_k^2\Big(h_k+\frac{1}{\gamma_k}\Big)\mathbb{E}\int_{t_k}^{t_{k+1}}|Z_s^{N,\varepsilon}-\bar{Z}_{t_k}^{N,\varepsilon}|^2\,ds\end{aligned}$$

$$
\begin{aligned}
&+ C\left(h_k + \frac{1}{\gamma_k}\right)\left(h_k^2 + \int_{t_k}^{t_{k+1}} \mathbb{E}|X_s - X_{t_k}|^2\,ds + h_k \mathbb{E}|X_{t_k} - X_{t_k}^n|^2\right)\\
&+ C\left(h_k + \frac{1}{\gamma_k}\right)\left(\int_{t_k}^{t_{k+1}} \mathbb{E}|Y_s^{N,\varepsilon} - Y_{t_{k+1}}^{N,\varepsilon}|^2\,ds + h_k \mathbb{E}|Y_{t_{k+1}}^{N,\varepsilon} - Y_{t_{k+1}}^{N,\varepsilon,n}|^2\right)\\
&+ (1+\eta)^{2/3} K_k^2 \left(h_k + \frac{1}{\gamma_k}\right)\mathbb{E}[\mathbb{E}_{t_k}(|Y_{t_{k+1}}^{N,\varepsilon} - Y_{t_{k+1}}^{N,\varepsilon,n}|^2)\\
&\qquad\qquad - |\mathbb{E}_{t_k}(Y_{t_{k+1}}^{N,\varepsilon} - Y_{t_{k+1}}^{N,\varepsilon,n})|^2]\\
&+ C K_k^2 \left(h_k + \frac{1}{\gamma_k}\right) h_k \mathbb{E}\int_{t_k}^{t_{k+1}} |f(s, X_s, Y_s^{N,\varepsilon}, Z_s^{N,\varepsilon})|^2\,ds.
\end{aligned}
$$

Taking $\gamma_k = (1+\eta)^{2/3} K_k^2$ and for h_k small enough, it gives

$$
\begin{aligned}
&\mathbb{E}|Y_{t_k}^{N,\varepsilon} - Y_{t_k}^{N,\varepsilon,n}|^2\\
&\quad \le (1 + Ch_k + (1+\eta)^{2/3} K_k^2 h_k)\mathbb{E}|Y_{t_{k+1}}^{N,\varepsilon} - Y_{t_{k+1}}^{N,\varepsilon,n}|^2 + Ch_k^2\\
&\qquad + Ch_k \max_{0\le k\le n} \mathbb{E}|X_{t_k} - X_{t_k}^n|^2\\
&\qquad + C\mathbb{E}\int_{t_k}^{t_{k+1}} |Z_s^{N,\varepsilon} - \bar{Z}_{t_k}^{N,\varepsilon}|^2\,ds + C\int_{t_k}^{t_{k+1}} \mathbb{E}|X_s - X_{t_k}|^2\,ds\\
&\qquad + C\int_{t_k}^{t_{k+1}} \mathbb{E}|Y_s^{N,\varepsilon} - Y_{t_{k+1}}^{N,\varepsilon}|^2\,ds + Ch_k \mathbb{E}\int_{t_k}^{t_{k+1}} f(s, X_s, Y_s^{N,\varepsilon}, Z_s^{N,\varepsilon})^2\,ds,
\end{aligned}
$$

because $K_k^2 h_k \le C(h_0 + h_k(T - t_{k+1})^{-1}) \le C\frac{\ln n}{n}$. The Gronwall's lemma gives us

$$
\begin{aligned}
&\mathbb{E}|Y_{t_k}^{N,\varepsilon} - Y_{t_k}^{N,\varepsilon,n}|^2\\
&\quad \le C\sum_{j=0}^{n-1}\left[\prod_{i=0}^{j-1}(1 + Ch_i + (1+\eta)^{2/3} K_i^2 h_i)\right]\\
&\qquad \times \left[h_j^2 + h_j \max_{0\le l\le n} \mathbb{E}|X_{t_l} - X_{t_l}^n|^2\right.\\
&\qquad\quad + \mathbb{E}\int_{t_j}^{t_{j+1}} (|Z_s^{N,\varepsilon} - \bar{Z}_{t_j}^{N,\varepsilon}|^2 + |X_s - X_{t_j}|^2 + |Y_s^{N,\varepsilon} - Y_{t_{j+1}}^{N,\varepsilon}|^2)\,ds\\
&\qquad\qquad \left. + h_j \mathbb{E}\int_{t_j}^{t_{j+1}} |f(s, X_s, Y_s^{N,\varepsilon}, Z_s^{N,\varepsilon})|^2\,ds\right]\\
&\qquad + \left[\prod_{i=0}^{n-1}(1 + Ch_i + (1+\eta)^{2/3} K_i^2 h_i)\right]\mathbb{E}|Y_{t_n}^{N,\varepsilon} - Y_{t_n}^{N,\varepsilon,n}|^2.
\end{aligned}
$$

Then, we apply Lemma 4.5.

$$\begin{aligned}
&\mathbb{E}|Y_{t_k}^{N,\varepsilon} - Y_{t_k}^{N,\varepsilon,n}|^2 \\
&\quad \le Cn^{(1+\eta)(K^2)^2 a} \\
&\qquad \times \Bigg[h_0 + \max_{0\le l\le n} \mathbb{E}|X_{t_l} - X_{t_l}^n|^2 \\
&\qquad\quad + \sum_{j=0}^{n} \mathbb{E}\bigg(\int_{t_j}^{t_{j+1}} |Z_s^{N,\varepsilon} - \bar{Z}_{t_j}^{N,\varepsilon}|^2 + |X_s - X_{t_j}|^2 + |Y_s^{N,\varepsilon} - Y_{t_{j+1}}^{N,\varepsilon}|^2 \, ds \bigg) \\
&\qquad\qquad + h_0 \mathbb{E} \int_0^{t_n} |f(s, X_s, Y_s^{N,\varepsilon}, Z_s^{N,\varepsilon})|^2 \, ds + \mathbb{E}|Y_{t_n}^{N,\varepsilon} - Y_{t_n}^{N,\varepsilon,n}|^2 \Bigg].
\end{aligned}$$

A classical estimation gives us $\mathbb{E}[|X_s - X_{t_j}|^2] \le |s - t_j|$. Moreover, since $Z^{N,\varepsilon}$ is bounded,

$$\begin{aligned}
&\mathbb{E} \int_0^{t_n} |f(s, X_s, Y_s^{N,\varepsilon}, Z_s^{N,\varepsilon})|^2 \, ds \\
&\qquad \le CT(1 + |Y^{N,\varepsilon}|_\infty) + C\mathbb{E}\bigg[\int_0^{t_n} |Z_s^{N,\varepsilon}|^4 \, ds \bigg] \\
&\qquad \le CT(1 + |Y^{N,\varepsilon}|_\infty) + Cn^{2b}\mathbb{E}\bigg[\int_0^{T} |Z_s^{N,\varepsilon}|^2 \, ds \bigg].
\end{aligned}$$

But we have an a priori estimate for $\mathbb{E}[\int_0^T |Z_s^{N,\varepsilon}|^2 \, ds]$ that does not depend on N and ε. So

$$\mathbb{E} \int_0^{t_n} |f(s, X_s, Y_s^{N,\varepsilon}, Z_s^{N,\varepsilon})|^2 \, ds \le Cn^{2b}. \tag{22}$$

With the same type of argument we also have

$$\mathbb{E}|Y_s^{N,\varepsilon} - Y_{t_{j+1}}^{N,\varepsilon}|^2 \le Ch_j n^{2b}. \tag{23}$$

If we add Zhang's path regularity Theorem 3.6, Propositions 4.6 and 4.7, we finally obtain

$$\mathbb{E}|Y_{t_k}^{N,\varepsilon} - Y_{t_k}^{N,\varepsilon,n}|^2 \le Cn^{(1+\eta)(K^2)^2 a} \frac{n^{2b} \ln n}{n} = C \frac{\ln n}{n^{1-2b-(1+\eta)(K^2)^2 a}}. \tag{24}$$

Now, let us deal with the error on Z. First of all, (17) gives us

$$\sum_{k=0}^{n-1} \mathbb{E}\bigg[\int_{t_k}^{t_{k+1}} |Z_{t_k}^{N,\varepsilon,n} - Z_t^{N,\varepsilon}|^2 \, dt \bigg] \le \sum_{k=0}^{n-1} \mathbb{E}\bigg[\int_{t_k}^{t_{k+1}} |\tilde{Z}_{t_k}^{N,\varepsilon,n} - Z_t^{N,\varepsilon}|^2 \, dt \bigg].$$

For $0 \le k \le n-1$, we can use (18) and (19) to obtain

$$\begin{aligned}
&\mathbb{E}\Big[\int_{t_k}^{t_{k+1}} |\tilde{Z}_{t_k}^{N,\varepsilon,n} - Z_t^{N,\varepsilon}|^2\,dt\Big] \\
&\quad\le \mathbb{E}\Big[\int_{t_k}^{t_{k+1}} |\bar{Z}_{t_k}^{N,\varepsilon} - Z_t^{N,\varepsilon}|^2\,dt\Big] \\
&\qquad + (1+\eta)^{2/3}\mathbb{E}[\mathbb{E}_{t_k}(|Y_{t_{k+1}}^{N,\varepsilon} - Y_{t_{k+1}}^{N,\varepsilon,n}|^2) - |\mathbb{E}_{t_k}(Y_{t_{k+1}}^{N,\varepsilon} - Y_{t_{k+1}}^{N,\varepsilon,n})|^2] \\
&\qquad + Ch_k\mathbb{E}\Big[\int_{t_k}^{t_{k+1}} |f(s, X_s, Y_s^{N,\varepsilon}, Z_s^{N,\varepsilon})|^2\,ds\Big].
\end{aligned}$$

Inequality (22) and estimates for Z give us

$$\begin{aligned}
&\sum_{k=0}^{n-1}\mathbb{E}\Big[\int_{t_k}^{t_{k+1}} |Z_{t_k}^{N,\varepsilon,n} - Z_t^{N,\varepsilon}|^2\,dt\Big] \\
&\quad\le \sum_{k=0}^{n-1}\mathbb{E}\Big[\int_{t_k}^{t_{k+1}} |\bar{Z}_{t_k}^{N,\varepsilon} - Z_t^{N,\varepsilon}|^2\,dt\Big] \\
&\qquad + (1+\eta)^{2/3}\sum_{k=0}^{n-1}\mathbb{E}[\mathbb{E}_{t_k}(|Y_{t_{k+1}}^{N,\varepsilon} - Y_{t_{k+1}}^{N,\varepsilon,n}|^2) - |\mathbb{E}_{t_k}(Y_{t_{k+1}}^{N,\varepsilon} - Y_{t_{k+1}}^{N,\varepsilon,n})|^2] \\
&\qquad + Ch_0\mathbb{E}\Big[\int_0^T |f(s, X_s, Y_s^{N,\varepsilon}, Z_s^{N,\varepsilon})|^2\,ds\Big] \qquad (25)\\
&\quad\le \sum_{k=0}^{n-1}\mathbb{E}\Big[\int_{t_k}^{t_{k+1}} |\bar{Z}_{t_k}^{N,\varepsilon} - Z_t^{N,\varepsilon}|^2\,dt\Big] \\
&\qquad + (1+\eta)^{2/3}\sum_{k=0}^{n-1}\mathbb{E}[\mathbb{E}_{t_k}(|Y_{t_k}^{N,\varepsilon} - Y_{t_k}^{N,\varepsilon,n}|^2) - |\mathbb{E}_{t_k}(Y_{t_{k+1}}^{N,\varepsilon} - Y_{t_{k+1}}^{N,\varepsilon,n})|^2] \\
&\qquad + C\mathbb{E}[|Y_{t_n}^{N,\varepsilon} - Y_{t_n}^{N,\varepsilon,n}|^2] + Ch_0n^{2b}
\end{aligned}$$

with an index change in the penultimate line. Then, by using (16) we get

$$\begin{aligned}
&(1+\eta)^{2/3}\mathbb{E}[\mathbb{E}_{t_k}(|Y_{t_k}^{N,\varepsilon} - Y_{t_k}^{N,\varepsilon,n}|^2) - |\mathbb{E}_{t_k}(Y_{t_{k+1}}^{N,\varepsilon} - Y_{t_{k+1}}^{N,\varepsilon,n})|^2] \\
&\quad\le C\gamma_k h_k\mathbb{E}|\mathbb{E}_{t_k}[Y_{t_{k+1}}^{N,\varepsilon} - Y_{t_{k+1}}^{N,\varepsilon,n}]|^2 \\
&\qquad + (1+\eta)K_k^2\Big(h_k + \frac{1}{\gamma_k}\Big)\mathbb{E}\int_{t_k}^{t_{k+1}} |Z_s^{N,\varepsilon} - Z_{t_k}^{N,\varepsilon,n}|^2\,ds \qquad (26)\\
&\qquad + C\Big(h_k + \frac{1}{\gamma_k}\Big)h_k\Big(h_k + \sup_{s\in[t_k,t_{k+1}]}\mathbb{E}[|X_s - X_{t_k}^n|^2 + |Y_s^{N,\varepsilon} - Y_{t_{k+1}}^{N,\varepsilon,n}|^2]\Big).
\end{aligned}$$

Thanks to (20), (21), (23) and a classical estimation on $\mathbb{E}[|X_s - X_{t_k}|^2]$ we have

$$\sup_{s\in[t_k,t_{k+1}]} \mathbb{E}[|X_s - X^n_{t_k}|^2 + |Y^{N,\varepsilon}_s - Y^{N,\varepsilon,n}_{t_{k+1}}|^2] \le C(h_k n^{2b} + \mathbb{E}[|Y^{N,\varepsilon}_{t_{k+1}} - Y^{N,\varepsilon,n}_{t_{k+1}}|^2]).$$

Let us set $\gamma_k = 3(1+\eta)K_k^2$. We recall that $h_k K_k^2 \le \frac{C\ln n}{n} \to 0$ when $n \to 0$. So, for n big enough, (26) becomes

$$\begin{aligned}(1+\eta)^{2/3}&\mathbb{E}[\mathbb{E}_{t_k}(|Y^{N,\varepsilon}_{t_k} - Y^{N,\varepsilon,n}_{t_k}|^2) - |\mathbb{E}_{t_k}(Y^{N,\varepsilon}_{t_{k+1}} - Y^{N,\varepsilon,n}_{t_{k+1}})|^2] \\ &\le \frac{C\ln n}{n}\mathbb{E}[|Y^{N,\varepsilon}_{t_{k+1}} - Y^{N,\varepsilon,n}_{t_{k+1}}|^2] + \frac{1}{2}\mathbb{E}\int_{t_k}^{t_{k+1}} |Z^{N,\varepsilon}_s - Z^{N,\varepsilon,n}_{t_k}|^2\,ds \\ &\quad + Ch_0 h_k n^{2b}.\end{aligned}$$

If we inject this last estimate in (25) and we use Theorem 3.6, we obtain

$$\begin{aligned}\frac{1}{2}\sum_{k=0}^{n-1}\mathbb{E}\Big[\int_{t_k}^{t_{k+1}} |Z^{N,\varepsilon,n}_{t_k} - Z^{N,\varepsilon}_t|^2\,dt\Big] \\ \le Ch_0 n^{2b} + C\ln n \sup_{0\le k\le n-1}\mathbb{E}[|Y^{N,\varepsilon}_{t_{k+1}} - Y^{N,\varepsilon,n}_{t_{k+1}}|^2].\end{aligned}$$

By using (24) and Proposition 4.7, we finally have

$$\begin{aligned}\sup_{0\le k\le 2n}\mathbb{E}[|Y^{N,\varepsilon,n}_{t_k} - Y^{N,\varepsilon}_{t_k}|^2] + \sum_{k=0}^{2n-1}\mathbb{E}\Big[\int_{t_k}^{t_{k+1}} |Z^{N,\varepsilon,n}_{t_k} - Z^{N,\varepsilon}_t|^2\,dt\Big] \\ \le C\frac{(\ln n)^2}{n^{1-2b-Ka}}\end{aligned}$$

with $K = 4(1+\eta)L^2_{f,z}M^2_{z,2}$. Since this estimate is true for every $\eta > 0$, we have proved the result. □

4.3. *Study of the global error* $e(N,\varepsilon,n)$. Let us study errors $e_1(N)$ and $e_2(N,\varepsilon)$.

PROPOSITION 4.10. *Let us assume that* (HX0) *and* (HY0) *hold. There exists a constant* $C > 0$ *such that*

$$e_2(N,\varepsilon) \le \frac{C}{n^{2a-4b}}.$$

PROOF. We just have to notice that

$$|f(t,X_t,Y^{N,\varepsilon}_t,Z^{N,\varepsilon}_t) - f(t,X_t,Y^{N,\varepsilon}_t,0)| \le C|Z^{N,\varepsilon}_t|^2$$

and $|Z_t^{N,\varepsilon}|$ is bounded by Cn^b. $\square$

For g_N we use the classical Lipschitz approximation

$$g_N(x) = \inf\{g(u) + N|x-u| | u \in \mathbb{R}^d\}.$$

PROPOSITION 4.11. *We assume that* (HX0) *holds and g is α-Hölder. Then, there exists a constant C such that*

$$e_1(N) \leq \frac{C}{n^{2b\alpha/(1-\alpha)}}.$$

PROOF. g_N is a N-Lipschitz function and $g_N \to g$ when $N \to +\infty$ uniformly on $\mathbb{R}^d$. More precisely, we have

$$|g - g_N|_\infty \leq \frac{C}{N^{\alpha/(1-\alpha)}}. \qquad \square$$

REMARK 4.12. For some explicit examples, it is possible to have a better convergence speed. For example, let us take $g(x) = (|x|^\alpha \mathbb{1}_{x\geq 0}) \wedge C$ and assume that σ is invertible. Then, we can use the fact that this function is not Lipschitz only in 0 and obtain

$$e_1(N) \leq \frac{C}{n^{2\alpha b/(1-\alpha)}} \mathbb{P}\big(X_T \in \big[0, N^{-1/(1-\alpha)}\big]\big)^{1/q} \leq \frac{C}{n^{(b/(1-\alpha))(2\alpha+1/q)}}.$$

REMARK 4.13. It is also possible to obtain convergence speed when g is not α-Hölder. For example, we assume that σ is invertible and we set $g(x) = \prod_{i=1}^d \mathbb{1}_{x_i>0}(x)$. Then

$$e_1(N) \leq C\left[\sum_{i=1}^d \mathbb{P}\big((X_T)_i \in [0, 1/N]\big)\right]^{1/q} \leq \frac{C}{N^{1/q}} = \frac{C}{n^{b/q}}.$$

Now we are able to gather all these errors.

THEOREM 4.14. *We assume that* (HX0), (HY0), (HY1) *and* (HX1) *or* (HX1′) *hold. We assume also that g is α-Hölder. Then for all $\eta > 0$, there exists a constant $C > 0$ that does not depend on n such that*

$$e(n) := e(N, \varepsilon, n) \leq \frac{C}{n^{2\alpha/((2-\alpha)(2+K)-2+2\alpha)}}$$

with $K = 4(1+\eta)L_{f,z}^2 M_{z,2}^2$.

PROOF. Thanks to Theorem 4.9, Propositions 4.10 and 4.11 we have

$$e(n) \leq \frac{C}{n^{1-2b-Ka}} + \frac{C}{n^{2a-4b}} + \frac{C}{n^{2\alpha b/(1-\alpha)}}.$$

Then we only need to set $a := \frac{1+2b}{2+K}$ and $b := \frac{1-\alpha}{(2-\alpha)(2+K)-2+2\alpha}$ to obtain the result. □

COROLLARY 4.15. *We assume that assumptions of Theorem 4.14 hold. Moreover, we assume that f has a sub-quadratic growth with respect to z; there exists $0 < \beta < 1$ such that, for all $t \in [0, T]$, $x \in \mathbb{R}^d$, $y \in R$, $z, z' \in \mathbb{R}^{1\times d}$,*

$$|f(t, x, y, z) - f(t, x, y, z')| \le \big(K_{f,z} + L_{f,z}(|z|^\beta + |z'|^\beta)\big)|z - z'|.$$

Then we are allowed to take K as small as we want. So, for all $\eta > 0$, there exists a constant $C > 0$ that does not depend on n such that

$$e(n) \le \frac{C}{n^{\alpha-\eta}}.$$

REMARK 4.16. We are able to specify Remark 4.8 in our case, when $a = \frac{1+2b}{2+K}$ and $b = \frac{1-\alpha}{(2-\alpha)(2+K)-2+2\alpha}$.

(1) When $K \le \frac{2-3\alpha}{2-\alpha}$, that is to say, when $\alpha < 2/3$ and K is sufficiently small, then we do not need to have a discretization grid on $[T - \varepsilon, T]$.

(2) When $K > \frac{2-3\alpha}{2-\alpha}$, then it is possible to take only $\lceil n^c \rceil$ discretization points on $[T - \varepsilon, T]$ with

$$c = 1 + \frac{3\alpha - 4}{(2-\alpha)(2+K) - 2 + 2\alpha}.$$

Theorem 4.14 is not interesting in practice because the speed of convergence depends strongly on K. But we see that the global error becomes $e(n) \le \frac{C}{n^{\alpha-\eta}}$ when we are allowed to choose K as small as we want. Under extra assumption we can show that we are allowed to take the constant $M_{z,2}$ as small as we want.

(HX2). b is bounded on $[0, T] \times \mathbb{R}^d$ by a constant M_b.

THEOREM 4.17. *We assume that* (HX0), (HY0), (HY1), (HX2) *and* (HX1) *or* (HX1′) *hold. We assume also that g is α-Hölder. Then for all $\eta > 0$, there exists a constant $C > 0$ that does not depend on n such that*

$$e(n) \le \frac{C}{n^{\alpha-\eta}}.$$

REMARK 4.18. With the assumptions of the previous theorem, it is also possible to have an estimate of the global error for examples given in Remarks 4.12 and 4.13. When $g(x) = (|x|^\alpha \mathbb{1}_{x\ge 0}) \wedge C$, we have

$$e(n) \le \frac{C}{n^{\alpha+(1-\alpha)/(1+2q)-\eta}}$$

and when $g(x) = \prod_{i=1}^{d} \mathbb{1}_{x_i>0}(x)$, we have

$$e(n) \leq \frac{C}{n^{1/(1+2q)-\eta}}.$$

PROOF OF THEOREM 4.17. First, we suppose that f is differentiable with respect to z. Thanks to Remark 3.4 we see that it is sufficient to show that

$$\mathbb{E}^{\mathbb{Q}^{N,\varepsilon}}\left[\int_t^T |Z_s^{N,\varepsilon}|^2\,ds|\mathcal{F}_t\right]$$

is small uniformly in ω, N and ε when t is close to T. We will obtain an estimation for this quantity by applying the same computation as [2] for the BMO norm estimate of Z, page 831. Thus, we have

$$\mathbb{E}^{\mathbb{Q}^{N,\varepsilon}}\left[\int_t^T |Z_s^{N,\varepsilon}|^2\,ds|\mathcal{F}_t\right] \leq \mathbb{E}^{\mathbb{Q}^{N,\varepsilon}}[|\varphi(Y_T^{N,\varepsilon}) - \varphi(Y_t^{N,\varepsilon})||\mathcal{F}_t] + C(T-t)$$

with $\varphi(x) = (e^{2c(x+m)} - 2c(x+m) - 1)/(2c^2)$, $m = |Y|_\infty$ and c that depends on constants in assumption (HY0) but does not depend on $\nabla_z f$. Let us notice that m, c and so φ do not depend on N and ε. Since Y is bounded, φ is a Lipschitz function, so

$$\mathbb{E}^{\mathbb{Q}^{N,\varepsilon}}\left[\int_t^T |Z_s^{N,\varepsilon}|^2\,ds|\mathcal{F}_t\right] \leq C\mathbb{E}^{\mathbb{Q}^{N,\varepsilon}}[|Y_T^{N,\varepsilon} - Y_t^{N,\varepsilon}||\mathcal{F}_t] + C(T-t).$$

We denote by $(Y^{N,\varepsilon,t,x}, Z^{N,\varepsilon,t,x})$ the solution of BSDE (12) when $X_t^{t,x} = x$. As usual, we set $X_s^{t,x} = x$ and $Z_s^{N,\varepsilon,t,x} = 0$ for $s \leq t$ and we define $u^{N,\varepsilon}(t,x) := Y_t^{N,\varepsilon,t,x}$. Then we give a proposition that we will prove in Appendix C.

PROPOSITION 4.19. *We assume that* (HX0), (HY0), (HY1), (HX2) *and* (HX1) *or* (HX1′) *hold. We assume also that g is uniformly continuous on $\mathbb{R}^d$. Then $u^{N,\varepsilon}$ is uniformly continuous on $[0,T] \times \mathbb{R}^d$ and there exists ω a concave modulus of continuity for all functions in $\{u^{N,\varepsilon} | N \in \mathbb{N}, \varepsilon > 0\}$, that is, ω does not depend on N and ε.*

Then

$$\begin{aligned}
&\mathbb{E}^{\mathbb{Q}^{N,\varepsilon}}[|Y_T^{N,\varepsilon} - Y_t^{N,\varepsilon}||\mathcal{F}_t] \\
&\quad= \mathbb{E}^{\mathbb{Q}^{N,\varepsilon}}[|u^{N,\varepsilon}(T, X_T) - u^{N,\varepsilon}(t, X_t)||\mathcal{F}_t] \\
&\quad\leq \mathbb{E}^{\mathbb{Q}^{N,\varepsilon}}\big[\mathbb{1}_{|\int_t^T \sigma(s)\,d\tilde{W}_s| \leq \nu} |u^{N,\varepsilon}(T, X_T) - u^{N,\varepsilon}(t, X_t)| \\
&\qquad\qquad + 2|Y^{N,\varepsilon}|_\infty \mathbb{1}_{|\int_t^T \sigma(s)\,d\tilde{W}_s| > \nu} |\mathcal{F}_t\big] \\
&\quad\leq \mathbb{E}^{\mathbb{Q}^{N,\varepsilon}}\big[\omega\big(|T-t| + \mathbb{1}_{|\int_t^T \sigma(s)\,d\tilde{W}_s| \leq \nu} |X_T - X_t|\big) \\
&\qquad\qquad + 2|Y^{N,\varepsilon}|_\infty \mathbb{1}_{|\int_t^T \sigma(s)\,d\tilde{W}_s| > \nu} |\mathcal{F}_t\big]
\end{aligned}$$

with $d\tilde{W}_s = dW_s - \nabla_z f^\varepsilon(s, X_s, Y_s^{N,\varepsilon}, Z_s^{N,\varepsilon})\,ds$. But,

$$\begin{aligned}
&\mathbb{1}_{|\int_t^T \sigma(s)\,d\tilde{W}_s|\le\nu}|X_T - X_t| \\
&\quad= \mathbb{1}_{|\int_t^T \sigma(s)\,d\tilde{W}_s|\le\nu}\left|\int_t^T b(s, X_s)\,ds \right. \\
&\qquad\qquad \left. + \int_t^T \nabla_z f^\varepsilon(s, X_s, Y_s^{N,\varepsilon}, Z_s^{N,\varepsilon})\,ds + \int_t^T \sigma(s)\,d\tilde{W}_s\right| \\
&\quad\le M_b(T-t) + \nu + C\int_t^T (1 + |Z_s^{N,\varepsilon}|)\,ds \\
&\quad\le C(T-t) + \nu + C(T-t)^{1/2}\left(\int_t^T |Z_s^{N,\varepsilon}|^2\,ds\right)^{1/2}.
\end{aligned}$$

Since ω is concave, we have by Jensen's inequality

$$\begin{aligned}
&\mathbb{E}^{\mathbb{Q}^{N,\varepsilon}}\big[\omega\big(|T-t| + \mathbb{1}_{|\int_t^T \sigma(s)\,d\tilde{W}_s|\le\nu}|X_T - X_t|\big)|\mathcal{F}_t\big] \\
&\quad\le \omega\left(C|T-t| + \nu + C(T-t)^{1/2}\mathbb{E}^{\mathbb{Q}^{N,\varepsilon}}\left[\left(\int_t^T |Z_s^{N,\varepsilon}|^2\,ds\right)^{1/2}\Big|\mathcal{F}_t\right]\right) \\
&\quad\le \omega\left(C|T-t| + \nu + C(T-t)^{1/2}\mathbb{E}^{\mathbb{Q}^{N,\varepsilon}}\left[\int_t^T |Z_s^{N,\varepsilon}|^2\,ds|\mathcal{F}_t\right]^{1/2}\right) \\
&\quad\le \omega\big(C|T-t| + \nu + C(T-t)^{1/2}\|Z^{N,\varepsilon}\|_{\mathrm{BMO}(\mathbb{Q})}\big).
\end{aligned}$$

But, $\|Z^{N,\varepsilon}\|_{\mathrm{BMO}(\mathbb{Q})}$ only depends on constants in assumption (HY0), so it is bounded uniformly in N and ε. Moreover, $|\int_t^T \sigma(s)\,d\tilde{W}_s|$ is independent of $\mathcal{F}_t$ so we have by the Markov inequality

$$\begin{aligned}
\mathbb{E}^{\mathbb{Q}^{N,\varepsilon}}\big[\mathbb{1}_{|\int_t^T \sigma(s)\,d\tilde{W}_s|>\nu}|\mathcal{F}_t\big] &= \mathbb{Q}^{N,\varepsilon}\left(\left|\int_t^T \sigma(s)\,d\tilde{W}_s\right| > \nu\right) \\
&\le \frac{C(T-t)^{1/2}}{\nu}.
\end{aligned}$$

Finally, we have

$$\begin{aligned}
\mathbb{E}^{\mathbb{Q}^{N,\varepsilon}}[|Y_T^{N,\varepsilon} - Y_t^{N,\varepsilon}||\mathcal{F}_t] &\le \omega(C|T-t|^{1/2} + \nu) + C\frac{(T-t)^{1/2}}{\nu} \\
&\le \omega(C|T-t|^{1/2} + |T-t|^{1/4}) + C|T-t|^{1/4}
\end{aligned}$$

by setting $\nu = |T-t|^{1/4}$ and $\mathbb{E}^{\mathbb{Q}^{N,\varepsilon}}[|Y_T^{N,\varepsilon} - Y_t^{N,\varepsilon}||\mathcal{F}_t] \to 0$ uniformly in ω, N and ε when $t \to T$. When f is not differentiable with respect to z but is only locally Lipschitz, then we can prove the result by a standard approximation. □

APPENDIX A: PROOF OF LEMMA 4.4

We have

$$\prod_{i=0}^{2n-1}(1+Mh_i)=\left(\prod_{i=0}^{n-1}(1+Mh_i)\right)\left(\prod_{i=n}^{2n-1}(1+Mh_i)\right).$$

First,

$$\prod_{i=n}^{2n-1}(1+Mh_i)\le\left(1+M\frac{T}{n}\right)^n\le C.$$

Moreover, for $0\le i\le n-1$,

$$h_i=t_{i+1}-t_i=Tn^{-ai/n}\left(1-e^{-(a\ln n)/n}\right)\le Tn^{-ai/n}a\frac{\ln n}{n}$$

thanks to the convexity of the exponential function. So

$$\begin{aligned}\prod_{i=0}^{n-1}(1+Mh_i)&\le\prod_{i=0}^{n-1}\left(1+MTan^{-ai/n}\frac{\ln n}{n}\right)\\&=\exp\left(\sum_{i=0}^{n-1}\ln\left(1+MTan^{-ai/n}\frac{\ln n}{n}\right)\right)\\&\le\exp\left(\sum_{i=0}^{n-1}MTa(n^{-a/n})^i\frac{\ln n}{n}\right)\\&\le\exp\left(MTa\frac{\ln n}{n}\left(\frac{1-(1/n^a)}{1-(1/n^{(a/n)})}\right)\right)\\&\le\exp\left(MTa\frac{\ln n}{n}\frac{n^{a/n}}{n^{a/n}-1}\right).\end{aligned}$$

But,

$$\frac{\ln n}{n}\frac{n^{a/n}}{n^{a/n}-1}\sim\frac{\ln n}{n}\frac{1}{(a\ln n)/n}\sim\frac{1}{a},$$

when $n\to+\infty$. Thus, we have shown the result.

APPENDIX B: PROOF OF LEMMA 4.5

Thanks to Lemma 4.4, we have

$$\begin{aligned}\frac{\prod_{i=0}^{n-1}(1+M_1h_i+M_2h_i/(T-t_{i+1}))}{\prod_{i=0}^{n-1}(1+M_2h_i/(T-t_{i+1}))}&=\prod_{i=0}^{n-1}\left(1+\frac{M_1}{1+M_2h_i/(T-t_{i+1})}h_i\right)\\&\le\prod_{i=0}^{n-1}(1+M_1h_i)\le C.\end{aligned}$$

So we just have to show that

$$\prod_{i=0}^{n-1}\left(1+M_2\frac{h_i}{T-t_{i+1}}\right)\le Cn^{aM_2}.$$

But

$$1+M_2\frac{h_i}{T-t_{i+1}}=1+M_2(n^{a/n}-1).$$

So

$$\begin{aligned}\prod_{i=0}^{n-1}\left(1+M_2\frac{h_i}{T-t_{i+1}}\right)&=\left(1+M_2(n^{a/n}-1)\right)^n\\&=\exp\left(n\ln\left(1+aM_2\frac{\ln n}{n}+O\left(\frac{\ln^2 n}{n^2}\right)\right)\right)\\&=\exp\left(aM_2\ln n+O\left(\frac{\ln^2 n}{n}\right)\right)\sim n^{aM_2},\end{aligned}$$

when $n\to+\infty$. Thus, we have shown the result.

APPENDIX C: PROOF OF PROPOSITION 4.19

We will prove this proposition as the authors of [9] do for their Proposition 4.2. In this proof we omit the superscript N,ε for u, Y and Z to be more readable. Let $x_0,x_0'\in\mathbb{R}^d$ and $t_0,t_0'\in[0,T]$. By an argument of symmetry we are allowed to suppose that $t_0\le t_0'$. We have

$$|u(t_0,x_0)-u(t_0',x_0')|\le|u(t_0,x_0)-u(t_0,x_0')|+|u(t_0,x_0')-u(t_0',x_0')|.$$

Let us begin with the first term. We will use a classical argument of linearization:

$$\begin{aligned}Y_t^{t_0,x_0}-Y_t^{t_0,x_0'}&=g_N(X_T^{t_0,x_0})-g_N(X_T^{t_0,x_0'})\\&\quad+\int_t^T\alpha_s(X_s^{t_0,x_0}-X_s^{t_0,x_0'})+\beta_s(Y_s^{t_0,x_0}-Y_s^{t_0,x_0'})\,ds\\&\quad-\int_t^T(Z_s^{t_0,x_0}-Z_s^{t_0,x_0'})\,d\tilde{W}_s\end{aligned}$$

with

$$\alpha_s:=\frac{f^\varepsilon(s,X_s^{t_0,x_0},Y_s^{t_0,x_0'},Z_s^{t_0,x_0'})-f^\varepsilon(s,X_s^{t_0,x_0'},Y_s^{t_0,x_0'},Z_s^{t_0,x_0'})}{X_s^{t_0,x_0}-X_s^{t_0,x_0'}},$$

if $X_s^{t_0,x_0}-X_s^{t_0,x_0'}\neq0$ and $\alpha_s=0$ elsewhere,

$$\beta_s:=\frac{f^\varepsilon(s,X_s^{t_0,x_0},Y_s^{t_0,x_0},Z_s^{t_0,x_0'})-f^\varepsilon(s,X_s^{t_0,x_0},Y_s^{t_0,x_0'},Z_s^{t_0,x_0'})}{Y_s^{t_0,x_0}-Y_s^{t_0,x_0'}},$$

if $X_s^{t_0,x_0} - X_s^{t_0,x_0'} \neq 0$ and $\beta_s = 0$ elsewhere,

$$\gamma_s := \frac{f^\varepsilon(s, X_s^{t_0,x_0}, Y_s^{t_0,x_0}, Z_s^{t_0,x_0}) - f^\varepsilon(s, X_s^{t_0,x_0}, Y_s^{t_0,x_0}, Z_s^{t_0,x_0'})}{|Z_s^{t_0,x_0} - Z_s^{t_0,x_0'}|^2} \times {}^t(Z_s^{t_0,x_0} - Z_s^{t_0,x_0'}),$$

if $Z_s^{t_0,x_0} - Z_s^{t_0,x_0'} \neq 0$ and $\gamma_s = 0$ elsewhere and $d\tilde{W}_s := dW_s - \gamma_s\,ds$. By a BMO argument, there exists a probability $\mathbb{Q}$ under which $\tilde{W}$ is a Brownian motion. Then we apply a classical transformation to obtain

$$\begin{aligned}&\mathbb{E}^{\mathbb{Q}}[e^{\int_{t_0}^t \beta_s\,ds}(Y_t^{t_0,x_0} - Y_t^{t_0,x_0'})]\\ &\quad= \mathbb{E}^{\mathbb{Q}}\Big[e^{\int_{t_0}^T \beta_s\,ds}\big(g_N(X_T^{t_0,x_0}) - g_N(X_T^{t_0,x_0'})\big)\\ &\qquad\quad+ \int_{t_0}^T \alpha_s e^{\int_{t_0}^s \beta_u\,du}(X_s^{t_0,x_0} - X_s^{t_0,x_0'})\,ds\Big]\end{aligned}$$

and

$$\begin{aligned}&|u(t_0,x_0) - u(t_0,x_0')|\\ &\quad\le C\bigg(\mathbb{E}^{\mathbb{Q}}[\omega(|X_T^{t_0,x_0} - X_T^{t_0,x_0'}|)] + \int_{t_0}^T \mathbb{E}^{\mathbb{Q}}[|X_s^{t_0,x_0} - X_s^{t_0,x_0'}|]\,ds\bigg)\end{aligned}$$

with ω a modulus of continuity of g that is also a modulus of continuity for g_N. We are allowed to suppose that ω is concave; indeed, there exist two positive constants a and b such that $\omega(x) \le ax + b$, then the concave hull of $x \mapsto \omega(x) \vee (\mathbb{1}_{x\ge 1}(ax + b))$ is also a modulus of continuity of g. So Jensen's inequality gives us

$$\begin{aligned}&|u(t_0,x_0) - u(t_0,x_0')|\\ &\quad\le C\bigg(\omega(\mathbb{E}^{\mathbb{Q}}[|X_T^{t_0,x_0} - X_T^{t_0,x_0'}|]) + \int_{t_0}^T \mathbb{E}^{\mathbb{Q}}[|X_s^{t_0,x_0} - X_s^{t_0,x_0'}|]\,ds\bigg).\end{aligned}$$

By using the fact that b is bounded we can prove the following proposition exactly as authors of [9] do for their Proposition 4.7.

PROPOSITION C.1. *$\exists C > 0$ that does not depend on N and ε such that $\forall t, t' \in [0,T]$, $\forall x, x' \in \mathbb{R}^d$, $\forall s \in [0,T]$,*

$$\mathbb{E}^{\mathbb{Q}}[|X_s^{t,x} - X_s^{t',x'}|] \le C(|x - x'| + |t - t'|^{1/2}).$$

Then,

$$|u(t_0,x_0) - u(t_0,x_0')| \le C\big(\omega(|x_0 - x_0'|) + |x_0 - x_0'|\big).$$

Now we will study the second term,

$$|u(t_0, x_0') - u(t_0', x_0')| = |Y_{t_0}^{t_0, x_0'} - Y_{t_0'}^{t_0', x_0'}|$$
$$\leq |Y_{t_0}^{t_0, x_0'} - Y_{t_0}^{t_0', x_0'}| + |Y_{t_0}^{t_0', x_0'} - Y_{t_0'}^{t_0', x_0'}|.$$

First,

$$|Y_{t_0}^{t_0', x_0'} - Y_{t_0'}^{t_0', x_0'}| \leq \left| \int_{t_0}^{t_0'} f(s, x_0', Y_s^{t_0', x_0'}, 0)\, ds \right| \leq C|t_0 - t_0'|.$$

Moreover, as for the first term we have

$$\mathbb{E}^{\mathbb{Q}}[e^{\int_{t_0}^t \beta_s\, ds}(Y_t^{t_0, x_0'} - Y_t^{t_0', x_0'})]$$
$$= \mathbb{E}^{\mathbb{Q}}\Big[e^{\int_{t_0}^T \beta_s\, ds}\big(g_N(X_T^{t_0, x_0'}) - g_N(X_T^{t_0', x_0'})\big)$$
$$+ \int_{t_0}^T \alpha_s e^{\int_{t_0}^s \beta_u\, du}(X_s^{t_0, x_0'} - X_s^{t_0', x_0'})\, ds \Big]$$

and

$$|Y_{t_0}^{t_0, x_0'} - Y_{t_0'}^{t_0', x_0'}| \leq C\big(\omega(|t_0 - t_0'|^{1/2}) + |t_0 - t_0'|^{1/2}\big).$$

Finally,

$$|u(t_0, x_0') - u(t_0', x_0')| \leq C\big(\omega(|t_0 - t_0'|^{1/2}) + |t_0 - t_0'|^{1/2}\big)$$

and

$$|u(t_0, x_0) - u(t_0', x_0')|$$
$$\leq C\big(\omega(|x_0 - x_0'|) + \omega(|t_0 - t_0'|^{1/2}) + |x_0 - x_0'| + |t_0 - t_0'|^{1/2}\big).$$

So u is uniformly continuous on $[0, T] \times \mathbb{R}^d$ and this function has a modulus of continuity that does not depend on N and ε. Moreover, we are allowed to suppose that this modulus of continuity is concave.

Acknowledgments. The author would like to thank his Ph.D. advisers, Philippe Briand and Ying Hu, an anonymous referee and the Associate Editor for their careful reading and helpful comments.

REFERENCES

[1] BOUCHARD, B. and TOUZI, N. (2004). Discrete-time approximation and Monte-Carlo simulation of backward stochastic differential equations. *Stochastic Process. Appl.* **111** 175–206. MR2056536

[2] BRIAND, P. and CONFORTOLA, F. (2008). BSDEs with stochastic Lipschitz condition and quadratic PDEs in Hilbert spaces. *Stochastic Process. Appl.* **118** 818–838. MR2411522

[3] BRIAND, P. and HU, Y. (2006). BSDE with quadratic growth and unbounded terminal value. *Probab. Theory Related Fields* **136** 604–618. MR2257138

[4] BRIAND, P. and HU, Y. (2008). Quadratic BSDEs with convex generators and unbounded terminal conditions. *Probab. Theory Related Fields* **141** 543–567. MR2391164

[5] CHERIDITO, P. and STADJE, M. (2010). BSΔEs and BSDEs with non-Lipschitz drivers: Comparison, convergence and robustness. Available at arXiv:1002.0175v1.

[6] CHEVANCE, D. (1997). Numerical methods for backward stochastic differential equations. In *Numerical Methods in Finance*. *Publ. Newton Inst*. 232–244. Cambridge Univ. Press, Cambridge. MR1470517

[7] DELARUE, F. and MENOZZI, S. (2006). A forward–backward stochastic algorithm for quasi-linear PDEs. *Ann. Appl. Probab.* **16** 140–184. MR2209339

[8] DELBAEN, F., HU, Y. and BAO, X. (2011). Backward SDEs with superquadratic growth. *Probab. Theory Related Fields* **150** 145–192.

[9] DELBAEN, F., HU, Y. and RICHOU, A. (2011). On the uniqueness of solutions to quadratic BSDEs with convex generators and unbounded terminal conditions. *Ann. Inst. H. Poincaré Probab. Statist.* To appear.

[10] FUHRMAN, M. and TESSITORE, G. (2002). The Bismut–Elworthy formula for backward SDEs and applications to nonlinear Kolmogorov equations and control in infinite dimensional spaces. *Stochastics Stochastics Rep*. **74** 429–464. MR1940495

[11] GOBET, E., LEMOR, J.-P. and WARIN, X. (2005). A regression-based Monte Carlo method to solve backward stochastic differential equations. *Ann. Appl. Probab.* **15** 2172–2202. MR2152657

[12] GOBET, E. and MAKHLOUF, A. (2010). $\mathbf{L}_2$-time regularity of BSDEs with irregular terminal functions. *Stochastic Process. Appl.* **120** 1105–1132. MR2639740

[13] HU, Y., IMKELLER, P. and MÜLLER, M. (2005). Utility maximization in incomplete markets. *Ann. Appl. Probab.* **15** 1691–1712. MR2152241

[14] IMKELLER, P. and DOS REIS, G. (2010). Path regularity and explicit convergence rate for BSDE with truncated quadratic growth. *Stochastic Process. Appl.* **120** 348–379. MR2584898

[15] IMKELLER, P., DOS REIS, G. and ZHANG, J. (2010). Results on numerics for FBSDE with drivers of quadratic growth. In *Contemporary Quantitative Finance* (A. Chiarella and C. Novikov, eds.). Springer, Berlin.

[16] KAZAMAKI, N. (1994). *Continuous Exponential Martingales and BMO*. *Lecture Notes in Math*. **1579**. Springer, Berlin. MR1299529

[17] KOBYLANSKI, M. (2000). Backward stochastic differential equations and partial differential equations with quadratic growth. *Ann. Probab.* **28** 558–602. MR1782267

[18] LEPELTIER, J. P. and SAN MARTÍN, J. (1998). Existence for BSDE with superlinear-quadratic coefficient. *Stochastics Stochastics Rep*. **63** 227–240. MR1658083

[19] ZHANG, J. (2004). A numerical scheme for BSDEs. *Ann. Appl. Probab.* **14** 459–488. MR2023027

[20] ZHANG, J. (2005). Representation of solutions to BSDEs associated with a degenerate FSDE. *Ann. Appl. Probab.* **15** 1798–1831. MR2152246

[21] ZHANG, J. (2001). Some fine properties of BSDE. Ph.D. thesis, Purdue Univ.

IRMAR
UNIVERSITÉ DE RENNES 1
CAMPUS DE BEAULIEU
35042 RENNES CEDEX
FRANCE
E-MAIL: adrien.richou@univ-rennes1.fr