Mean square rate of convergence for random walk approximation of forward-backward SDEs

Christel Geiss, Céline Labart, Antti Luoto

Applied Probability Trust (5 March 2020)

MEAN SQUARE RATE OF CONVERGENCE FOR RANDOM WALK APPROXIMATION OF FORWARD-BACKWARD SDES

CHRISTEL GEISS,* *Department of Mathematics and Statistics, University of Jyvaskyla*

CÉLINE LABART,** *Univ. Grenoble Alpes, Univ. Savoie Mont Blanc, CNRS, LAMA*

ANTTI LUOTO,* *Department of Mathematics and Statistics, University of Jyvaskyla*

Abstract

Let (Y, Z) denote the solution to a forward-backward SDE. If one constructs a random walk B^n from the underlying Brownian motion B by Skorohod embedding, one can show L_2-convergence of the corresponding solutions (Y^n, Z^n) to (Y, Z). We estimate the rate of convergence in dependence of smoothness properties, especially for a terminal condition function in $C^{2,\alpha}$. The proof relies on an approximative representation of Z^n and uses the concept of discretized Malliavin calculus. Moreover, we use growth and smoothness properties of the PDE associated to the FBSDE as well as of the finite difference equations associated to the approximating stochastic equations. We derive these properties by probabilistic methods.

Keywords: Backward stochastic differential equations; approximation scheme; finite difference equation; convergence rate; random walk approximation

2010 Mathematics Subject Classification: Primary 60H10; 60H35; 60G50

Secondary 60H30

1. Introduction

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space carrying the standard Brownian motion $B = (B_t)_{t\geq 0}$ and assume that $(\mathcal{F}_t)_{t\geq 0}$ is the augmented natural filtration. Let (Y, Z) be the solution of the forward-backward SDE (FBSDE)

$$\begin{aligned} X_s &= x + \int_0^s b(r, X_r)dr + \int_0^s \sigma(r, X_r)dB_r, \\ Y_s &= g(X_T) + \int_s^T f(r, X_r, Y_r, Z_r)dr - \int_s^T Z_r dB_r, \qquad 0 \leq s \leq T. \end{aligned} \tag{1}$$

* Postal address: P.O.Box 35 (MaD) FI-40014 University of Jyvaskyla, Finland

** Postal address: Univ. Grenoble Alpes, Univ. Savoie Mont Blanc, CNRS, LAMA, 73000 Chambéry, France

Let (Y^n, Z^n) be the solution of the FBSDE if the Brownian motion B is replaced by a scaled random walk B^n given by

$$B_t^n = \sqrt{h} \sum_{i=1}^{[t/h]} \varepsilon_i, \qquad 0 \le t \le T, \tag{2}$$

where $h = \frac{T}{n}$ and $(\varepsilon_i)_{i=1,2,\dots}$ is a sequence of i.i.d. Rademacher random variables. Then (Y^n, Z^n) solves the discretized FBSDE

$$\begin{aligned} X_s^n &= x + \int_{(0,s]} b(r, X_{r-}^n) d[B^n]_r + \int_{(0,s]} \sigma(r, X_{r-}^n) dB_r^n, \\ Y_s^n &= g(X_T^n) + \int_{(s,T]} f(r, X_{r-}^n Y_{r-}^n, Z_{r-}^n) d[B^n]_r - \int_{(s,T]} Z_{r-}^n dB_r^n, \quad 0 \le s \le T. \end{aligned} \tag{3}$$

The approximation of BSDEs using random walk has been investigated by many authors, also numerically (see, for example, [5], [25], [29], [31], [32], [33], [16]). In 2001, Briand et al. [5] have shown weak convergence of (Y^n, Z^n) to (Y, Z) for a Lipschitz continuous generator f and a terminal condition in L_2. The rate of convergence of this method remained an open problem. Bouchard and Touzi in [7] and Zhang in [41] proposed instead of random walk an approach based on the dynamic programming equation, for which they established a rate of convergence. But this approach involves conditional expectations. Various methods to approximate these conditional expectations have been developed ([23], [17], [14]). Also forward methods have been introduced to approximate (1): a branching diffusion method ([26]), a multilevel Picard approximation ([40]) and Wiener chaos expansion ([6]). Many extensions of (1) have been considered, among them schemes for reflected BSDEs ([3], [13]), high order schemes ([10], [9]), fully-coupled BSDEs ([18], [8]), quadratic BSDEs ([12]), BSDEs with jumps ([22]) and McKean-Vlasov BSDEs ([1], [15], [11]).

The aim of this paper is to study the rate of the L_2-approximation of (Y_t^n, Z_t^n) to (Y_t, Z_t) when X satisfies (1). For this, we generate the random walk B^n by Skorohod embedding from the Brownian motion B. In this case the L_p-convergence of B^n to B is of order $h^{\frac{1}{4}}$ for any $p > 0$. The special case $X = B$ has already been studied in [21], assuming a locally α-Hölder continuous terminal function g and a Lipschitz continuous generator. An estimate for the rate of convergence was obtained which is of order $h^{\frac{\alpha}{4}}$ for the L_2-norm of $Y_t^n - Y_t$, and of order $\frac{h^{\frac{\alpha}{4}}}{\sqrt{T-t}}$ for the L_2-norm of $Z_t^n - Z_t$.

In the present paper, where we assume that X is a solution of the SDE in (1), rather strong conditions on the smoothness and boundedness on f and g and also on b and σ are needed. In Theorem 3.1, the main result of the paper, we show that the convergence rate for (Y_t^n, Z_t^n) to

(Y_t, Z_t) in L_2 is of order $h^{\frac{1}{4}\wedge\frac{\alpha}{2}}$ provided that g'' is locally α-Hölder continuous. To the best of our knowledge, these are the first cases a convergence rate for the approximation of forward-backward SDEs using random walk has been obtained.

Remark 1.1. For the diffusion setting – in contrast to the case $X = B$ – we can derive the convergence rate for (Y_t^n, Z_t^n) to (Y_t, Z_t) in L_2 only under strong smoothness conditions on the coefficients which include also that g'' is locally α-Hölder continuous (see Assumption 2.3 below). These requirements appear to be necessary. This becomes visible in Subsection 2.2.2 where we introduce a discretized Malliavin weight to obtain a representation $\hat{Z}^n$ for Z^n. While it holds that $\hat{Z}^n = Z^n$ when $X = B$, in our case $\hat{Z}^n$ does not coincide with Z^n. However, one can show that the difference $\hat{Z}_t^n - Z_t^n$ converges to 0 in L_2 as $n \to \infty$ using a Hölder continuity property (see (62) in Remark 4.1) for the space derivative of the generator in (3). For this Hölder continuity property to hold one needs enough smoothness in space from the solution u^n to the finite difference equation associated to the discretized FBSDE (3). Provided that Assumption 2.3 holds we show the smoothness properties for u^n in Proposition 4.2 applying methods known for Lévy driven BSDEs.

The paper is organized as follows: Section 2 contains the setting, main assumptions and the approximative representation $\hat{Z}^n$ of Z^n. Our main results about the approximation rate for the case of no generator (i.e. $f = 0$) and for the general case are in Section 3. One can see that in contrast to what is known for time discretization schemes, for random walk schemes the Lipschitz generator seems to cause more difficulties than the terminal condition: while in the case $f = 0$ we need that g' is locally α-Hölder continuous, in the case $f \neq 0$ is this property is required for g''. In Section 4 we recall some needed facts about Malliavin weights, about the regularity of solutions to BSDEs and properties of the associated PDEs. Finally, we sketch how to prove growth and smoothness properties of solutions to the finite difference equation associated to the discretized FBSDE. Section 5 contains technical results which mainly arise from the fact that the construction of the random walk by Skorohod embedding forces us to compare our processes on different 'time lines', one coming from the stopping times of the Skorohod embedding, and the other one is ruled by the equidistant deterministic times due to the quadratic variation process $[B^n]$.

2. Preliminaries

2.1. The SDE and its approximation scheme

We introduce

$$X_t = x + \int_0^t b(s, X_s)ds + \int_0^t \sigma(s, X_s)dB_s, \qquad 0 \le t \le T$$

and its discretized counterpart

$$X^n_{t_k} = x + h\sum_{j=1}^{k} b(t_j, X^n_{t_{j-1}}) + \sqrt{h}\sum_{j=1}^{k} \sigma(t_j, X^n_{t_{j-1}})\varepsilon_j, \qquad t_j := j\tfrac{T}{n},\ j = 0, ..., n, \tag{4}$$

where $(\varepsilon_i)_{i=1,2,...}$ is a sequence of i.i.d. Rademacher random variables. Letting $\mathcal{G}_k := \sigma(\varepsilon_i : 1 \le i \le k)$ with $\mathcal{G}_0 := \{\emptyset, \Omega\}$, it follows that the associated discrete-time random walk $(B^n_{t_k})_{k=0}^n$ is $(\mathcal{G}_k)_{k=0}^n$-adapted. Recall (2) and $h = \frac{T}{n}$. If we extend the sequence $(X^n_{t_k})_{k\ge0}$ to a process in continuous time by defining $X^n_t := X^n_{t_k}$ for $t \in [t_k, t_{k+1})$, it is the solution of the forward SDE (3).

We formulate our first assumptions. Assumption 2.1 (ii) will be not used explicitely for our estimates but it is required for Theorem 4.1 below.

Assumption 2.1.

(i) $b, \sigma \in C_b^{0,2}([0,T] \times \mathbb{R})$, *in the sense that the derivatives of order* $k = 0, 1, 2$ *w.r.t. the space variable are continuous and bounded on* $[0,T] \times \mathbb{R}$,

(ii) the first and second derivatives of b *and* σ *w.r.t. the space variable are assumed to be* γ*-Hölder continuous (for some* $\gamma \in (0,1]$, *w.r.t. the parabolic metric* $d((t,x),(\bar t,\bar x)) = (|t-\bar t|+|x-\bar x|^2)^{\frac12}$*) on all compact subsets of* $[0,T] \times \mathbb{R}$.

(iii) b, σ *are* $\frac12$*-Hölder continuous in time, uniformly in space,*

(iv) $\sigma(t,x) \ge \delta > 0$ *for all* (t,x).

Assumption 2.2.

(i) g *is locally Hölder continuous with order* $\alpha \in (0,1]$ *and polynomially bounded* ($p_0 \ge 0, C_g > 0$) *in the following sense*

$$\forall (x,\bar x) \in \mathbb{R}^2, \quad |g(x) - g(\bar x)| \le C_g(1 + |x|^{p_0} + |\bar x|^{p_0})|x - \bar x|^\alpha. \tag{5}$$

(ii) The function $[0,T] \times \mathbb{R}^3 : (t,x,y,z) \mapsto f(t,x,y,z)$ *satisfies*

$$|f(t,x,y,z) - f(\bar t,\bar x,\bar y,\bar z)| \le L_f(\sqrt{t-\bar t} + |x-\bar x| + |y-\bar y| + |z-\bar z|). \tag{6}$$

Notice that (5) implies

$$|g(x)| \leq K(1+|x|^{p_0+1}) =: \Psi(x), \quad x \in \mathbb{R}, \tag{7}$$

for some $K > 0$. From the continuity of f we conclude that

$$K_f := \sup_{0\leq t\leq T} |f(t,0,0,0)| < \infty.$$

Notation:

- $\|\cdot\|_p := \|\cdot\|_{L_p(\mathbb{P})}$ for $p \geq 1$ and for $p = 2$ simply $\|\cdot\|$.
- If a is a function, $C(a)$ represents a generic constant which depends on a and possibly also on its derivatives.
- $\mathbb{E}_{0,x} := \mathbb{E}(\cdot|X_0 = x)$.
- Let ϕ be a $C^{0,1}([0,T]\times\mathbb{R})$ function. ϕ_x denotes $\partial_x\phi$, the partial derivative of ϕ w.r.t. x.

2.2. The FBSDE and its approximation scheme

Recall the FBSDE (1) and its approximation (3). The backward equation in (3) can equivalently be written in the form

$$Y^n_{t_k} = g(X^n_T) + h\sum_{m=k}^{n-1} f(t_{m+1}, X^n_{t_m}, Y^n_{t_m}, Z^n_{t_m}) - \sqrt{h}\sum_{m=k}^{n-1} Z^n_{t_m}\varepsilon_{m+1}, \quad 0 \leq k \leq n, \tag{8}$$

if one puts $X^n_r := X^n_{t_m}$, $Y^n_r := Y^n_{t_m}$ and $Z^n_r := Z^n_{t_m}$ for $r \in [t_m, t_{m+1})$.

Remark 2.1. Equations (3) and (8) do not contain any orthogonal part to the random walk B^n since we are in a special case where the orthogonal part is zero. Indeed, for $(\varepsilon_k)_{k=1,\cdots,n}$ following the Rademacher law assume ($\mathcal{G}_k := \sigma(\varepsilon_i, i = 1,\cdots,k)$) as filtration, and let for the $\mathcal{G}_n$-measurable random variable $F(\varepsilon_1,...,\varepsilon_n)$ hold the representation

$$F(\varepsilon_1,...,\varepsilon_n) = c + \sum_{m=1}^{n} h_m\varepsilon_m + N_n,$$

where $(h_m)_{m=1}^n$ is predictable and $(N_m)_{m=1}^n$ a martingale orthogonal to $(B^n_{t_m})_{m=1}^n$ given by $B^n_{t_m} = \sqrt{h}(\varepsilon_1 + ... + \varepsilon_m)$. By definition, orthogonality of the martingales N and B^n means that their product is a martingale, i.e. we have

$$\mathbb{E}[N_{k+1}B^n_{t_{k+1}}|\mathcal{G}_k] = N_kB^n_{t_k},$$

and since $N_kB^n_{t_k} = \mathbb{E}[N_{k+1}B^n_{t_k}|\mathcal{G}_k]$, this implies especially that $\mathbb{E}N_{k+1}\varepsilon_{k+1} = 0$. Assume N_{k+1} is given by $N_{k+1} = H(\varepsilon_1,...,\varepsilon_{k+1})$. Then $0 = \mathbb{E}H(\varepsilon_1,...,\varepsilon_{k+1})\varepsilon_{k+1} = \frac{1}{2}[H(\varepsilon_1,...,\varepsilon_k,1) - H(\varepsilon_1,...,\varepsilon_k,-1)]$

implying that N_{k+1} is $\mathcal{G}_k$-measurable since $H(\varepsilon_1,...,\varepsilon_k,1) = H(\varepsilon_1,...,\varepsilon_k,-1)$, and therefore the martingale $(N_m)_{m=1}^n$ is identically zero. (See also [5, page 3] or [34, Proposition1.7.5].)

One can derive an equation for $Z^n = (Z^n_{t_k})_{k=0}^{n-1}$ if one multiplies (8) by ε_{k+1} and takes the conditional expectation w.r.t. $\mathcal{G}_k$, so that

$$Z^n_{t_k} = \frac{\mathbb{E}^{\mathcal{G}}_k\left(g(X^n_T)\varepsilon_{k+1}\right)}{\sqrt{h}} + \mathbb{E}^{\mathcal{G}}_k\left(\sqrt{h}\sum_{m=k+1}^{n-1} f(t_{m+1}, X^n_{t_m}, Y^n_{t_m}, Z^n_{t_m})\varepsilon_{k+1}\right), \; 0 \le k \le n-1, \quad (9)$$

where $\mathbb{E}^{\mathcal{G}}_k := \mathbb{E}(\cdot|\mathcal{G}_k)$.

Remark 2.2. For n large enough, the BSDE (3) has a unique solution (Y^n, Z^n) (see [36, Proposition 1.2]), and $(Y^n_{t_k}, Z^n_{t_k})_{k=0}^{n-1}$ is adapted to the filtration $(\mathcal{G}_k)_{k=0}^{n-1}$.

2.2.1. *Representation for Z* We will use the following representation for Z, due to Ma and Zhang (see [30, Theorem 4.2])

$$Z_t = \mathbb{E}_t\left(g(X_T)N^t_T + \int_t^T f(s, X_s, Y_s, Z_s)N^t_s ds\right)\sigma(t, X_t), \quad 0 \le t \le T \quad (10)$$

where $\mathbb{E}_t := \mathbb{E}(\cdot|\mathcal{F}_t)$, and for all $s \in (t, T]$, we have (cf. Lemma 4.1)

$$N^t_s = \frac{1}{s-t}\int_t^s \frac{\nabla X_r}{\sigma(r, X_r)\nabla X_t} dB_r, \quad (11)$$

where $\nabla X = (\nabla X_s)_{s\in[0,T]}$ is the variational process i.e. it solves

$$\nabla X_s = 1 + \int_0^s b_x(r, X_r)\nabla X_r dr + \int_0^s \sigma_x(r, X_r)\nabla X_r dB_r, \quad (12)$$

with $(X_s)_{s\in[0,T]}$ given in (1).

Remark 2.3. In the following we will assume that g'' exists. In such a case we have the following representation for Z:

$$Z_t = \mathbb{E}_t\left(g'(X_T)\nabla X_T + \int_t^T f(s, X_s, Y_s, Z_s)N^t_s ds\right)\sigma(t, X_t), \quad 0 \le t \le T. \quad (13)$$

2.2.2. *Approximation for Z^n* In this section we state the discrete counterpart to (10), which, in the general case of a forward process X, does not coincide with Z^n (given by (9)). In contrast to the continuous-time case, where the variational process and the Malliavin derivative are connected by $\frac{\nabla X_t}{\nabla X_s} = \frac{D_s X_t}{\sigma(s,X_s)}$ $(s \le t)$, we can not expect equality for the corresponding expressions if we use the discretized version of the processes $(\nabla X_t)_t$ and $(D_s X_t)_{s\le t}$ introduced in (15). This counterpart $\hat{Z}^n$ to Z is a key tool in the proof of the convergence of Z^n to Z. As we will see in the proof of

Theorem 3.1, the study of $\|Z^n_{t_k} - Z_{t_k}\|$ goes through the study of $\|Z^n_{t_k} - \hat{Z}^n_{t_k}\|$ and $\|\hat{Z}^n_{t_k} - Z_{t_k}\|$.

Before defining the discretized version of $(\nabla X_t)_t$ and $(D_s X_t)_{s\le t}$, we shortly introduce the discretized Malliavin derivative and refer the reader to [4] for more information on this topic.

Definition 2.1. (Definition of $T_{m,+}$, $T_{m,-}$ and $\mathcal{D}^n_m$.) *For any function $F : \{-1,1\}^n \to \mathbb{R}$, the mappings $T_{m,+}$ and $T_{m,-}$ are defined by*

$$T_{m,\pm}F(\varepsilon_1,\dots,\varepsilon_n) := F(\varepsilon_1,\dots,\varepsilon_{m-1},\pm 1,\varepsilon_{m+1},\dots,\varepsilon_n), \qquad 1\le m\le n.$$

For any $\xi = F(\varepsilon_1,\dots,\varepsilon_n)$, the discretized Malliavin derivative is defined by

$$\mathcal{D}^n_m\xi := \frac{\mathbb{E}[\xi\varepsilon_m|\sigma((\varepsilon_l)_{l\in\{1,\dots,n\}\setminus\{m\}})]}{\sqrt{h}} = \frac{T_{m,+}\xi - T_{m,-}\xi}{2\sqrt{h}}, \quad 1\le m\le n. \tag{14}$$

Definition 2.2. (Definition of $\phi^{(k,l)}_x$.) *Let ϕ be a $C^{0,1}([0,T]\times\mathbb{R})$ function. We denote*

$$\phi^{(k,l)}_x := \frac{\mathcal{D}^n_k\phi(t_l, X^n_{t_{l-1}})}{\mathcal{D}^n_k X^n_{t_{l-1}}} := \int_0^1 \phi_x(t_l, \vartheta T_{k,+}\, X^n_{t_{l-1}} + (1-\vartheta)T_{k,-}\, X^n_{t_{l-1}})d\vartheta.$$

If $\mathcal{D}^n_k X^n_{t_{\ell-1}} \neq 0$ the second $'{:=}'$ holds as an identity.

We are now able to define the discretized version of $(\nabla X_t)_t$ and $(D_s X_t)_{s\le t}$.

Definition 2.3. (Discretized processes $(\nabla X^{n,t_k,x}_{t_m})_{m\in\{k,\dots,n\}}$ and $(\mathcal{D}^n_k X^n_{t_m})_{m\in\{k,\dots,n\}}$.) *For all m in $\{k,\dots,n\}$ we define*

$$\nabla X^{n,t_k,x}_{t_m} = 1 + h\sum_{l=k+1}^{m} b_x(t_l, X^{n,t_k,x}_{t_{l-1}})\nabla X^{n,t_k,x}_{t_{l-1}} + \sqrt{h}\sum_{l=k+1}^{m}\sigma_x(t_l, X^{n,t_k,x}_{t_{l-1}})\nabla X^{n,t_k,x}_{t_{l-1}}\varepsilon_l, \quad 0\le k\le n,$$

$$\mathcal{D}^n_k X^n_{t_m} = \sigma(t_k, X^n_{t_{k-1}}) + h\sum_{l=k+1}^{m} b^{(k,l)}_x \mathcal{D}^n_k X^n_{t_{l-1}} + \sqrt{h}\sum_{l=k+1}^{m}\sigma^{(k,l)}_x(\mathcal{D}^n_k X^n_{t_{l-1}})\varepsilon_l, \quad 0<k\le n. \tag{15}$$

Remark 2.4. (i) Although $\nabla X^{n,t_k,X^n_{t_k}}_{t_m}$ is not equal to $\frac{\mathcal{D}^n_{k+1}X^n_{t_m}}{\sigma(t_{k+1},X^n_{t_k})}$, we can show that the difference of these terms converges in L_p (see Lemma 5.4).

(ii) With the notation introduced above, (9) rewrites to

$$Z^n_{t_k} = \mathbb{E}^{\mathcal{G}}_k\left(\mathcal{D}^n_{k+1}g(X^n_T)\right) + \mathbb{E}^{\mathcal{G}}_k\left(h\sum_{m=k+1}^{n-1}\mathcal{D}^n_{k+1}f(t_{m+1}, X^n_{t_m}, Y^n_{t_m}, Z^n_{t_m})\right). \tag{16}$$

In order to define the discrete counterpart to (10), we first define the discrete counterpart to $(N^t_s)_{s\in[t,T]}$ given in (11):

$$N^{n,t_k}_{t_\ell} := \sqrt{h}\sum_{m=k+1}^{\ell}\frac{\nabla X^{n,t_k,X^n_{t_k}}_{t_{m-1}}}{\sigma(t_m, X^n_{t_{m-1}})}\frac{\varepsilon_m}{t_\ell - t_k}, \quad k<\ell\le n. \tag{17}$$

Notice that there is some constant $\widehat{\kappa}_2 > 0$ depending on b, σ, T, δ such that

$$\left(\mathbb{E}_k^{\mathcal{G}} |N_{t_\ell}^{n,t_k}|^2\right)^{\frac{1}{2}} \leq \frac{\widehat{\kappa}_2}{(t_\ell - t_k)^{\frac{1}{2}}}, \quad 0 \leq k < \ell \leq n. \tag{18}$$

Definition 2.4. (Discrete counterpart to (13).) *Let the process* $\hat{Z}^n = (\hat{Z}_{t_k}^n)_{k=0}^{n-1}$ *be defined by*

$$\hat{Z}_{t_k}^n := \mathbb{E}_k^{\mathcal{G}} \left(\mathcal{D}_{k+1}^n g(X_T^n)\right) + \mathbb{E}_k^{\mathcal{G}} \left(h \sum_{m=k+1}^{n-1} f(t_{m+1}, X_{t_m}^n, Y_{t_m}^n, Z_{t_m}^n) N_{t_m}^{n,t_k} \right) \sigma(t_{k+1}, X_{t_k}^n), \tag{19}$$

Remark 2.5. In (19) We could have used also the approximate expression $\mathbb{E}_k^{\mathcal{G}}(g(X_T^n) N_{t_n}^{n,t_k} \sigma(t_{k+1}, X_{t_k}^n))$, but since we will assume that g'' exists, we work with the correct term.

The study of the convergence $\mathbb{E}_{0,x}^{\mathcal{G}} |Z_{t_k}^n - \hat{Z}_{t_k}^n|^2$ requires stronger assumptions on the coefficients b, σ, f and g.

Assumption 2.3. *Assumptions 2.1 and 2.2 hold. Additionally, we assume that all first and second derivatives w.r.t. the variables* x, y, z *of* $b(t,x), \sigma(t,x)$ *and* $f(t,x,y,z)$ *exist and are bounded Lipschitz functions w.r.t. these variables, uniformly in time. Moreover,* g'' *satisfies* (5).

Proposition 2.1. *If Assumption 2.3 holds, then*

$$\mathbb{E}_{0,x}^{\mathcal{G}} |Z_{t_k}^n - \hat{Z}_{t_k}^n|^2 \leq C_{2.1} \hat{\Psi}^2(x) h^\alpha,$$

where $\mathbb{E}_{0,x}^{\mathcal{G}} := \mathbb{E}^{\mathcal{G}}(\cdot | X_0 = x)$, *the function* $\hat{\Psi}$ *is defined in* (61) *below, and* $C_{2.1}$ *depends on* b, σ, f, g, T, p_0 *and* δ.

Proof. According to [5, Proposition 5.1] one has the representations

$$Y_{t_m}^n = u^n(t_m, X_{t_m}^n), \quad \text{and} \quad Z_{t_m}^n = \mathcal{D}_{m+1}^n u^n(t_{m+1}, X_{t_{m+1}}^n), \tag{20}$$

where u^n is the solution of the finite difference equation (43) with terminal condition $u^n(t_n, x) = g(x)$. Notice that by the definition of $\mathcal{D}_{m+1}^n$ in (14) the expression $\mathcal{D}_{m+1}^n u^n(t_{m+1}, X_{t_{m+1}}^n)$ depends in fact on $X_{t_m}^n$. Hence we can put

$$\begin{aligned} f(t_{m+1}, X_{t_m}^n, Y_{t_m}^n, Z_{t_m}^n) &= f(t_{m+1}, X_{t_m}^n, u^n(t_m, X_{t_m}^n), \mathcal{D}_{m+1}^n u^n(t_{m+1}, X_{t_{m+1}}^n)) \\ &=: F^n(t_{m+1}, X_{t_m}^n). \end{aligned}$$

From (19) and (16) we conclude that (we use $\mathbb{E} := \mathbb{E}^{\mathcal{G}}_{0,x}$ for $\|\cdot\|$)

$$\begin{aligned}
&\|Z^n_{t_k} - \hat{Z}^n_{t_k}\| \\
&= \Bigg\| \mathbb{E}^{\mathcal{G}}_k \Bigg(h \sum_{m=k+1}^{n-1} \mathcal{D}^n_{k+1} f(t_{m+1}, X^n_{t_m}, Y^n_{t_m}, Z^n_{t_m}) \Bigg) \\
&\qquad - \mathbb{E}^{\mathcal{G}}_k \Bigg(h \sum_{m=k+1}^{n-1} f(t_{m+1}, X^n_{t_m}, Y^n_{t_m}, Z^n_{t_m}) N^{n,t_k}_{t_m} \sigma(t_{k+1}, X^n_{t_k}) \Bigg) \Bigg\| \\
&\le \sum_{m=k+1}^{n-1} \frac{h}{m-k} \sum_{\ell=k+1}^{m} \left\| \mathbb{E}^{\mathcal{G}}_k \left[\mathcal{D}^n_{k+1} F^n(t_{m+1}, X^n_{t_m}) - \mathcal{D}^n_{\ell} F^n(t_{m+1}, X^n_{t_m}) \frac{\sigma(t_{k+1}, X^n_{t_k}) \nabla X^{n,t_k,X^n_{t_k}}_{t_{\ell-1}}}{\sigma(t_\ell, X^n_{t_{\ell-1}})} \right] \right\|.
\end{aligned}$$

With the notation introduced in Definition 2.2 applied to F^n,

$$\begin{aligned}
&\left\| \mathcal{D}^n_{k+1} F^n(t_{m+1}, X^n_{t_m}) - \mathcal{D}^n_{\ell} F^n(t_{m+1}, X^n_{t_m}) \frac{\sigma(t_{k+1}, X^n_{t_k}) \nabla X^{n,t_k,X^n_{t_k}}_{t_{\ell-1}}}{\sigma(t_\ell, X^n_{t_{\ell-1}})} \right\| \\
&\quad \le \quad \|(\mathcal{D}^n_{k+1} X^n_{t_m})(F^{n,(k+1,m+1)}_x - F^{n,(\ell,m+1)}_x)\| \\
&\qquad + \left\| F^{n,(\ell,m+1)}_x \left((\mathcal{D}^n_{k+1} X^n_{t_m}) - (\mathcal{D}^n_{\ell} X^n_{t_m}) \frac{\sigma(t_{k+1}, X^n_{t_k}) \nabla X^{n,t_k,X^n_{t_k}}_{t_{\ell-1}}}{\sigma(t_\ell, X^n_{t_{\ell-1}})} \right) \right\| \\
&\quad =: \quad A_1 + A_2.
\end{aligned}$$

For A_1 we use Definition 2.2 again and exploit the fact that

$$x \mapsto F^n_x(t_{m+1}, x) := \partial_x f(t_{m+1}, x, u^n(t_m, x), \mathcal{D}^n_{m+1} u^n(t_{m+1}, X^{n,t_m,x}_{t_{m+1}}))$$

is locally α-Hölder continuous according to (62). By Hlder's inequality and Lemma 5.4 (i) and (iii),

$$\begin{aligned}
A_1 \le \|\mathcal{D}^n_{k+1} X^n_{t_m}\|_4 \int_0^1 &\|F^n_x(t_{m+1}, \vartheta T_{k+1,+} X^n_{t_m} + (1-\vartheta) T_{k+1,-} X^n_{t_m}) \\
&- F^n_x(t_{m+1}, \vartheta T_{\ell,+} X^n_{t_m} + (1-\vartheta) T_{\ell,-} X^n_{t_m})\|_4 d\vartheta \le C(b,\sigma,f,g,T,p_0) \hat{\Psi}(x) h^{\frac{\alpha}{2}}.
\end{aligned}$$

For the estimate of A_2 we notice that by our assumptions the L_4-norm of $F^{n,(\ell,m+1)}_x$ is bounded by $C\Psi^2(x)$, so that it suffices to estimate

$$\begin{aligned}
&\left\| (\mathcal{D}^n_{k+1} X^n_{t_m}) - (\mathcal{D}^n_{\ell} X^n_{t_m}) \frac{\sigma(t_{k+1}, X^n_{t_k}) \nabla X^{n,t_k,X^n_{t_k}}_{t_{\ell-1}}}{\sigma(t_\ell, X^n_{t_{\ell-1}})} \right\|_4 \\
&\le \left\| (\mathcal{D}^n_{k+1} X^n_{t_m}) - \frac{\sigma(t_{k+1}, X^n_{t_k})\, \mathcal{D}^n_{\ell} X^n_{t_m}}{\sigma(t_\ell, X^n_{t_{\ell-1}})} \frac{\mathcal{D}^n_{k+1} X^n_{t_{\ell-1}}}{\sigma(t_{k+1}, X^n_{t_k})} \right\|_4 \\
&\qquad + \left\| \frac{\sigma(t_{k+1}, X^n_{t_k})\, \mathcal{D}^n_{\ell} X^n_{t_m}}{\sigma(t_\ell, X^n_{t_{\ell-1}})} \left(\nabla X^{n,t_k,X^n_{t_k}}_{t_{\ell-1}} - \frac{\mathcal{D}^n_{k+1} X^n_{t_{\ell-1}}}{\sigma(t_{k+1}, X^n_{t_k})} \right) \right\|_4 .
\end{aligned} \tag{21}$$

The second expression on the r.h.s. of (21) is bounded by $C(b,\sigma,T,\delta)h^{\frac{1}{2}}$ as a consequence of Lemma 5.4 (ii)-(iii). To show that also the first expression is bounded by $C(b,\sigma,T,\delta)h^{\frac{1}{2}}$, we rewrite it using (15) and get

$$
\begin{aligned}
&\left|\frac{\mathcal{D}^n_\ell X^n_{t_m}}{\sigma(t_\ell, X^n_{t_{\ell-1}})}\mathcal{D}^n_{k+1}X^n_{t_{\ell-1}} - \mathcal{D}^n_{k+1}X^n_{t_m}\right| \\
&= \Bigg|\left(1+\sum_{l=\ell+1}^{m}\frac{\mathcal{D}^n_\ell X^n_{t_{l-1}}}{\sigma(t_\ell, X^n_{t_{\ell-1}})}(b_x^{(\ell,l)}h+\sigma_x^{(\ell,l)}\sqrt{h}\varepsilon_l)\right) \\
&\quad\times\left(\sigma(t_{k+1},X^n_{t_k})+\sum_{l=k+2}^{\ell-1}\mathcal{D}^n_{k+1}X^n_{t_{l-1}}(b_x^{(k+1,l)}h+\sigma_x^{(k+1,l)}\sqrt{h}\varepsilon_l)\right) \\
&\quad-\left(\sigma(t_{k+1},X^n_{t_k})+\left(\sum_{l=k+2}^{\ell-1}+\sum_{l=\ell}^{m}\right)\mathcal{D}^n_{k+1}X^n_{t_{l-1}}(b_x^{(k+1,l)}h+\sigma_x^{(k+1,l)}\sqrt{h}\varepsilon_l)\right)\Bigg| \\
&\le \left|\mathcal{D}^n_{k+1}X^n_{t_{\ell-1}}(b_x^{(k+1,\ell)}h+\sigma_x^{(k+1,\ell)}\sqrt{h}\varepsilon_\ell)\right| \\
&\quad+\left|\sum_{l=\ell+1}^{m}\left[\frac{\mathcal{D}^n_\ell X^n_{t_{l-1}}}{\sigma(t_\ell, X^n_{t_{\ell-1}})}\mathcal{D}^n_{k+1}X^n_{t_{\ell-1}} - \mathcal{D}^n_{k+1}X^n_{t_{l-1}}\right]\left(b_x^{(\ell,l)}h+\sigma_x^{(\ell,l)}\sqrt{h}\varepsilon_l\right)\right| \\
&\quad+\left|\sum_{l=\ell+1}^{m}\mathcal{D}^n_{k+1}X^n_{t_{l-1}}\left[b_x^{(\ell,l)}h+\sigma_x^{(\ell,l)}\sqrt{h}\varepsilon_l-\left(b_x^{(k+1,l)}h+\sigma_x^{(k+1,l)}\sqrt{h}\varepsilon_l\right)\right]\right|. \qquad (22)
\end{aligned}
$$

We take the L_4-norm of (22) and apply the BDG inequality and Hölder's inequality. The second term on the r.h.s. of (22) will be used for Gronwall's lemma, while the first and the last one can be bounded by $C(b,\sigma,T)h^{\frac{1}{2}}$, by using Lemma 5.4-(iii). For the last term we also use the Lipschitz continuity of b_x and σ_x in space and Lemma 5.4-(i). □

3. Main results

In order to compute the mean square distance between the solution to (1) and the solution to (3) we construct the random walk B^n from the Brownian motion B by Skorohod embedding. Let

$$\tau_0 := 0 \quad \text{and} \quad \tau_k := \inf\{t > \tau_{k-1} : |B_t - B_{\tau_{k-1}}| = \sqrt{h}\}, \quad k \ge 1. \qquad (23)$$

Then $(B_{\tau_k} - B_{\tau_{k-1}})_{k=1}^{\infty}$ is a sequence of i.i.d. random variables with

$$\mathbb{P}(B_{\tau_k} - B_{\tau_{k-1}} = \pm\sqrt{h}) = \tfrac{1}{2},$$

which means that $\sqrt{h}\varepsilon_k \overset{d}{=} B_{\tau_k} - B_{\tau_{k-1}}$. We will denote by $\mathbb{E}_{\tau_k}$ the conditional expectation w.r.t. $\mathcal{F}_{\tau_k} := \mathcal{G}_k$. In this case we also use the notation $\mathcal{X}_{\tau_k} := X^n_{t_k}$ for all $k = 0,\dots,n$, so that (4)

turns into

$$\mathcal{X}_{\tau_k} = x + \sum_{j=1}^{k} b(t_j, \mathcal{X}_{\tau_{j-1}})h + \sum_{j=1}^{k} \sigma(t_j, \mathcal{X}_{\tau_{j-1}})(B_{\tau_j} - B_{\tau_{j-1}}), \quad 0 \le k \le n.$$

Assumption 3.1. *We assume that the random walk B^n in (3) is given by*

$$B^n_t = \sum_{k=1}^{[t/h]} (B_{\tau_k} - B_{\tau_{k-1}}), \qquad 0 \le t \le T,$$

where the τ_k, $k = 1, ..., n$ are taken from (23).

Remark 3.1. Note that for $p > 0$ there exists a $C(p) > 0$ such that for all $k = 1, \dots, n$ it holds

$$\tfrac{1}{C(p)}(t_k h)^{\frac{1}{4}} \le (\mathbb{E}|B_{\tau_k} - B_{t_k}|^p)^{\frac{1}{p}} \le C(p)(t_k h)^{\frac{1}{4}}.$$

The upper estimate is given in Lemma 5.1. For $p \in [4, \infty)$ the lower estimate follows from [2, Proposition 5.3]. We get the lower estimate for $p \in (0, 4)$ by choosing $0 < \theta < 1$ and $0 < p < p_1$ such that $\frac{1}{4} = \frac{1-\theta}{p} + \frac{\theta}{p_1}$. Then it holds by the log-convexity of L_p norms (see, for example [35, Lemma 1.11.5]) that

$$\|B_{\tau_k} - B_{t_k}\|_p^{1-\theta} \ge \frac{\|B_{\tau_k} - B_{t_k}\|_4}{\|B_{\tau_k} - B_{t_k}\|_{p_1}^{\theta}} \ge \frac{C(4)^{-1}(t_k h)^{\frac{1}{4}}}{\left(C(p_1)(t_k h)^{\frac{1}{4}}\right)^{\theta}} \ge \left(C(p)(t_k h)^{\frac{1}{4}}\right)^{1-\theta}.$$

Since for $t \in [t_k, t_{k+1})$ it holds $B^n_t = B_{\tau_k}$ and $\|B_t - B_{t_k}\|_p \le C(p)h^{\frac{1}{2}}$, we have for any $p > 0$ that

$$\sup_{0 \le t \le T} \|B^n_t - B_t\|_p = O(h^{\frac{1}{4}}). \tag{24}$$

Proposition 3.1 states the convergence rate of (Y_v, Z_v) to (Y^n_v, Z^n_v) in L_2 when $f = 0$ and Theorem 3.1 generalizes this result for any f which satisfies Assumption 2.3.

Proposition 3.1. *Let Assumptions 2.1 and 3.1 hold. If $f = 0$ and $g \in C^1$ is such that g' is a locally α-Hölder continuous function in the sense of (5), then for all $0 \le v < T$, we have (for sufficiently large n) that*

$$\mathbb{E}_{0,x}|Y_v - Y^n_v|^2 \le C^y_{3.1}\Psi(x)^2 h^{\frac{1}{2}}, \quad \text{and} \quad \mathbb{E}_{0,x}|Z_v - Z^n_v|^2 \le C^z_{3.1}\Psi(x)^2 h^{\frac{\alpha}{2}},$$

where $C^y_{3.1} = C(C_g, b, \sigma, T, p_0, \delta)$ and $C^z_{3.1} = C(C_{g'}, b, \sigma, T, p_0, \delta)$.

Theorem 3.1. *Let Assumptions 2.3 and 3.1 be satisfied. Then for all $v \in [0, T)$ and large enough n, we have*

$$\mathbb{E}_{0,x}|Y_v - Y^n_v|^2 + \mathbb{E}_{0,x}|Z_v - Z^n_v|^2 \le C_{3.1}\hat{\Psi}(x)^2 h^{\frac{1}{2}\wedge\alpha}$$

with $C_{3.1} = C(b, \sigma, f, g, T, p_0, \delta)$ and $\hat{\Psi}$ is given in (61).

Remark 3.2. As noticed above, the filtration $\mathcal{G}_k$ coincides with $\mathcal{F}_{\tau_k}$, for all $k = 0, \dots, n$. The expectation $\mathbb{E}_{0,x}$ appearing in Proposition 3.1 and in Theorem 3.1 is defined on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

Remark 3.3. In order to avoid too much notation for the dependencies of the constants, if for example only g is mentioned and not C_g, this means that the estimate might depend also on the bounds of the derivatives of g.

From (24) one can see that the convergence rates stated in Proposition 3.1 and Theorem 3.1 are the natural ones for this approach. The results are proved in the next two sections. In both proofs, we will use the following remark.

Remark 3.4. Since the process $(X_t)_{t\geq 0}$ is strong Markov we can express conditional expectations with the help of an independent copy of B denoted by $\tilde{B}$, for example $\mathbb{E}_{\tau_k} g(X_T^n) = \tilde{\mathbb{E}} g(\tilde{\mathcal{X}}_{\tau_n}^{\tau_k, \mathcal{X}_{\tau_k}})$ for $0 \leq k \leq n$, where

$$\tilde{\mathcal{X}}_{\tau_n}^{\tau_k, \mathcal{X}_{\tau_k}} = \mathcal{X}_{\tau_k} + \sum_{j=k+1}^{n} b(t_j, \tilde{\mathcal{X}}_{\tau_{j-1}}^{\tau_k, \mathcal{X}_{\tau_k}}) h + \sum_{j=k+1}^{n} \sigma(t_j, \tilde{\mathcal{X}}_{\tau_{j-1}}^{\tau_k, \mathcal{X}_{\tau_k}}) (\tilde{B}_{\tilde{\tau}_{j-k}} - \tilde{B}_{\tilde{\tau}_{j-k-1}}), \tag{25}$$

(we define $\tilde{\tau}_k := 0$ and $\tilde{\tau}_j := \inf\{t > \tilde{\tau}_{j-1} : |\tilde{B}_t - \tilde{B}_{\tilde{\tau}_{j-1}}| = \sqrt{h}\}$ for $j \geq 1$ and $\tau_n := \tau_k + \tilde{\tau}_{n-k}$ for $n \geq k$). In fact, to represent the conditional expectations $\mathbb{E}_{t_k}$ and $\mathbb{E}_{\tau_k}$ we work here with $\tilde{\mathbb{E}}$ and the Brownian motions B' and B'', respectively, given by

$$B'_t = B_{t\wedge t_k} + \tilde{B}_{(t-t_k)^+} \quad \text{and} \quad B''_t = B_{t\wedge \tau_k} + \tilde{B}_{(t-\tau_k)^+}, \quad t \geq 0. \tag{26}$$

3.1. Proof of Proposition 3.1: the approximation rates for the zero generator case

To shorten the notation, we use $\mathbb{E} := \mathbb{E}_{0,x}$. Let us first deal with the error of Y. If v belongs to $[t_k, t_{k+1})$ we have $Y_v^n = Y_{t_k}^n$. Then

$$\mathbb{E}|Y_v - Y_v^n|^2 \leq 2(\mathbb{E}|Y_v - Y_{t_k}|^2 + \mathbb{E}|Y_{t_k} - Y_{t_k}^n|^2).$$

Using Theorem 4.1 we bound $\|Y_v - Y_{t_k}\|$ by

$$C_{4.1}^y \Psi(x)(v - t_k)^{\frac{1}{2}} = C(C_g, b, \sigma, T, p_0, \delta)\Psi(x)(v - t_k)^{\frac{1}{2}}$$

(since $\alpha = 1$ can be chosen when g is locally Lipschitz continuous). It remains to bound

$$\mathbb{E}|Y_{t_k} - Y_{t_k}^n|^2 \quad = \quad \mathbb{E}|\mathbb{E}_{t_k} g(X_T) - \mathbb{E}_{\tau_k} g(X_T^n)|^2 = \mathbb{E}|\tilde{\mathbb{E}} g(\tilde{X}_{t_n}^{t_k, X_{t_k}}) - \tilde{\mathbb{E}} g(\tilde{\mathcal{X}}_{\tau_n}^{\tau_k, \mathcal{X}_{\tau_k}})|^2.$$

By (5) and the Cauchy-Schwarz inequality ($\Psi_1 := C_g(1 + |\tilde{X}_{t_n}^{t_k, X_{t_k}}|^{p_0} + |\tilde{\mathcal{X}}_{\tau_n}^{\tau_k, \mathcal{X}_{\tau_k}}|^{p_0})$),

$$|\tilde{\mathbb{E}} g(\tilde{X}_{t_n}^{t_k, X_{t_k}}) - \tilde{\mathbb{E}} g(\tilde{\mathcal{X}}_{\tau_n}^{\tau_k, \mathcal{X}_{\tau_k}})|^2 \quad \leq \quad (\tilde{\mathbb{E}}(\Psi_1 |\tilde{X}_{t_n}^{t_k, X_{t_k}} - \tilde{\mathcal{X}}_{\tau_n}^{\tau_k, \mathcal{X}_{\tau_k}}|))^2 \leq \tilde{\mathbb{E}}(\Psi_1^2) \tilde{\mathbb{E}} |\tilde{X}_{t_n}^{t_k, X_{t_k}} - \tilde{\mathcal{X}}_{\tau_n}^{\tau_k, \mathcal{X}_{\tau_k}}|^2.$$

Finally, we get by Lemma 5.2-(v) that

$$\mathbb{E}|Y_{t_k} - Y^n_{t_k}|^2 \leq \left(\mathbb{E}\tilde{\mathbb{E}}(\Psi_1^4)\right)^{\frac{1}{2}} \left(\mathbb{E}\tilde{\mathbb{E}}|\tilde{X}^{t_k,X_{t_k}}_{t_n} - \tilde{\mathcal{X}}^{\tau_k,\mathcal{X}_{\tau_k}}_{\tau_n}|^4\right)^{\frac{1}{2}} \leq C(C_g, b, \sigma, T, p_0)\Psi(x)^2 h^{\frac{1}{2}}.$$

Let us now deal with the error of Z. We use $\|Z_v - Z^n_v\| \leq \|Z_v - Z_{t_k}\| + \|Z_{t_k} - Z^n_{t_k}\|$ and the representation

$$Z_t = \sigma(t, X_t)\tilde{\mathbb{E}}(g'(\tilde{X}^{t,X_t}_T)\nabla\tilde{X}^{t,X_t}_T)$$

(see Theorem 4.2), where

$$\begin{aligned} \tilde{X}^{t,x}_s &= x + \int_t^s b(r, \tilde{X}^{t,x}_r)dr + \int_t^s \sigma(r, \tilde{X}^{t,x}_r)d\tilde{B}_{r-t}, \\ \nabla\tilde{X}^{t,x}_s &= 1 + \int_t^s b_x(r, \tilde{X}^{t,x}_r)\nabla\tilde{X}^{t,x}_r dr + \int_t^s \sigma_x(r, \tilde{X}^{t,x}_r)\nabla\tilde{X}^{t,x}_r d\tilde{B}_{r-t}, \quad 0 \leq t \leq s \leq T. \end{aligned} \tag{27}$$

For the first term we get by the assumption on g and Lemma 5.2-(i) and (iii)

$$\begin{aligned} \|Z_v - Z_{t_k}\| &= \|\sigma(v, X_v)\tilde{\mathbb{E}}(g'(\tilde{X}^{v,X_v}_T)\nabla\tilde{X}^{v,X_v}_T) - \sigma(t_k, X_{t_k})\tilde{\mathbb{E}}(g'(\tilde{X}^{t_k,X_{t_k}}_T)\nabla\tilde{X}^{t_k,X_{t_k}}_T)\| \\ &\leq \|\sigma(v, X_v) - \sigma(t_k, X_{t_k})\|_4 \|\tilde{\mathbb{E}}(g'(\tilde{X}^{v,X_v}_T)\nabla\tilde{X}^{v,X_v}_T)\|_4 \\ &\quad + \|\sigma\|_\infty \|\tilde{\mathbb{E}}(g'(\tilde{X}^{v,X_v}_T)\nabla\tilde{X}^{v,X_v}_T) - \tilde{\mathbb{E}}(g'(\tilde{X}^{t_k,X_{t_k}}_T)\nabla\tilde{X}^{v,X_v}_T)\| \\ &\quad + \|\sigma\|_\infty \|\tilde{\mathbb{E}}(g'(\tilde{X}^{t_k,X_{t_k}}_T)\nabla\tilde{X}^{v,X_v}_T) - \tilde{\mathbb{E}}(g'(\tilde{X}^{t_k,X_{t_k}}_T)\nabla\tilde{X}^{t_k,X_{t_k}}_T)\| \\ &\leq C(C_{g'}, b, \sigma, T, p_0)\Psi(x)\Big[h^{\frac{1}{2}} + \|X_v - X_{t_k}\|_4 + \left(\mathbb{E}\tilde{\mathbb{E}}|\tilde{X}^{v,X_v}_T - \tilde{X}^{t_k,X_{t_k}}_T|^{4\alpha}\right)^{\frac{1}{4}} \\ &\quad + \left(\mathbb{E}\tilde{\mathbb{E}}|\nabla\tilde{X}^{v,X_v}_T - \nabla\tilde{X}^{t_k,X_{t_k}}_T|^4\right)^{\frac{1}{4}}\Big] \\ &\leq C(C_{g'}, b, \sigma, T, p_0)\Psi(x)h^{\frac{\alpha}{2}}. \end{aligned}$$

We compute the second term using $Z^n_{t_k}$ as given in (16). Hence, with the notation from Definition 2.2,

$$\begin{aligned} \|Z_{t_k} - Z^n_{t_k}\|^2 &= \mathbb{E}\left|\sigma(t_k, X_{t_k})\tilde{\mathbb{E}}g'(\tilde{X}^{t_k,X_{t_k}}_{t_n})\nabla\tilde{X}^{t_k,X_{t_k}}_{t_n} - \tilde{\mathbb{E}}\mathcal{D}^n_{k+1}g(\tilde{\mathcal{X}}^{\tau_k,\mathcal{X}_{\tau_k}}_{\tau_n})\right|^2 \\ &\leq \|\sigma\|^2_\infty \,\mathbb{E}\left|\tilde{\mathbb{E}}(g'(\tilde{X}^{t_k,X_{t_k}}_{t_n})\nabla\tilde{X}^{t_k,X_{t_k}}_{t_n}) - \frac{\tilde{\mathbb{E}}\mathcal{D}^n_{k+1}g(\tilde{\mathcal{X}}^{\tau_k,\mathcal{X}_{\tau_k}}_{\tau_n})}{\sigma(t_k, X_{t_k})}\right|^2 \\ &= \|\sigma\|^2_\infty \,\mathbb{E}\left|\tilde{\mathbb{E}}(g'(\tilde{X}^{t_k,X_{t_k}}_{t_n})\nabla\tilde{X}^{t_k,X_{t_k}}_{t_n}) - \tilde{\mathbb{E}}\left(g^{(k+1,n+1)}_x \frac{\mathcal{D}^n_{k+1}\tilde{\mathcal{X}}^{\tau_k,\mathcal{X}_{\tau_k}}_{\tau_n}}{\sigma(t_k, X_{t_k})}\right)\right|^2. \end{aligned}$$

We insert $\pm\tilde{\mathbb{E}}(g_x^{(k+1,n+1)}\nabla\tilde{X}_{t_n}^{t_k,X_{t_k}})$ and get by the Cauchy-Schwarz inequality that

$$\begin{aligned}&\Big|\tilde{\mathbb{E}}(g'(\tilde{X}_{t_n}^{t_k,X_{t_k}})\nabla\tilde{X}_{t_n}^{t_k,X_{t_k}}) - \tilde{\mathbb{E}}\Big(g_x^{(k+1,n+1)}\frac{\mathcal{D}_{k+1}^n\tilde{\mathcal{X}}_{\tau_n}^{\tau_k,\mathcal{X}_{\tau_k}}}{\sigma(t_k,X_{t_k})}\Big)\Big|^2\\ &\quad\le 2\tilde{\mathbb{E}}|g'(\tilde{X}_{t_n}^{t_k,X_{t_k}}) - g_x^{(k+1,n+1)}|^2\tilde{\mathbb{E}}|\nabla\tilde{X}_{t_n}^{t_k,X_{t_k}}|^2 + 2\tilde{\mathbb{E}}|g_x^{(k+1,n+1)}|^2\tilde{\mathbb{E}}\Big|\nabla\tilde{X}_{t_n}^{t_k,X_{t_k}} - \frac{\mathcal{D}_{k+1}^n\tilde{\mathcal{X}}_{\tau_n}^{\tau_k,\mathcal{X}_{\tau_k}}}{\sigma(t_k,X_{t_k})}\Big|^2.\end{aligned} \tag{28}$$

For the estimate of $\tilde{\mathbb{E}}|\nabla\tilde{X}_{t_n}^{t_k,X_{t_k}}|^2$ we use Lemma 5.2. Since g' satisfies (5) we proceed with

$$\begin{aligned}&\tilde{\mathbb{E}}|g'(\tilde{X}_{t_n}^{t_k,X_{t_k}}) - g_x^{(k+1,n+1)}|^2\\ &\quad\le \int_0^1\tilde{\mathbb{E}}\Big|g'(\tilde{X}_{t_n}^{t_k,X_{t_k}}) - g'(\vartheta T_{k+1,+}\tilde{\mathcal{X}}_{\tau_n}^{\tau_k,\mathcal{X}_{\tau_k}} + (1-\vartheta)T_{k+1,-}\tilde{\mathcal{X}}_{\tau_n}^{\tau_k,\mathcal{X}_{\tau_k}})\Big|^2 d\vartheta\\ &\quad\le \int_0^1(\tilde{\mathbb{E}}\Psi_1^4)^{\frac12}\Big[\tilde{\mathbb{E}}\Big|\tilde{X}_{t_n}^{t_k,X_{t_k}} - \vartheta T_{k+1,+}\tilde{\mathcal{X}}_{\tau_n}^{\tau_k,\mathcal{X}_{\tau_k}} - (1-\vartheta)T_{k+1,-}\tilde{\mathcal{X}}_{\tau_n}^{\tau_k,\mathcal{X}_{\tau_k}}\Big|^{4\alpha}\Big]^{\frac12}d\vartheta,\end{aligned}$$

where $\Psi_1 := C_{g'}(1 + |\tilde{X}_{t_n}^{t_k,X_{t_k}}|^{p_0} + |\vartheta T_{k+1,+}\tilde{\mathcal{X}}_{\tau_n}^{\tau_k,\mathcal{X}_{\tau_k}} + (1-\vartheta)T_{k+1,-}\tilde{\mathcal{X}}_{\tau_n}^{\tau_k,\mathcal{X}_{\tau_k}}|^{p_0})$. For $\tilde{\mathbb{E}}\Psi_1^4$ and

$$\begin{aligned}&\tilde{\mathbb{E}}\Big|\tilde{X}_{t_n}^{t_k,X_{t_k}} - (\vartheta T_{k+1,+}\tilde{\mathcal{X}}_{\tau_n}^{\tau_k,\mathcal{X}_{\tau_k}} + (1-\vartheta)T_{k+1,-}\tilde{\mathcal{X}}_{\tau_n}^{\tau_k,\mathcal{X}_{\tau_k}})\Big|^{4\alpha}\\ &\quad\le 8\Big(\vartheta^{2\alpha}\tilde{\mathbb{E}}\Big|\tilde{X}_{t_n}^{t_k,X_{t_k}} - T_{k+1,+}\tilde{\mathcal{X}}_{\tau_n}^{\tau_k,\mathcal{X}_{\tau_k}}\Big|^{4\alpha} + (1-\vartheta)^{2\alpha}\tilde{\mathbb{E}}\Big|\tilde{X}_{t_n}^{t_k,X_{t_k}} - T_{k+1,-}\tilde{\mathcal{X}}_{\tau_n}^{\tau_k,\mathcal{X}_{\tau_k}}\Big|^{4\alpha}\Big)\\ &\quad\le C(b,\sigma,T)h^{2\alpha} + C(b,\sigma,T)(|X_{t_k} - \mathcal{X}_{\tau_k}|^{4\alpha} + h^\alpha),\end{aligned}$$

we use Lemma 5.4 and Lemma 5.2-(v). For the last term in (28) we notice that

$$\mathbb{E}\tilde{\mathbb{E}}|g_x^{(k+1,n+1)}|^4 \le C(C_{g'},b,\sigma,T,p_0)\Psi^4(x).$$

By Lemma 5.2 we have $\mathbb{E}\tilde{\mathbb{E}}|\nabla\tilde{X}_{t_n}^{t_k,X_{t_k}} - \nabla\tilde{\mathcal{X}}_{\tau_n}^{\tau_k,\mathcal{X}_{\tau_k}}|^p \le C(b,\sigma,T,p)h^{\frac{p}{4}}$, and by Lemma 5.4,

$$\begin{aligned}&\mathbb{E}\tilde{\mathbb{E}}\Big|\nabla\tilde{\mathcal{X}}_{\tau_n}^{\tau_k,\mathcal{X}_{\tau_k}} - \frac{\mathcal{D}_{k+1}^n\tilde{\mathcal{X}}_{\tau_n}^{\tau_k,\mathcal{X}_{\tau_k}}}{\sigma(t_k,X_{t_k})}\Big|^p\\ &\le C(p)\mathbb{E}\Big|\nabla X_{t_n}^{n,t_k,X_{t_k}^n} - \frac{\mathcal{D}_{k+1}^nX_{t_n}^n}{\sigma(t_{k+1},X_{t_k}^n)}\Big|^p + C(p)\mathbb{E}\Big|\frac{\mathcal{D}_{k+1}^nX_{t_n}^n}{\sigma(t_{k+1},X_{t_k}^n)} - \frac{\mathcal{D}_{k+1}^nX_{t_n}^n}{\sigma(t_k,X_{t_k})}\Big|^p \le C(b,\sigma,T,p,\delta)h^{\frac{p}{4}}.\end{aligned}$$

Consequently, $\|Z_{t_k} - Z_{t_k}^n\|^2 \le C(C_{g'},b,\sigma,T,p_0,\delta)\Psi^2(x)h^{\frac{\alpha}{2}}$.

3.2. Proof of Theorem 3.1: the approximation rates for the general case

Let $u : [0,T)\times\mathbb{R}\to\mathbb{R}$ be the solution of the PDE (37) associated to (1). We use the representations $Y_s = u(s,X_s)$ and $Z_s = \sigma(s,X_s)u_x(s,X_s)$ stated in Theorem 4.2 and define

$$F(s,x) := f(s,x,u(s,x),\sigma(s,x)u_x(s,x)). \tag{29}$$

From (1) and (3) we conclude

$$\begin{aligned}\|Y_{t_k} - Y^n_{t_k}\| &\leq \|\mathbb{E}_{t_k} g(X_T) - \mathbb{E}_{\tau_k} g(X^n_T)\| \\ &\quad + \left\| \mathbb{E}_{t_k} \int_{t_k}^T f(s, X_s, Y_s, Z_s) ds - h\mathbb{E}_{\tau_k} \sum_{m=k}^{n-1} f(t_{m+1}, X^n_{t_m}, Y^n_{t_m}, Z^n_{t_m}) \right\|,\end{aligned}$$

where Proposition 3.1 provides the estimate for the terminal condition. We decompose the generator term as follows:

$$\begin{aligned}&\mathbb{E}_{t_k} f(s, X_s, Y_s, Z_s) - \mathbb{E}_{\tau_k} f(t_{m+1}, X^n_{t_m}, Y^n_{t_m}, Z^n_{t_m}) \\ &= [\mathbb{E}_{t_k} f(s, X_s, Y_s, Z_s) - \mathbb{E}_{t_k} f(t_m, X_{t_m}, Y_{t_m}, Z_{t_m})] + [\mathbb{E}_{t_k} F(t_m, X_{t_m}) - \mathbb{E}_{\tau_k} F(t_m, X^n_{t_m})] \\ &\quad + [\mathbb{E}_{\tau_k} F(t_m, X^n_{t_m}) - \mathbb{E}_{\tau_k} F(t_m, X_{t_m})] + [\mathbb{E}_{\tau_k} f(t_m, X_{t_m}, Y_{t_m}, Z_{t_m}) - \mathbb{E}_{\tau_k} f(t_{m+1}, X^n_{t_m}, Y^n_{t_m}, Z^n_{t_m})] \\ &=: d_1(s,m) + d_2(m) + d_3(m) + d_4(m).\end{aligned}$$

We use

$$\begin{aligned}&\left\| \mathbb{E}_{t_k} \int_{t_k}^T f(s, X_s, Y_s, Z_s) ds - h\mathbb{E}_{\tau_k} \sum_{m=k}^{n-1} f(t_{m+1}, X^n_{t_m}, Y^n_{t_m}, Z^n_{t_m}) \right\| \\ &\qquad \leq \sum_{m=k}^{n-1} \left(\left\| \int_{t_m}^{t_{m+1}} d_1(s,m) ds \right\| + h \sum_{i=2}^4 \|d_i(m)\| \right)\end{aligned}$$

and estimate the expressions on the right hand side. For the function F defined in (29) we use Assumption 2.3 (which implies that (5) holds for $\alpha = 1$) to derive by Theorem 4.2 and the mean value theorem that for $x_1, x_2 \in \mathbb{R}$ there exist $\xi \in [\min\{x_1, x_2\}, \max\{x_1, x_2\}]$ such that

$$\begin{aligned}|F(t, x_1) - F(t, x_2)| &= |f(t, x_1, u(t, x_1), \sigma(t, x_1) u_x(t, x_1)) - f(t, x_2, u(t, x_2), \sigma(t, x_2) u_x(t, x_2))| \\ &\leq C(L_f, \sigma) \left(1 + c^2_{4.2} \Psi(\xi) + \frac{c^3_{4.2} \Psi(\xi)}{(T-t)^{\frac{1}{2}}} \right) |x_1 - x_2| \\ &\leq C(L_f, c^{2,3}_{4.2}, \sigma, T)(1 + |x_1|^{p_0+1} + |x_2|^{p_0+1}) \frac{|x_1 - x_2|}{(T-t)^{\frac{1}{2}}}. \qquad (30)\end{aligned}$$

By (6), standard estimates on (X_s), Theorem 4.1-(i) and Proposition 4.1 for $p = 2$ we immediately get

$$\begin{aligned}\|d_1(s,m)\| &\leq C(L_f, C^y_{4.1}, C_{4.1}, b, \sigma, T) \Psi(x)\, h^{\frac{1}{2}} \\ &= C(b, \sigma, f, g, T, p_0, \delta) \Psi(x)\, h^{\frac{1}{2}}.\end{aligned}$$

For the estimate of d_2 one exploits

$$\mathbb{E}_{t_k} F(t_m, X_{t_m}) - \mathbb{E}_{\tau_k} F(t_m, X^n_{t_m}) = \tilde{\mathbb{E}} F(t_m, \tilde{X}^{t_k, X_{t_k}}_{t_m}) - \tilde{\mathbb{E}} F(t_m, \tilde{X}^{n, t_k, X^n_{t_k}}_{t_m})$$

and then uses (30) and Lemma 5.2-(v). This gives

$$\|d_2(m)\| \quad \leq \quad C(L_f, c^{2,3}_{4.2}, b, \sigma, T, p_0)\Psi(x)\frac{1}{(T-t_m)^{\frac{1}{2}}}h^{\frac{1}{4}}.$$

For d_3 we start with Jensen's inequality and continue then similarly as above to get

$$\|d_3(m)\| \leq \|F(t_m, X^n_{t_m}) - F(t_m, X_{t_m})\| \leq C(L_f, c^{2,3}_{4.2}, b, \sigma, T, p_0)\Psi(x)\frac{1}{(T-t_m)^{\frac{1}{2}}}h^{\frac{1}{4}},$$

and for the last term we get

$$\|d_4(m)\| \quad \leq \quad L_f(h^{\frac{1}{2}} + \|X_{t_m} - X^n_{t_m}\| + \|Y_{t_m} - Y^n_{t_m}\| + \|Z_{t_m} - Z^n_{t_m}\|).$$

This implies

$$\|Y_{t_k} - Y^n_{t_k}\| \leq C\Psi(x)h^{\frac{1}{4}} + hL_f\sum_{m=k}^{n-1}(\|Y_{t_m} - Y^n_{t_m}\| + \|Z_{t_m} - Z^n_{t_m}\|), \tag{31}$$

where $C = C(L_f, C^y_{3.1}, C^y_{4.1}, C_{4.1}, c^{2,3}_{4.2}, b, \sigma, T, p_0) = C(b, \sigma, f, g, T, p_0, \delta)$.

For $\|Z_{t_k} - Z^n_{t_k}\|$ we use the representations (13), (16) and the approximation (19) as well as Proposition 2.1. Instead of $N^{n,t_k}_{t_n}$ we will use here the notation $N^{n,\tau_k}_{\tau_n}$ to indicate its measurability w.r.t. the filtration $(\mathcal{F}_t)$. It holds that

$$\begin{aligned}
\|Z^n_{t_k} - Z_{t_k}\| \quad &\leq \quad \|Z^n_{t_k} - \hat{Z}^n_{t_k}\| + \|Z_{t_k} - \hat{Z}^n_{t_k}\| \\
&\leq \quad C_{2.1}\hat{\Psi}(x)h^{\frac{\alpha}{2}} + \|\sigma(t_k, X_{t_k})\tilde{\mathbb{E}}g'(\tilde{X}^{t_k,X_{t_k}}_{t_n})\nabla\tilde{X}^{t_k,X_{t_k}}_{t_n} - \tilde{\mathbb{E}}\mathcal{D}^n_{k+1}g(\tilde{X}^{n,t_k,X^n_{t_k}}_{t_n})\| \\
&\quad + \Big\|\mathbb{E}_{t_k}\int_{t_{k+1}}^{T} f(s, X_s, Y_s, Z_s)N^{t_k}_s ds\,\sigma(t_k, X_{t_k}) \\
&\qquad -\mathbb{E}_{\tau_k}h\sum_{m=k+1}^{n-1} f(t_{m+1}, X^n_{t_m}, Y^n_{t_m}, Z^n_{t_m})N^{n,\tau_k}_{\tau_m}\sigma(t_{k+1}, X^n_{t_k})\Big\| \\
&\quad + \Big\|\mathbb{E}_{t_k}\int_{t_k}^{t_{k+1}} f(s, X_s, Y_s, Z_s)N^{t_k}_s ds\,\sigma(t_k, X_{t_k})\Big\|.
\end{aligned} \tag{32}$$

For the terminal condition Proposition 3.1 provides

$$\|\sigma(t_k, X_{t_k})\tilde{\mathbb{E}}g'(\tilde{X}^{t_k,X_{t_k}}_{t_n})\nabla\tilde{X}^{t_k,X_{t_k}}_{t_n} - \tilde{\mathbb{E}}\mathcal{D}^n_{k+1}g(\tilde{X}^{n,t_k,X^n_{t_k}}_{t_n})\| \leq (C^z_{3.1})^{\frac{1}{2}}\Psi(x)h^{\frac{1}{4}}. \tag{33}$$

We continue with the generator terms and use F defined in (29) to decompose the difference

$$\begin{aligned}
&\mathbb{E}_{t_k}f(s, X_s, Y_s, Z_s)N^{t_k}_s\sigma(t_k, X_{t_k}) - \mathbb{E}_{\tau_k}f(t_{m+1}, X^n_{t_m}, Y^n_{t_m}, Z^n_{t_m})N^{n,\tau_k}_{\tau_m}\sigma(t_{k+1}, X^n_{t_k}) \\
&\quad = \quad \mathbb{E}_{t_k}f(s, X_s, Y_s, Z_s)N^{t_k}_s\sigma(t_k, X_{t_k}) - \mathbb{E}_{t_k}f(t_m, X_{t_m}, Y_{t_m}, Z_{t_m})N^{t_k}_{t_m}\sigma(t_k, X_{t_k}) \\
&\qquad +\mathbb{E}_{t_k}F(t_m, X_{t_m})N^{t_k}_{t_m}\sigma(t_k, X_{t_k}) - \mathbb{E}_{\tau_k}F(t_m, X^n_{t_m})N^{n,\tau_k}_{\tau_m}\sigma(t_{k+1}, X^n_{t_k}) \\
&\qquad +\mathbb{E}_{\tau_k}\left[[F(t_m, X^n_{t_m}) - F(t_m, X_{t_m})]N^{n,\tau_k}_{\tau_m}\sigma(t_{k+1}, X^n_{t_k})\right] \\
&\qquad +\mathbb{E}_{\tau_k}\left[[f(t_m, X_{t_m}, Y_{t_m}, Z_{t_m}) - f(t_{m+1}, X^n_{t_m}, Y^n_{t_m}, Z^n_{t_m})]N^{n,\tau_k}_{\tau_m}\sigma(t_{k+1}, X^n_{t_k})\right] \\
&\quad =: \quad \mathtt{t}_1(s, m) + \mathtt{t}_2(m) + \mathtt{t}_3(m) + \mathtt{t}_4(m)
\end{aligned}$$

where $s \in [t_m, t_{m+1})$. For $\mathtt{t}_1$ we use that $\mathbb{E}_{t_k} f(t_m, X_{t_k}, Y_{t_k}, Z_{t_k})(N_s^{t_k} - N_{t_m}^{t_k}) = 0$, so that

$$\begin{aligned}\|\mathtt{t}_1(s,m)\| &\leq \|\mathbb{E}_{t_k} f(s, X_s, Y_s, Z_s) N_s^{t_k} \sigma(t_k, X_{t_k}) - \mathbb{E}_{t_k} f(t_m, X_{t_m}, Y_{t_m}, Z_{t_m}) N_s^{t_k} \sigma(t_k, X_{t_k})\| \\ &\quad + \|\mathbb{E}_{t_k}(f(t_m, X_{t_m}, Y_{t_m}, Z_{t_m}) - f(t_m, X_{t_k}, Y_{t_k}, Z_{t_k}))(N_s^{t_k} - N_{t_m}^{t_k})\sigma(t_k, X_{t_k})\|.\end{aligned}$$

As before, we rewrite the conditional expectations with the help of the independent copy $\tilde{B}$. Then

$$\begin{aligned}&\mathbb{E}_{t_k} f(s, X_s, Y_s, Z_s) N_s^{t_k} - \mathbb{E}_{t_k} f(t_m, X_{t_m}, Y_{t_m}, Z_{t_m}) N_s^{t_k} \\ &\quad = \tilde{\mathbb{E}}[(f(s, \tilde{X}_s^{t_k, X_{t_k}}, \tilde{Y}_s^{t_k, X_{t_k}}, \tilde{Z}_s^{t_k, X_{t_k}}) - f(t_m, \tilde{X}_{t_m}^{t_k, X_{t_k}}, \tilde{Y}_{t_m}^{t_k, X_{t_k}}, \tilde{Z}_{t_m}^{t_k, X_{t_k}}))\tilde{N}_s^{t_k}]\end{aligned}$$

and

$$\begin{aligned}&\mathbb{E}_{t_k}(f(t_m, X_{t_m}, Y_{t_m}, Z_{t_m}) - f(t_m, X_{t_k}, Y_{t_k}, Z_{t_k}))(N_s^{t_k} - N_{t_m}^{t_k}) \\ &\quad = \tilde{\mathbb{E}}[(f(t_m, \tilde{X}_{t_m}^{t_k, X_{t_k}}, \tilde{Y}_{t_m}^{t_k, X_{t_k}}, \tilde{Z}_{t_m}^{t_k, X_{t_k}}) - f(t_m, X_{t_k}, Y_{t_k}, Z_{t_k}))(\tilde{N}_s^{t_k} - \tilde{N}_{t_m}^{t_k})].\end{aligned}$$

We apply the conditional Hölder inequality, and from the estimates (36) and $\tilde{\mathbb{E}}|\tilde{N}_s^{t_k} - \tilde{N}_{t_m}^{t_k}|^2 \leq C(b, \sigma, T, \delta)\frac{h}{(s-t_k)^2}$ we get

$$\begin{aligned}\|\mathtt{t}_1(s,m)\| &\leq \frac{\kappa_2 \|\sigma\|_\infty}{(s-t_k)^{\frac{1}{2}}} \|f(s, X_s, Y_s, Z_s) - f(t_m, X_{t_m}, Y_{t_m}, Z_{t_m})\| \\ &\quad + C(b, \sigma, T, \delta)\frac{h^{\frac{1}{2}}}{s - t_k} \|f(t_m, X_{t_m}, Y_{t_m}, Z_{t_m}) - f(t_k, X_{t_k}, Y_{t_k}, Z_{t_k})\| \\ &\leq C(L_f, C^y_{4.1}, C_{4.1}, \kappa_2, b, \sigma, T, p_0, \delta)\Psi(x)\frac{h^{\frac{1}{2}}}{(s-t_k)^{\frac{1}{2}}},\end{aligned}$$

since for $0 \leq t < s \leq T$ we have by Theorem 4.1 and Proposition 4.1 that

$$\|f(s, X_s, Y_s, Z_s) - f(t, X_t, Y_t, Z_t)\| \leq C(L_f, C^y_{4.1}, C_{4.1}, b, \sigma, T, p_0)\Psi(x)(s-t)^{\frac{1}{2}}. \tag{34}$$

For the estimate of $\mathtt{t}_2$ Lemma 5.2, Lemma 5.3, (30) and (36) yield

$$\begin{aligned}\|\mathtt{t}_2(m)\| &= \|\tilde{\mathbb{E}} F(t_m, \tilde{X}_{t_m}^{t_k, X_{t_k}})\tilde{N}_{t_m}^{t_k}\sigma(t_k, X_{t_k}) - \tilde{\mathbb{E}} F(t_m, \tilde{\mathcal{X}}_{\tau_m}^{\tau_k, \mathcal{X}_{\tau_k}})\tilde{N}_{\tau_m}^{n,\tau_k}\sigma(t_{k+1}, \mathcal{X}_{\tau_k})\| \\ &\leq \frac{C(\kappa_2, \sigma)}{(t_m - t_k)^{\frac{1}{2}}}\left(\mathbb{E}\tilde{\mathbb{E}}(F(t_m, \tilde{X}_{t_m}^{t_k, X_{t_k}}) - F(t_m, \tilde{\mathcal{X}}_{\tau_m}^{\tau_k, \mathcal{X}_{\tau_k}}))^2\right)^{\frac{1}{2}} \\ &\quad + (\mathbb{E}\tilde{\mathbb{E}}|F(t_m, \tilde{\mathcal{X}}_{\tau_m}^{\tau_k, \mathcal{X}_{\tau_k}}) - F(t_m, \mathcal{X}_{\tau_k})|^2 \tilde{\mathbb{E}}|\tilde{N}_{t_m}^{t_k}\sigma(t_k, X_{t_k}) - \tilde{N}_{\tau_m}^{n,\tau_k}\sigma(t_{k+1}, \mathcal{X}_{\tau_k})|^2)^{\frac{1}{2}} \\ &\leq C(L_f, c^{2,3}_{4.2}, \kappa_2, b, \sigma, T, p_0, \delta)\frac{\Psi(x)}{(T-t_m)^{\frac{1}{2}}}\frac{h^{\frac{1}{4}}}{(t_m - t_k)^{\frac{1}{2}}}.\end{aligned}$$

For $\mathtt{t}_3$ we use the conditional Hölder inequality, (30), (18) and Lemma 5.2:

$$\begin{aligned}\|\mathtt{t}_3(m)\| &= \left\|\mathbb{E}_{\tau_k}\left[[F(t_m, X_{t_m}^n) - F(t_m, X_{t_m})]N_{\tau_m}^{n,\tau_k}\sigma(t_{k+1}, \mathcal{X}_{\tau_k})\right]\right\| \\ &\leq \frac{C(\widehat{\kappa}_2, \sigma)}{(t_m - t_k)^{\frac{1}{2}}}\left\|F(t_m, X_{t_m}^n) - F(t_m, X_{t_m})\right\| \\ &\leq C(L_f, c^{2,3}_{4.2}, b, \sigma, T, p_0, \delta)\frac{\Psi(x)}{(T-t_m)^{\frac{1}{2}}}\frac{h^{\frac{1}{4}}}{(t_m - t_k)^{\frac{1}{2}}}.\end{aligned}$$

The term $\mathtt{t}_4$ can be estimated as follows:

$$\begin{aligned}\|\mathtt{t}_4(m)\| &= \left\|\mathbb{E}_{\tau_k}\left[[f(t_m,X_{t_m},Y_{t_m},Z_{t_m})-f(t_{m+1},X^n_{t_m},Y^n_{t_m},Z^n_{t_m})]N^{n,\tau_k}_{\tau_m}\sigma(t_{k+1},\mathcal{X}_{\tau_k})\right]\right\| \\ &\leq \frac{C(L_f,b,\sigma,T,\delta)}{(t_m-t_k)^{\frac12}}(h^{\frac12}+\|X_{t_m}-X^n_{t_m}\|+\|Y_{t_m}-Y^n_{t_m}\|+\|Z_{t_m}-Z^n_{t_m}\|).\end{aligned}$$

Finally, for the remaining term of the estimate of $\|Z_{t_k}-Z^n_{t_k}\|$, we use (34) and (36) to get

$$\begin{aligned}\left\|\mathbb{E}_{t_k}f(s,X_s,Y_s,Z_s)N^{t_k}_s\,\sigma(t_k,X_{t_k})\right\| &= \|\mathbb{E}_{t_k}[(f(s,X_s,Y_s,Z_s)-f(s,X_{t_k},Y_{t_k},Z_{t_k}))N^{t_k}_s]\,\sigma(t_k,X_{t_k})\| \\ &\leq C(L_f,C^y_{4.1},C_{4.1},\kappa_2,b,\sigma,T,p_0)\Psi(x).\end{aligned}$$

Consequently, from (32), (33), the estimates for the remaining term and for $\mathtt{t}_1,...,\mathtt{t}_4$ it follows that

$$\begin{aligned}\|Z_{t_k}-Z^n_{t_k}\| &\leq C_{2.1}\hat\Psi(x)h^{\frac\alpha2}+(C^z_{3.1})^{\frac12}\Psi(x)h^{\frac14}+C(L_f,C^y_{4.1},C_{4.1},b,\sigma,T,p_0,\kappa_2)\Psi(x)h \\ &\quad+C(L_f,C^y_{4.1},C_{4.1},\kappa_2,b,\sigma,T,p_0,\delta)\Psi(x)h^{\frac12}\int_{t_k}^T\frac{ds}{(s-t_k)^{\frac12}} \\ &\quad+C(L_f,c^{2,3}_{4.2},\kappa_2,b,\sigma,T,p_0,\delta)h\sum_{m=k+1}^{n-1}\frac{\Psi(x)}{(T-t_m)^{\frac12}}\frac{h^{\frac14}}{(t_m-t_k)^{\frac12}} \\ &\quad+C(L_f,b,\sigma,T,\delta)h\sum_{m=k+1}^{n-1}(\|Y_{t_m}-Y^n_{t_m}\|+\|Z_{t_m}-Z^n_{t_m}\|)\frac{1}{(t_m-t_k)^{\frac12}} \\ &\leq C(C_{2.1},C^z_{3.1})\hat\Psi(x)h^{\frac\alpha2\wedge\frac14}+C(L_f,c^{2,3}_{4.2},C^y_{4.1},C_{4.1},\kappa_2,b,\sigma,T,p_0,\delta)\Psi(x)h^{\frac14} \\ &\quad+C(L_f,b,\sigma,T,\delta)\sum_{m=k+1}^{n-1}(\|Y_{t_m}-Y^n_{t_m}\|+\|Z_{t_m}-Z^n_{t_m}\|)\frac{1}{(t_m-t_k)^{\frac12}}h.\end{aligned}$$

Then we use (31) and the above estimate to get

$$\begin{aligned}&\|Y_{t_k}-Y^n_{t_k}\|+\|Z_{t_k}-Z^n_{t_k}\| \\ &\leq C(C_{2.1},C^z_{3.1})\hat\Psi(x)h^{\frac\alpha2\wedge\frac14}+C(L_f,C^y_{3.1},C^y_{4.1},C_{4.1},c^{2,3}_{4.2},\kappa_2,b,\sigma,T,p_0,\delta)\Psi(x)h^{\frac14} \\ &\quad+C(L_f,b,\sigma,T,\delta)\sum_{m=k+1}^{n-1}(\|Y_{t_m}-Y^n_{t_m}\|+\|Z_{t_m}-Z^n_{t_m}\|)\frac{1}{(t_m-t_k)^{\frac12}}h.\end{aligned}$$

Consequently, summarizing the dependencies, there is a $C=C(b,\sigma,f,g,T,p_0,\delta)$ such that

$$\|Y_{t_k}-Y^n_{t_k}\|+\|Z_{t_k}-Z^n_{t_k}\| \leq C\hat\Psi(x)h^{\frac\alpha2\wedge\frac14}.$$

By Theorem 4.1 (note that by Assumption 2.3 on g we have $\alpha=1$) it follows that

$$\|Y_v-Y^n_v\|\leq\|Y_v-Y_{t_k}\|+\|Y_{t_k}-Y^n_{t_k}\|\leq C^y_{4.1}\Psi(x)h^{\frac12}+\hat\Psi(x)h^{\frac\alpha2\wedge\frac14}.$$

while Proposition 4.1 implies that

$$\|Z_v - Z_{t_k}\| \leq C_{4.1}\Psi(x)h^{\frac{1}{2}},$$

and hence we have

$$\mathbb{E}_{0,x}|Y_v - Y_v^n|^2 + \mathbb{E}_{0,x}|Z_v - Z_v^n|^2 \leq C_{3.1}\hat{\Psi}(x)^2 h^{\frac{1}{2}\wedge\alpha}$$

with $C_{3.1} = C_{3.1}(b, \sigma, f, g, T, p_0, \delta)$.

4. Some properties of solutions to BSDEs and their associated PDEs

4.1. Malliavin weights

We use the SDE from (1) started in (t, x),

$$X_s^{t,x} = x + \int_t^s b(r, X_r^{t,x})dr + \int_t^s \sigma(r, X_r^{t,x})dB_r, \quad 0 \leq t \leq s \leq T \tag{35}$$

and recall the Malliavin weight and its properties from [20, Subsection 1.1 and Remark 3].

Lemma 4.1. *Let $H : \mathbb{R} \to \mathbb{R}$ be a polynomially bounded Borel function. If Assumption 2.1 holds and $X^{t,x}$ is given by (35) then setting*

$$G(t, x) := \mathbb{E}H(X_T^{t,x})$$

implies that $G \in C^{1,2}([0, T) \times \mathbb{R})$. Especially it holds for $0 \leq t \leq r < T$ that

$$\partial_x G(r, X_r^{t,x}) = \mathbb{E}[H(X_T^{t,x})N_T^{r,(t,x)}|\mathcal{F}_r^t],$$

where $(\mathcal{F}_r^t)_{r\in[t,T]}$ is the augmented natural filtration of $(B_r^{t,0})_{r\in[t,T]}$,

$$N_T^{r,(t,x)} = \frac{1}{T-r}\int_r^T \frac{\nabla X_s^{t,x}}{\sigma(s, X_s^{t,x})\nabla X_r^{t,x}}dB_s,$$

and $\nabla X_s^{t,x}$ is given in (12). Moreover, for $q \in (0, \infty)$ there exists a $\kappa_q > 0$ such that a.s.

$$(\mathbb{E}[|N_T^{r,(t,x)}|^q|\mathcal{F}_r^t])^{\frac{1}{q}} \leq \frac{\kappa_q}{(T-r)^{\frac{1}{2}}} \quad \text{and} \quad \mathbb{E}[N_T^{r,(t,x)}|\mathcal{F}_r^t] = 0 \text{ a.s.} \tag{36}$$

and we have

$$\|\partial_x G(r, X_r^{t,x})\|_{L_p(\mathbb{P})} \leq \kappa_q \frac{\|H(X_T^{t,x}) - \mathbb{E}[H(X_T^{t,x})|\mathcal{F}_r^t]\|_p}{\sqrt{T-r}}$$

for $1 < q, p < \infty$ with $\frac{1}{p} + \frac{1}{q} = 1$.

4.2. Regularity of solutions to BSDEs

The following result originates from [20, Theorem 1] where also path dependent cases were included. We formulate it only for our Markovian setting but use $\mathbb{P}_{t,x}$ since we are interested in an estimate for all $(t,x) \in [0,T) \times \mathbb{R}$. A sketch of a proof of this formulation can be found in [21].

Theorem 4.1. *Let Assumption 2.1 and 2.2 hold. Then for any $p \in [2,\infty)$ the following assertions are true.*

(i) There exists a constant $C^y_{4.1} > 0$ such that for $0 \le t < s \le T$ and $x \in \mathbb{R}$,

$$\|Y_s - Y_t\|_{L_p(\mathbb{P}_{t,x})} \le C^y_{4.1} \Psi(x) \left(\int_t^s (T-r)^{\alpha-1} dr \right)^{\frac{1}{2}},$$

(ii) there exists a constant $C^z_{4.1} > 0$ such that for $0 \le t < s < T$ and $x \in \mathbb{R}$,

$$\|Z_s - Z_t\|_{L_p(\mathbb{P}_{t,x})} \le C^z_{4.1} \Psi(x) \left(\int_t^s (T-r)^{\alpha-2} dr \right)^{\frac{1}{2}}.$$

The constants $C^y_{4.1}$ and $C^z_{4.1}$ depend on $(L_f, K_f, C_g, c^{1,2}_{4.2}, \kappa_q, b, \sigma, T, p_0, p)$, and $\Psi(x)$ is defined in (7).

4.3. Properties of the associated PDE

The theorem below collects properties of the solution to the PDE associated to the FBSDE (1). For a proof see [42, Theorem 3.2], [43] and [21, Theorem 5.4].

Theorem 4.2. *Consider the FBSDE (1) and let Assumptions 2.1 and 2.2 hold. Then for the solution u of the associated PDE*

$$\begin{cases} u_t(t,x) + \frac{\sigma^2(t,x)}{2} u_{xx}(t,x) + b(t,x) u_x(t,x) + f(t,x,u(t,x),\sigma(t,x)u_x(t,x)) = 0, \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad t \in [0,T), x \in \mathbb{R}, \\ u(T,x) = g(x), \quad x \in \mathbb{R} \end{cases} \tag{37}$$

we have

(i) $Y_t = u(t,X_t)$ a.s., where $u(t,x) = \mathbb{E}_{t,x}\Big(g(X_T) + \int_t^T f(r,X_r,Y_r,Z_r)dr\Big)$ and $|u(t,x)| \le c^1_{4.2}\Psi(x)$ with Ψ given in (7), where $c^1_{4.2}$ depends on L_f, K_f, C_g, T, p_0 and on the bounds and Lipschitz constants of b and σ.

(ii) (a) $\partial_x u$ exists and is continuous in $[0,T) \times \mathbb{R}$,

(b) $Z^{t,x}_s = u_x(s, X^{t,x}_s)\sigma(s, X^{t,x}_s)$ a.s.,

(c) $|u_x(t,x)| \le \frac{c^2_{4.2}\Psi(x)}{(T-t)^{\frac{1-\alpha}{2}}}$,

where $c^2_{4.2}$ *depends on* $L_f, K_f, C_g, T, p_0, \kappa_2 = \kappa_2(b,\sigma,T,\delta)$ *and on the bounds and Lipschitz constants of* b *and* σ*, and hence* $c^2_{4.2} = c^2_{4.2}(L_f, K_f, C_g, b, \sigma, T, p_0, \delta)$.

(iii) *(a)* $\partial_x^2 u$ *exists and is continuous in* $[0,T)\times\mathbb{R}$,

(b) $|\partial_x^2 u(t,x)| \le \frac{c^3_{4.2}\Psi(x)}{(T-t)^{1-\frac{\alpha}{2}}}$,

where $c^3_{4.2}$ *depends on* $L_f, C_g, T, p_0, \kappa_2 = \kappa_2(b,\sigma,T,\delta), C^y_{4.1}, C^z_{4.1}$ *and on the bounds and Lipschitz constants of* b *and* σ*, and hence* $c^3_{4.2} = c^3_{4.2}(L_f, K_f, C_g, b, \sigma, T, p_0, \delta)$.

Using Assumption 2.3 we are now in the position to improve the bound on $\|Z_s - Z_t\|_{L_p(\mathbb{P}_{t,x})}$ given in Theorem 4.1.

Proposition 4.1. *If Assumption 2.3 holds, then there exists a constant* $C_{4.1} > 0$ *such that for* $0 \le t < s \le T$ *and* $x \in \mathbb{R}$,

$$\|Z_s - Z_t\|_{L_p(\mathbb{P}_{t,x})} \le C_{4.1}\Psi(x)(s-t)^{\frac{1}{2}},$$

where $C_{4.1}$ *depends on* $c^{2,3}_{4.2}, b, \sigma, f, g, T, p_0, p$*, and hence* $C_{4.1} = C_{4.1}(b,\sigma,f,g,T,p_0,p,\delta)$.

Proof. From $Z^{t,x}_s = u_x(s, X^{t,x}_s)\sigma(s, X^{t,x}_s)$ and $\nabla Y^{t,x}_s = \partial_x u(s, X^{t,x}_s) = u_x(s, X^{t,x}_s)\nabla X^{t,x}_s$ we conclude

$$Z^{t,x}_s = \frac{\nabla Y^{t,x}_s}{\nabla X^{t,x}_s}\sigma(s, X^{t,x}_s), \quad 0 \le t \le s \le T. \tag{38}$$

It is well-known (see e.g. [19]) that the solution ∇Y of the linear BSDE

$$\nabla Y_s = g'(X_T)\nabla X_T + \int_s^T f_x(\Theta_r)\nabla X_r + f_y(\Theta_r)\nabla Y_r + f_z(\Theta_r)\nabla Z_r dr - \int_s^T \nabla Z_r dB_r, \quad 0 \le s \le T, \tag{39}$$

can be represented as

$$\begin{aligned}\frac{\nabla Y_s}{\nabla X_s} &= \mathbb{E}_s\Big[g'(X_T)\nabla X_T\Gamma^s_T + \int_s^T f_x(\Theta_r)\nabla X_r\Gamma^s_r dr\Big]\frac{1}{\nabla X_s}\\ &= \tilde{\mathbb{E}}\Big[g'(\tilde{X}^{s,X_s}_T)\nabla\tilde{X}^{s,X_s}_T\tilde{\Gamma}^{s,X_s}_T + \int_s^T f_x(\tilde{\Theta}^{s,X_s}_r)\nabla\tilde{X}^{s,X_s}_r\tilde{\Gamma}^{s,X_s}_r dr\Big], \quad 0 \le t \le s \le T,\end{aligned} \tag{40}$$

where $\Theta_r := (r, X_r, Y_r, Z_r)$ and Γ^s denotes the adjoint process given by

$$\Gamma^s_r = 1 + \int_s^r f_y(\Theta_u)\Gamma^s_u du + \int_s^r f_z(\Theta_u)\Gamma^s_u dB_u, \quad s \le r \le T,$$

and

$$\tilde{\Gamma}^{t,x}_s = 1 + \int_t^s f_y(\tilde{\Theta}^{t,x}_r)\tilde{\Gamma}^{t,x}_r dr + \int_t^s f_z(\tilde{\Theta}^{t,x}_r)\tilde{\Gamma}^{t,x}_r d\tilde{B}_r, \quad t \le s \le T,\ x \in \mathbb{R}$$

where $\tilde{B}$ denotes an independent copy of B. Notice that $\nabla X_t^{t,x} = 1$, so that

$$\frac{\nabla Y_t^{t,x}}{\nabla X_t^{t,x}} = \nabla Y_t^{t,x} = \tilde{\mathbb{E}}\Big[g'(\tilde{X}_T^{t,x})\nabla \tilde{X}_T^{t,x}\tilde{\Gamma}_T^{t,x} + \int_t^T f_x(\tilde{\Theta}_r^{t,x})\nabla \tilde{X}_r^{t,x}\tilde{\Gamma}_r^{t,x}dr\Big].$$

Then, by (38),

$$\|Z_s - Z_t\|_{L_p(\mathbb{P}_{t,x})} \le C(\sigma)\Big[\Big\|\frac{\nabla Y_s}{\nabla X_s} - \frac{\nabla Y_t}{\nabla X_t}\Big\|_{L_p(\mathbb{P}_{t,x})} + \|\nabla Y_t\|_{L_{2p}(\mathbb{P}_{t,x})}[(s-t)^{\frac{1}{2}} + \|X_s^{t,x} - x\|_{L_{2p}(\mathbb{P}_{t,x})}]\Big].$$

Since $(\nabla Y_s, \nabla Z_s)$ is the solution to the linear BSDE (39) with bounded f_x, f_y, f_z, we have that $\|\nabla Y_t\|_{L_{2p}(\mathbb{P}_{t,x})} \le C(b,\sigma,f,g,T,p)$. Obviously, $\|X_s^{t,x} - x\|_{L_{2p}(\mathbb{P}_{t,x})} \le C(b,\sigma,T,p)(s-t)^{\frac{1}{2}}$. So it remains to show that

$$\Big\|\frac{\nabla Y_s}{\nabla X_s} - \frac{\nabla Y_t}{\nabla X_t}\Big\|_{L_p(\mathbb{P}_{t,x})} \le C\Psi(x)(s-t)^{\frac{1}{2}}.$$

We intend to use (40) in the following. There is a certain degree of freedom how to connect B and $\tilde{B}$ in order to compute conditional expectations. Here, unlike in (26), we define the processes

$$B'_u = B_{u\wedge s} + \tilde{B}_{u\vee s} - \tilde{B}_s \quad \text{and} \quad B''_u = B_{u\wedge t} + \tilde{B}_{u\vee t} - \tilde{B}_t, \quad u \ge 0$$

as driving Brownian motions for $\frac{\nabla Y_s}{\nabla X_s}$ and $\frac{\nabla Y_t}{\nabla X_t}$, respectively. This will especially simplify the estimate for $\tilde{\mathbb{E}}|\tilde{\Gamma}_T^{s,X_s} - \tilde{\Gamma}_T^{t,x}|^q$ below. From the above relations we get for $(X_s := X_s^{t,x})$

$$\begin{aligned}
\Big\|\frac{\nabla Y_s}{\nabla X_s} - \frac{\nabla Y_t}{\nabla X_t}\Big\|_{L_p(\mathbb{P}_{t,x})} &\le \Big\|\tilde{\mathbb{E}}\Big[g'(\tilde{X}_T^{s,X_s})\nabla \tilde{X}_T^{s,X_s}\tilde{\Gamma}_T^{s,X_s} - g'(\tilde{X}_T^{t,x})\nabla \tilde{X}_T^{t,x}\tilde{\Gamma}_T^{t,x}\Big]\Big\|_p \\
&\quad + \int_t^s \Big\|\tilde{\mathbb{E}}\Big[f_x(\tilde{\Theta}_r^{t,x})\nabla \tilde{X}_r^{t,x}\tilde{\Gamma}_r^{t,x}\Big]\Big\|_p dr \\
&\quad + \Big\|\int_s^T \tilde{\mathbb{E}}\Big[f_x(\tilde{\Theta}_r^{s,X_s})\nabla \tilde{X}_r^{s,X_s}\tilde{\Gamma}_r^{s,X_s} - f_x(\tilde{\Theta}_r^{t,x})\nabla \tilde{X}_r^{t,x}\tilde{\Gamma}_r^{t,x}\Big]dr\Big\|_p \\
&=: J_1 + J_2 + J_3.
\end{aligned}$$

Since g' is Lipschitz continuous and of polynomial growth, the estimate $J_1 \le C(b,\sigma,g,T,p)\Psi(x)(s-t)^{\frac{1}{2}}$ follows by Hölder's inequality and the L_q -boundedness for any $q > 0$ of all the factors, as well as from the estimates for $\tilde{X}_T^{s,X_s} - \tilde{X}_T^{t,x}$ and $\nabla \tilde{X}_T^{s,X_s} - \nabla \tilde{X}_T^{t,x}$ like in Lemma 5.2. For the Γ differences we first apply the inequalities of Hölder and BDG:

$$\begin{aligned}
\tilde{\mathbb{E}}|\tilde{\Gamma}_T^{s,X_s} - \tilde{\Gamma}_T^{t,x}|^q \quad &\le C(T,q)\ \Big[(s-t)^{q-1}\tilde{\mathbb{E}}\int_t^s |f_y(\tilde{\Theta}_r^{s,X_s})\tilde{\Gamma}_r^{s,X_s}|^q dr + \tilde{\mathbb{E}}\Big(\int_t^s |f_z(\tilde{\Theta}_r^{s,X_s})\tilde{\Gamma}_r^{s,X_s}|^2 dr\Big)^{\frac{q}{2}} \\
&\quad + \tilde{\mathbb{E}}\int_s^T |f_y(\tilde{\Theta}_r^{s,X_s})\tilde{\Gamma}_r^{s,X_s} - f_y(\tilde{\Theta}_r^{t,x})\tilde{\Gamma}_r^{t,x}|^q dr \\
&\quad + \tilde{\mathbb{E}}\Big(\int_s^T |f_z(\tilde{\Theta}_r^{s,X_s})\tilde{\Gamma}_r^{s,X_s} - f_z(\tilde{\Theta}_r^{t,x})\tilde{\Gamma}_r^{t,x}|^2 dr\Big)^{\frac{q}{2}}\Big]
\end{aligned}$$

Since f_y and f_z are bounded we have $\tilde{\mathbb{E}}|\tilde{\Gamma}_r^{s,X_s}|^q + \tilde{\mathbb{E}}|\tilde{\Gamma}_r^{t,x}|^q \leq C(f,T,q)$. Similar to (30), since f_x, f_y, f_z are Lipschitz continuous w.r.t. the space variables,

$$\begin{aligned}|f_x(\tilde{\Theta}_r^{s,X_s}) - f_x(\tilde{\Theta}_r^{t,x})| &= |f_x(r,\tilde{X}_r^{s,X_s}, u(r,\tilde{X}_r^{s,X_s}), \sigma(r,\tilde{X}_r^{s,X_s})u_x(r,\tilde{X}_r^{s,X_s})) \\ &\quad - f_x(r,\tilde{X}_r^{t,x}, u(r,\tilde{X}_r^{t,x}), \sigma(r,\tilde{X}_r^{t,x})u_x(r,\tilde{X}_r^{t,x}))| \\ &\leq C(c_{4.2}^{2,3},\sigma,f,T)(1+|\tilde{X}_r^{s,X_s}|^{p_0+1} + |\tilde{X}_r^{t,x}|^{p_0+1})\frac{|\tilde{X}_r^{s,X_s} - \tilde{X}_r^{t,x}|}{(T-r)^{\frac{1}{2}}},\end{aligned}$$

so that Lemma 5.2 yields

$$\tilde{\mathbb{E}}|f_x(\tilde{\Theta}_r^{s,X_s}) - f_x(\tilde{\Theta}_r^{t,x})|^q \leq C(c_{4.2}^{2,3}, b,\sigma,f,T,p_0,q)(1+|X_s|^{p_0+1} + |x|^{p_0+1})^q \frac{|X_s - x|^q + |s-t|^{\frac{q}{2}}}{(T-r)^{\frac{1}{2}}}.$$

The same holds for $|f_y(\tilde{\Theta}_r^{s,X_s}) - f_y(\tilde{\Theta}_r^{t,x})|$ and $|f_z(\tilde{\Theta}_r^{s,X_s}) - f_z(\tilde{\Theta}_r^{t,x})|$. Applying these inequalities and Gronwall's lemma, we arrive at

$$\|\tilde{\mathbb{E}}[\tilde{\Gamma}_T^{s,X_s} - \tilde{\Gamma}_T^{t,x}]\|_p \leq C(c_{4.2}^{2,3}, b,\sigma,f,g,T,p_0,p)\Psi(x)|s-t|^{\frac{1}{2}}$$

for $p > 0$.

For $J_2 \leq C(t-s)$ it is enough to realise that the integrand is bounded. The estimate for J_3 follows similarly to that of J_1. □

4.4. Properties of the solution to the finite difference equation

Recall the definition of $\mathcal{D}_m^n$ given in (14). By (4),

$$X_{t_{m+1}}^{n,t_m,x} = x + hb(t_{m+1},x) + \sqrt{h}\sigma(t_{m+1},x)\varepsilon_{m+1}, \tag{41}$$

so that

$$T_{m+1,\pm}u^n(t_{m+1}, X_{t_{m+1}}^{n,t_m,x}) = u^n(t_{m+1}, x + hb(t_{m+1},x) \pm \sqrt{h}\sigma(t_{m+1},x)). \tag{42}$$

While for the solution to the PDE (37) one can observe in Theorem 4.2 the well-known smoothing property which implies that u is differentiable on $[0,T)\times\mathbb{R}$ even though g is only Hölder continuous, in following proposition, for the solution u^n to the finite difference equation we have to require from g the same regularity as we want for u^n.

Proposition 4.2. *Let Assumption 2.3 hold and assume that u^n is a solution of*

$$\begin{aligned}&u^n(t_m,x) - hf(t_{m+1},x,u^n(t_m,x),\mathcal{D}_{m+1}^n u^n(t_{m+1},X_{t_{m+1}}^{n,t_m,x})) \\ &\quad = \frac{1}{2}[T_{m+1,+}u^n(t_{m+1},X_{t_{m+1}}^{n,t_m,x}) + T_{m+1,-}u^n(t_{m+1},X_{t_{m+1}}^{n,t_m,x})], \quad m=0,\dots,n-1,\end{aligned} \tag{43}$$

with terminal condition $u^n(t_n,x)=g(x)$. *Then, for sufficiently small* h, *the map* $x\mapsto u^n(t_m,x)$ *is* C^2, *and it holds*

$$|u^n(t_m,x)|+|u^n_x(t_m,x)|\le C_{u^n,1}\,\Psi(x),\quad |u^n_{xx}(t_m,x)|\le C_{u^n,2}\,\Psi^2(x)$$

and

$$|u^n_{xx}(t_m,x)-u^n_{xx}(t_m,\bar x)|\le C_{u^n,3}\,(1+|x|^{6p_0+7}+|\bar x|^{6p_0+7})|x-\bar x|^{\alpha},\tag{44}$$

uniformly in $m=0,\dots,n-1$. *The constants* $C_{u^n,1}$, $C_{u^n,2}$ *and* $C_{u^n,3}$ *depend on the bounds of* f,g,b,σ *and their derivatives and on* T *and* p_0.

Proof. **Step 1.** From (43), since g is C^2 and f_y is bounded, for sufficiently small h we conclude by induction (backwards in time) that $u^n_x(t_m,x)$ exists for $m=0,...,n-1$, and that it holds

$$\begin{aligned}u^n_x(t_m,x) \;=\;& hf_x(t_{m+1},x,u^n(t_m,x),\mathcal{D}^n_{m+1}u^n(t_{m+1},X^{n,t_m,x}_{t_{m+1}}))\\ &+hf_y(t_{m+1},x,u^n(t_m,x),\mathcal{D}^n_{m+1}u^n(t_{m+1},X^{n,t_m,x}_{t_{m+1}}))\,u^n_x(t_m,x)\\ &+hf_z(t_{m+1},x,u^n(t_m,x),\mathcal{D}^n_{m+1}u^n(t_{m+1},X^{n,t_m,x}_{t_{m+1}}))\,\partial_x\mathcal{D}^n_{m+1}u^n(t_{m+1},X^{n,t_m,x}_{t_{m+1}})\\ &+\tfrac12\Big(\partial_xT_{m+1,+}u^n(t_{m+1},X^{n,t_m,x}_{t_{m+1}})+\partial_xT_{m+1,-}u^n(t_{m+1},X^{n,t_m,x}_{t_{m+1}})\Big).\end{aligned}$$

Similarly one can show that $u^n_{xx}(t_m,x)$ exists and solves the derivative of the previous equation.

Step 2. As stated in the proof of Proposition 2.1, the finite difference equation (43) is the associated equation to (8) in the sense that we have the representations (20). We will use that $u^n(t_m,x)=Y^{n,t_m,x}_{t_m}$ and exploit the BSDE

$$\begin{aligned}Y^{n,t_m,x}_{t_m} \;=\; g(X^{n,t_m,x}_T)+\int_{(t_m,T]}f(s,X^{n,t_m,x}_{s-},Y^{n,t_m,x}_{s-},Z^{n,t_m,x}_{s-})d[B^n]_s\\ -\int_{(t_m,T]}Z^{n,t_m,x}_{s-}dB^n_s,\end{aligned}\tag{45}$$

where we will drop the superscript t_m,x from now on. For $u^n_x(t_m,x)$ we will consider

$$\begin{aligned}\nabla Y^n_{t_m}:=\partial_xY^n_{t_m} \;=\; g'(X^n_T)\partial_xX^n_T+\int_{(t_m,T]}f_x\partial_xX^n_{s-}+f_y\partial_xY^n_{s-}+f_z\partial_xZ^n_{s-}d[B^n]_s\\ -\int_{(t_m,T]}\partial_xZ^n_{s-}dB^n_s.\end{aligned}\tag{46}$$

Similarly as in the proof of [30, Theorem 3.1] the BSDE (46) can be derived from (45) as a limit of difference quotients w.r.t. x. Notice that the generator of (46) is random but has the same Lipschitz constant and linear growth bound as f. Assumption 2.3 allows us to find a $p_0\ge0$ and a $K>0$ such that

$$|g(x)|+|g'(x)|+|g''(x)|\le K(1+|x|^{p_0+1})=\Psi(x).$$

In order to get estimates simultaneously for (45) and (46) we show the following lemma.

Lemma 4.2. *We fix n and assume a BSDE*

$$\mathsf{Y}_{t_k} \;=\; \xi^n + \int_{(t_k,T]} \mathsf{f}(s,\mathsf{X}_{s-},\mathsf{Y}_{s-},\mathsf{Z}_{s-})d[B^n]_s - \int_{(t_k,T]} \mathsf{Z}_{s-}dB^n_s, \quad m \le k \le n, \tag{47}$$

with $\xi^n = g(X_T^{n,t_m,x})$ or $\xi^n = g'(X_T^{n,t_m,x})\partial_x X_T^{n,t_m,x}$ and $\mathsf{X}_s := X_s^{n,t_m,x}$ or $\mathsf{X}_s := \partial_x X_s^{n,t_m,x}$ such that $\mathsf{f} : \Omega \times [0,T] \times \mathbb{R}^3 \to \mathbb{R}$ is measurable and satisfies

$$\begin{aligned} |\mathsf{f}(\omega,t,x,y,z) - \mathsf{f}(\omega,t,x',y',z')| &\le L_f(|x-x'| + |y-y'| + |z-z'|), \\ |\mathsf{f}(\omega,t,x,y,z)| &\le (K_f + L_f)(1 + |x| + |y| + |z|). \end{aligned} \tag{48}$$

Then for any $p \ge 2$,

(i) $\mathbb{E}|\mathsf{Y}_{t_k}|^p + \frac{\gamma_p}{4}\mathbb{E}\int_{(t_k,T]} |\mathsf{Y}_{s-}|^{p-2}|\mathsf{Z}_{s-}|^2 d[B^n]_s \le C\Psi^p(x)$, *for* $k = m, ..., n$ *and some* $\gamma_p > 0$,

(ii) $\mathbb{E}\sup_{t_m<s\le T} |\mathsf{Y}_{s-}|^p \le C\Psi^p(x)$,

(iii) $\mathbb{E}\Big(\int_{(t_m,T]} |\mathsf{Z}_{s-}|^2 d[B^n]_s\Big)^{\frac{p}{2}} \le C\Psi^p(x)$,

for some constant $C = C(b,\sigma,f,g,T,p,p_0)$.

Proof. (i) By Itô's formula (see [24, Theorem 4.57]) we get for $p \ge 2$

$$\begin{aligned} |\mathsf{Y}_{t_k}|^p \;=\; & |\xi^n|^p - p\int_{(t_k,T]} \mathsf{Y}_{s-}|\mathsf{Y}_{s-}|^{p-2}\mathsf{Z}_{s-}dB^n_s + p\int_{(t_k,T]} \mathsf{Y}_{s-}|\mathsf{Y}_{s-}|^{p-2}\mathsf{f}(s,\mathsf{X}_{s-},\mathsf{Y}_{s-},\mathsf{Z}_{s-})d[B^n]_s \\ & - \sum_{s\in(t_k,T]} [|\mathsf{Y}_s|^p - |\mathsf{Y}_{s-}|^p - p\mathsf{Y}_{s-}|\mathsf{Y}_{s-}|^{p-2}(\mathsf{Y}_s - \mathsf{Y}_{s-})]. \end{aligned} \tag{49}$$

Following the proof of [27, Proposition 2] (which is carried out there in the Lévy process setting but can be done also for martingales with jumps, like B^n) we can use the estimate

$$-\sum_{s\in(t_k,T]} [|\mathsf{Y}_s|^p - |\mathsf{Y}_{s-}|^p - p\mathsf{Y}_{s-}|\mathsf{Y}_{s-}|^{p-2}(\mathsf{Y}_s - \mathsf{Y}_{s-})] \le -\gamma_p \sum_{s\in(t_k,T]} |\mathsf{Y}_{s-}|^{p-2}(\mathsf{Y}_s - \mathsf{Y}_{s-})^2$$

where $\gamma_p > 0$ is computed in [38, Lemma A4]. Since

$$\mathsf{Y}_{t_{\ell+1}} - \mathsf{Y}_{t_{\ell+1}-} = \mathsf{f}(t_{\ell+1},\mathsf{X}_{t_\ell},\mathsf{Y}_{t_\ell},\mathsf{Z}_{t_\ell})h - \mathsf{Z}_{t_\ell}\sqrt{h}\varepsilon_{\ell+1}$$

we have

$$
\begin{aligned}
&-\sum_{s\in(t_k,T]}\left[|\mathsf{Y}_s|^p-|\mathsf{Y}_{s-}|^p-p\mathsf{Y}_{s-}|\mathsf{Y}_{s-}|^{p-2}(\mathsf{Y}_s-\mathsf{Y}_{s-})\right]\\
&\quad\le\quad -\gamma_p\sum_{\ell=k}^{n-1}|\mathsf{Y}_{t_\ell}|^{p-2}\left(\mathsf{f}(t_{\ell+1},\mathsf{X}_{t_\ell},\mathsf{Y}_{t_\ell},\mathsf{Z}_{t_\ell})h-\mathsf{Z}_{t_\ell}\sqrt{h}\varepsilon_{\ell+1}\right)^2\\
&\quad=\quad -\gamma_p\,h\int_{(t_k,T]}|\mathsf{Y}_{s-}|^{p-2}\,\mathsf{f}^2(s,\mathsf{X}_{s-},\mathsf{Y}_{s-},\mathsf{Z}_{s-})d[B^n]_s-\gamma_p\int_{(t_k,T]}|\mathsf{Y}_{s-}|^{p-2}|\mathsf{Z}_{s-}|^2d[B^n]_s\\
&\qquad+2\gamma_p\int_{(t_k,T]}|\mathsf{Y}_{s-}|^{p-2}\,\mathsf{f}(s,\mathsf{X}_{s-},\mathsf{Y}_{s-},\mathsf{Z}_{s-})\mathsf{Z}_{s-}(B^n_s-B^n_{s-})d[B^n]_s.
\end{aligned}
$$

Hence we get from (49)

$$
\begin{aligned}
|\mathsf{Y}_{t_k}|^p\quad\le\quad&|\xi^n|^p-p\int_{(t_k,T]}\mathsf{Y}_{s-}|\mathsf{Y}_{s-}|^{p-2}\mathsf{Z}_{s-}dB^n_s+p\int_{(t_k,T]}\mathsf{Y}_{s-}|\mathsf{Y}_{s-}|^{p-2}\,\mathsf{f}(s,\mathsf{X}_{s-},\mathsf{Y}_{s-},\mathsf{Z}_{s-})d[B^n]_s\\
&-\gamma_p\int_{(t_k,T]}|\mathsf{Y}_{s-}|^{p-2}\,|\mathsf{Z}_{s-}|^2d[B^n]_s\\
&+2\gamma_p\int_{(t_k,T]}|\mathsf{Y}_{s-}|^{p-2}\,\mathsf{f}(s,\mathsf{X}_{s-},\mathsf{Y}_{s-},\mathsf{Z}_{s-})\mathsf{Z}_{s-}(B^n_s-B^n_{s-})d[B^n]_s.
\end{aligned}
$$

From Young's inequality and (48) we conclude that there is a $c'=c'(p,K_f,L_f,\gamma_p)>0$ such that

$$
p|\mathsf{Y}_{s-}|^{p-1}\,|\mathsf{f}(s,\mathsf{X}_{s-},\mathsf{Y}_{s-},\mathsf{Z}_{s-})|\le\tfrac{\gamma_p}{4}|\mathsf{Y}_{s-}|^{p-2}\,|\mathsf{Z}_{s-}|^2+c'(1+|\mathsf{X}_{s-}|^p+|\mathsf{Y}_{s-}|^p),
$$

and for $\sqrt{h}<\frac{1}{8(L_f+K_f)}$ we find a $c''=c''(p,L_f,K_f,\gamma_p)>0$ such that

$$
2\gamma_p\sqrt{h}|\mathsf{Y}_{s-}|^{p-2}\,|\mathsf{f}(s,\mathsf{X}_{s-},\mathsf{Y}_{s-},\mathsf{Z}_{s-})||\mathsf{Z}_{s-}|\le\tfrac{\gamma_p}{4}|\mathsf{Y}_{s-}|^{p-2}\,|\mathsf{Z}_{s-}|^2+c''\,(1+|\mathsf{X}_{s-}|^p+|\mathsf{Y}_{s-}|^p).
$$

Then for $c=c'+c''$ we have

$$
\begin{aligned}
|\mathsf{Y}_{t_k}|^p\quad\le\quad&|\xi^n|^p-p\int_{(t_k,T]}\mathsf{Y}_{s-}|\mathsf{Y}_{s-}|^{p-2}\,\mathsf{Z}_{s-}dB^n_s+c\int_{(t_k,T]}1+|\mathsf{X}_{s-}|^p+|\mathsf{Y}_{s-}|^pd[B^n]_s\\
&-\tfrac{\gamma_p}{2}\int_{(t_k,T]}|\mathsf{Y}_{s-}|^{p-2}\,|\mathsf{Z}_{s-}|^2d[B^n]_s.
\end{aligned}\tag{50}
$$

By standard methods, approximating the terminal condition and the generator by bounded functions, it follows that for any $a>0$

$$
\mathbb{E}\sup_{t_k\le s\le T}|\mathsf{Y}_s|^a<\infty\quad\text{ and }\quad\mathbb{E}\left(\int_{(t_k,T]}|\mathsf{Z}_{s-}|^2d[B^n]_s\right)^{\frac{a}{2}}<\infty.
$$

Hence $\int_{(t_k,T]}\mathsf{Y}_{s-}|\mathsf{Y}_{s-}|^{p-2}\mathsf{Z}_{s-}dB^n_s$ has expectation zero. Taking the expectation in (50) yields

$$
\mathbb{E}|\mathsf{Y}_{t_k}|^p+\tfrac{\gamma_p}{2}\mathbb{E}\int_{(t_k,T]}|\mathsf{Y}_{s-}|^{p-2}|\mathsf{Z}_{s-}|^2d[B^n]_s\le\mathbb{E}|\xi^n|^p+c\mathbb{E}\int_{(t_k,T]}1+|\mathsf{X}_{s-}|^p+|\mathsf{Y}_{s-}|^pd[B^n]_s.\tag{51}
$$

Since $\mathbb{E}|\xi^n|^p$ and $\mathbb{E}\int_{(t_k,T]} 1+|\mathsf{X}_{s-}|^p d[B^n]_s$ are polynomially bounded in x, Gronwall's lemma gives

$$\|\mathsf{Y}_{t_k}\|_p \le C(b,\sigma,f,g,T,p,p_0)(1+|x|^{p_0+1}), \quad k=m,...,n,$$

and inserting this into (51) yields

$$\Big(\mathbb{E}\int_{(t_k,T]} |\mathsf{Y}_{s-}|^{p-2}|\mathsf{Z}_{s-}|^2 d[B^n]_s\Big)^{\frac{1}{p}} \le C(b,\sigma,f,g,T,p,p_0)(1+|x|^{p_0+1}), \quad k=m,...,n-1.$$

(ii) From (50) we derive by the inequality of BDG and Young's inequality that for $t_m \le t_k \le T$

$$\begin{aligned}
&\mathbb{E}\sup_{t_k<s\le T}|\mathsf{Y}_{s-}|^p \\
&\le \mathbb{E}|\xi^n|^p + C(p)\mathbb{E}\left(\int_{(t_k,T]}|\mathsf{Y}_{s-}|^{2p-2}|\mathsf{Z}_{s-}|^2 d[B^n]_s\right)^{\frac12} + c\mathbb{E}\int_{(t_k,T]} 1+|\mathsf{X}_{s-}|^p+|\mathsf{Y}_{s-}|^p d[B^n]_s \\
&\le \mathbb{E}|\xi^n|^p + c\mathbb{E}\int_{(t_k,T]} 1+|\mathsf{X}_{s-}|^p d[B^n]_s + C(p)\mathbb{E}\left[\sup_{t_k<s\le T}|\mathsf{Y}_{s-}|^{\frac{p}{2}}\left(\int_{(t_k,T]}|\mathsf{Y}_{s-}|^{p-2}|\mathsf{Z}_{s-}|^2 d[B^n]_s\right)^{\frac12}\right] \\
&\quad + c\mathbb{E}\int_{(t_k,T]}|\mathsf{Y}_{s-}|^p d[B^n]_s \\
&\le \mathbb{E}|\xi^n|^p + c\mathbb{E}\int_{(t_k,T]} 1+|\mathsf{X}_{s-}|^p d[B^n]_s + C(p)\mathbb{E}\int_{(t_k,T]}|\mathsf{Y}_{s-}|^{p-2}|\mathsf{Z}_{s-}|^2 d[B^n]_s \\
&\quad + \mathbb{E}\sup_{t_k<s\le T}|\mathsf{Y}_{s-}|^p\big(\tfrac14 + c(T-t_k)\big).
\end{aligned}$$

We assume that h is sufficiently small so that we find a t_k with $c(T-t_k)<\frac14$. We rearrange the inequality to have $\mathbb{E}\sup_{t_k<s\le T}|\mathsf{Y}_{s-}|^p$ on the l.h.s., and from (i) we conclude that

$$\begin{aligned}
\mathbb{E}\sup_{t_k<s\le T}|\mathsf{Y}_{s-}|^p &\le 2\mathbb{E}|\xi^n|^p + 2c\mathbb{E}\int_{(t_k,T]} 1+|\mathsf{X}_{s-}|^p d[B^n]_s + 2C(p)\mathbb{E}\int_{(t_k,T]}|\mathsf{Y}_{s-}|^{p-2}|\mathsf{Z}_{s-}|^2 d[B^n]_s \\
&\le C(b,\sigma,f,g,T,p,p_0)(1+|x|^{(p_0+1)p}).
\end{aligned}$$

Now we may repeat the above step for $\mathbb{E}\sup_{t_\ell<s\le t_k}|\mathsf{Y}_{s-}|^p$ with $c(t_k-t_\ell)<\frac14$ and $\xi^n=\mathsf{Y}_T$ replaced by Y_{t_k}, and continue doing so until we eventually get assertion (ii).

(iii) We proceed from (47),

$$\sup_{k\le\ell\le n}\Big|\int_{(t_\ell,T]}\mathsf{Z}_{s-}dB^n_s\Big|^p \le C(p)\Big(|\xi^n|^p + \sup_{k\le\ell\le n}|\mathsf{Y}_{t_\ell}|^p + \Big(\int_{(t_k,T]}|\mathsf{f}(s,\mathsf{X}_{s-},\mathsf{Y}_{s-},\mathsf{Z}_{s-})|\,d[B^n]_s\Big)^p\Big),$$

so that by (48) and the inequalities of BDG and Hölder we have that

$$\begin{aligned}
&\mathbb{E}\Big(\int_{(t_k,T]}|\mathsf{Z}_{s-}|^2 d[B^n]_s\Big)^{\frac{p}{2}} \\
&\le C(p)\Big(\mathbb{E}|\xi^n|^p + \mathbb{E}\sup_{k\le\ell\le n}|\mathsf{Y}_{t_\ell}|^p\Big) + C(p,L_f,K_f)\mathbb{E}\left(\int_{(t_k,T]} 1+|\mathsf{X}_{s-}|+|\mathsf{Y}_{s-}|d[B^n]_s\right)^p \\
&\quad + C(p,L_f,K_f)(T-t_k)^{\frac{p}{2}}\mathbb{E}\left(\int_{(t_k,T]}|\mathsf{Z}_{s-}|^2 d[B^n]_s\right)^{\frac{p}{2}}.
\end{aligned}$$

Hence for $C(p, L_f, K_f)(T-t_k)^{\frac{p}{2}} < \frac{1}{2}$ we derive from assertion (ii) and from the growth properties of the other terms that

$$\mathbb{E}\Big(\int_{(t_k,T]} |\mathsf{Z}_{s-}|^2 d[B^n]_s\Big)^{\frac{p}{2}} \leq C(b,\sigma,f,g,T,p,p_0)(1+|x|^{(p_0+1)p}). \tag{52}$$

Repeating this procedure eventually yields (iii). □

Step 3. Applying Lemma 4.2 to (45) and (46) we see that for all $m = 0, ..., n$ we have

$$|u^n(t_m,x)| = |Y^{n,t_m,x}_{t_m}| = (\mathbb{E}(Y^{n,t_m,x}_{t_m})^2)^{\frac{1}{2}} \leq C(b,\sigma,f,g,T,p_0)(1+|x|^{p_0+1})$$

and

$$|u^n_x(t_m,x)| = (\mathbb{E}(\partial_x Y^{n,t_m,x}_{t_m})^2)^{\frac{1}{2}} \leq C(b,\sigma,f,g,T,p_0)(1+|x|^{p_0+1}). \tag{53}$$

Our next aim is to show that $u^n_{xx}(t_m,x)$ is locally Lipschitz in x. We first show that $u^n_{xx}(t_m,x)$ has polynomial growth. We introduce the BSDE which describes $u^n_{xx}(t_m,x)$ and denote for simplicity

$$f(t,x_1,x_2,x_3) := f(t,x,y,z) \quad \text{and} \quad D^a := \partial^{i_1}_{x_1}\partial^{i_2}_{x_2}\partial^{i_3}_{x_3} \quad \text{with} \quad a := (i_1,i_2,i_3)$$

and consider

$$\begin{aligned}
\partial^2_x Y^n_{t_m} &= g''(X^n_T)(\partial_x X^n_T)^2 + g'(X^n_T)\partial^2_x X^n_T \\
&\quad + \int_{(t_m,T]} \sum_{\substack{a\in\{0,1,2\}^3 \\ i_1+i_2+i_3=2}} (D^a f)(s, X^n_{s-}, Y^n_{s-}, Z^n_{s-})(\partial_x X^n_{s-})^{i_1}(\partial_x Y^n_{s-})^{i_2}(\partial_x Z^n_{s-})^{i_3} d[B^n]_s \\
&\quad + \int_{(t_m,T]} \sum_{\substack{a\in\{0,1\}^3 \\ i_1+i_2+i_3=1}} (D^a f)(s, X^n_{s-}, Y^n_{s-}, Z^n_{s-})(\partial^2_x X^n_{s-})^{i_1}(\partial^2_x Y^n_{s-})^{i_2}(\partial^2_x Z^n_{s-})^{i_3} d[B^n]_s \\
&\quad - \int_{(t_m,T]} \partial^2_x Z^n_{s-} dB^n_s.
\end{aligned} \tag{54}$$

We denote the generator of this BSDE by $\hat{f}$ and notice that it is of the structure

$$\hat{f}(\omega,t,x,y,z) = f_0(\omega,t) + f_1(\omega,t)x + f_2(\omega,t)y + f_3(\omega,t)z.$$

Here $f_0(\omega,t)$ denotes the integrand of the first integral on the r.h.s of (54), and from the previous results one concludes that $\mathbb{E}(\int_{(t_m,T]} |f_0(s-)|d[B^n]_s)^p < \infty$. The functions $f_1(t) = (D^{(1,0,0)}f)(t,\cdot) = (\partial_x f)(t,\cdot)$ as well as $f_2(t) = (\partial_y f)(t,\cdot)$ and $f_3(t) = (\partial_z f)(t,\cdot)$ are bounded by our assumptions. We put

$$\hat{\xi}^n := g''(X^n_T)(\partial_x X^n_T)^2 + g'(X^n_T)\partial^2_x X^n_T.$$

Denoting the solution by $(\hat{\mathsf{Y}}, \hat{\mathsf{Z}})$ we get for $C(f_3)(T-t_m) \le \frac{1}{2}$ that

$$\mathbb{E}|\hat{\mathsf{Y}}_{t_m}|^2 + \frac{1}{2}\mathbb{E}\int_{(t_m,T]} |\hat{\mathsf{Z}}_{s-}|^2 d[B^n]_s \\ \le C\Big[\mathbb{E}|\hat{\xi}^n|^2 + \mathbb{E}\Big(\int_{(t_m,T]} |f_0(s-)|d[B^n]_s\Big)^2 + \mathbb{E}\int_{(t_m,T]} |\hat{\mathsf{X}}_{s-}|^2 + |\hat{\mathsf{Y}}_{s-}|^2 d[B^n]_s\Big]. \tag{55}$$

Now we derive the polynomial growth $\mathbb{E}|\hat{\xi}^n|^2 \le C\Psi^2(x)$ from the properties of g' and g'' and from the fact that $\mathbb{E}\sup_{t_m<s\le T}|\partial_x^j X_s^n|^p$ is bounded for $j=1,2$ under our assumptions. Then the estimate

$$\mathbb{E}\Big(\int_{(t_m,T]} |f_0(s-)|d[B^n]_s\Big)^2 \le C\Psi^4(x)$$

can be derived from Lemma 4.2(ii)-(iii), so that Gronwall's lemma implies

$$|\hat{\mathsf{Y}}_{t_m}^{t_m,x}| = |u_{xx}(t_m,x)| \le C\Psi^2(x). \tag{56}$$

Finally, to show (44), one uses (54) and derives an inequality as in (55) but now for the difference $\partial_x^2 Y_{t_m}^{n,t_m,x} - \partial_x^2 Y_{t_m}^{n,t_m,\bar{x}}$.

Before proving it, let us state the following lemma.

Lemma 4.3. *Let Assumption 2.3 hold. We have*

$$\Big(\mathbb{E}\sup_s |Z_{s-}^{n,t_m,x} - Z_{s-}^{n,t_m,\bar{x}}|^p\Big)^{1/p} \le C(\Psi(x)^2 + \Psi(\bar{x})^2)|x-\bar{x}|, \quad p \ge 2, \tag{57}$$

$$\mathbb{E}\Big(\int_{(t_m,T]} |\partial_x Z_{s-}^{n,t_m,x} - \partial_x Z_{s-}^{n,t_m,\bar{x}}|^2 d[B^n]_s\Big)^{\frac{p}{2}} \le C(\Psi^{4p}(x) + \Psi^{4p}(\bar{x}))|x-\bar{x}|^p, \quad p \ge 2, \tag{58}$$

$$\mathbb{E}\left(\int_{(t_m,T]} |\partial_x^2 Z_{s-}^{n,t_m,x}|^2 d[B^n]_s\right)^{\frac{p}{2}} \le C\Psi^{4p}(x), \quad p \ge 2, \tag{59}$$

for some constant $C = C(b,\sigma,f,g,T,p,p_0)$.

Proof of Lemma 4.3. (57): Introduce $G(t_{k+1},x) := \mathcal{D}_{k+1}^n u^n(t_{k+1}, X_{t_{k+1}}^{n,t_k,x})$. Using relations (41)–(42) and the bounds (53) and (56) for u_x^n and u_{xx}^n, respectively, one obtains

$$|G(t_{k+1},x) - G(t_{k+1},\bar{x})| \le C(1 + |x|^{2(p_0+1)} + |\bar{x}|^{2(p_0+1)})|x-\bar{x}|, \quad x,\bar{x} \in \mathbb{R},$$

uniformly in t_{k+1}. Since $Z_{t_k}^{n,t_m,x} = \mathcal{D}_{k+1}^n u^n(t_{k+1}, X_{t_{k+1}}^{n,t_k,\eta}) = G(t_{k+1},\eta)$ where $\eta = X_{t_k}^{n,t_m,x}$, the previous bound yields

$$|Z_{t_k}^{n,t_m,x} - Z_{t_k}^{n,t_m,\bar{x}}| \le C(1 + |X_{t_k}^{n,t_m,x}|^{2(p_0+1)} + |X_{t_k}^{n,t_m,\bar{x}}|^{2(p_0+1)})|X_{t_k}^{n,t_m,x} - X_{t_k}^{n,t_m,\bar{x}}|$$

uniformly for each $t_m \le t_k < T$. Inequality (57) then follows by applying the Cauchy-Schwarz inequality and standard L_p-estimates for the process X^n.

(58): This can be shown similarly as Lemma 4.2-(iii) considering the BSDE for the difference $\partial_x Y^{n,t_m,x}_{t_m} - \partial_x Y^{n,t_m,\bar{x}}_{t_m}$ instead of (46) itself.

(59): This one gets repeating again the proof of Lemma 4.2-(iii) but now for the BSDE (54). □

By our assumptions we have

$$\mathbb{E}|\hat{\xi}^{n,t_m,x} - \hat{\xi}^{n,t_m,\bar{x}}|^2 \le C(\Psi^2(x) + \Psi^2(\bar{x}))(1 + |x|^2 + |\bar{x}|^2)|x - \bar{x}|^{2\alpha},$$

where we use $|x - \bar{x}|^2 \le C(1 + |x|^2 + |\bar{x}|^2)|x - \bar{x}|^{2\alpha}$. The term $|x - \bar{x}|^2$ appears for example in the estimate of $(\partial_x X^{n,t_m,x}_T)^2 - (\partial_x X^{n,t_m,\bar{x}}_T)^2$. To see that

$$\mathbb{E}\Big(\int_{(t_m,T]} |f_0^{t_m,x}(s-) - f_0^{t_m,\bar{x}}(s-)| d[B^n]_s \Big)^2 \le C(\Psi^{10}(x) + \Psi^{10}(\bar{x}))(1 + |x|^2 + |\bar{x}|^2)|x - \bar{x}|^{2\alpha},$$

we check the terms with the highest polynomial growth. For example, we have to deal with terms like $\mathbb{E}\Big(\int_{(t_m,T]} |Z^{n,t_m,x}_{s-} - Z^{n,t_m,\bar{x}}_{s-}|\, |\partial_x Z^{n,t_m,x}_{s-}|^2 d[B^n]_s \Big)^2$ and $\mathbb{E}\Big(\int_{(t_m,T]} |\partial_x Z^{n,t_m,x}_{s-}|^2 - |\partial_x Z^{n,t_m,\bar{x}}_{s-}|^2 d[B^n]_s \Big)^2$. We bound the first term by using (52) and (57)

$$\begin{aligned}
&\mathbb{E}\Big(\int_{(t_m,T]} |Z^{n,t_m,x}_{s-} - Z^{n,t_m,\bar{x}}_{s-}|\, |\partial_x Z^{n,t_m,x}_{s-}|^2 d[B^n]_s \Big)^2 \\
&\quad \le \quad (\mathbb{E} \sup_s |Z^{n,t_m,x}_{s-} - Z^{n,t_m,\bar{x}}_{s-}|^4)^{\frac{1}{2}} \Big(\mathbb{E}\Big(\int_{(t_m,T]} |\partial_x Z^{n,t_m,x}_{s-}|^2 d[B^n]_s \Big)^4 \Big)^{\frac{1}{2}} \\
&\quad \le \quad C(\Psi^4(x) + \Psi^4(\bar{x}))|x - \bar{x}|^2 \Psi^4(x).
\end{aligned}$$

We bound the second term by using (52) and (58)

$$\begin{aligned}
&\mathbb{E}\Big(\int_{(t_m,T]} |\partial_x Z^{n,t_m,x}_{s-}|^2 - |\partial_x Z^{n,t_m,\bar{x}}_{s-}|^2 d[B^n]_s \Big)^2 \\
&\quad \le \quad C\mathbb{E} \int_{(t_m,T]} |\partial_x Z^{n,t_m,x}_{s-}|^2 + |\partial_x Z^{n,t_m,\bar{x}}_{s-}|^2 d[B^n]_s \int_{(t_m,T]} |\partial_x Z^{n,t_m,x}_{s-} - \partial_x Z^{n,t_m,\bar{x}}_{s-}|^2 d[B^n]_s \\
&\quad \le \quad C(\Psi^2(x) + \Psi^2(\bar{x}))(\Psi^8(x) + \Psi^8(\bar{x}))|x - \bar{x}|^2 \\
&\quad \le \quad C(\Psi^{10}(x) + \Psi^{10}(\bar{x}))(|x|^{2-2\alpha} + |\bar{x}|^{2-2\alpha})|x - \bar{x}|^{2\alpha} \\
&\quad \le \quad C(\Psi^{10}(x) + \Psi^{10}(\bar{x}))(1 + |x|^2 + |\bar{x}|^2)|x - \bar{x}|^{2\alpha},
\end{aligned}$$

While all the other terms can be easily estimated using the results we have obtained already, for

$$\mathbb{E}\Big(\int_{(t_m,T]} |(f_3^{t_m,x}(s-) - f_3^{t_m,\bar{x}}(s-))\partial_x^2 Z^{n,t_m,x}_{s-}| d[B^n]_s \Big)^2 \le C(\Psi^{12}(x) + \Psi^{12}(\bar{x}))(1 + |x|^2 + |\bar{x}|^2)|x - \bar{x}|^{2\alpha}$$

we need the bound (59).

The result follows then from Gronwall's lemma.

Remark 4.1. Under Assumption 2.3 we conclude that by Proposition 4.2 there exists a constant $C = C(b, \sigma, f, g, T, p, p_0) > 0$ such that

$$\begin{aligned}
|u^n(t_m, x) - u^n(t_m, \bar{x})| &\leq C(1 + \Psi(x) + \Psi(\bar{x}))|x - \bar{x}|, \\
|\mathcal{D}^n_{m+1} u^n(t_{m+1}, X^{n,t_m,x}_{t_{m+1}}) - \mathcal{D}^n_{m+1} u^n(t_{m+1}, X^{n,t_m,x}_{t_{m+1}})| &\leq C(1 + \Psi^2(x) + \Psi^2(\bar{x}))|x - \bar{x}|, \\
|u^n_x(t_m, x) - u^n_x(t_m, \bar{x})| &\leq C(1 + \Psi^2(x) + \Psi^2(\bar{x}))|x - \bar{x}|, \\
|\partial_x \mathcal{D}^n_{m+1} u^n(t_{m+1}, X^{n,t_m,x}_{t_{m+1}}) - \partial_x \mathcal{D}^n_{m+1} u^n(t_{m+1}, X^{n,t_m,\bar{x}}_{t_{m+1}})| &\leq C(1 + \hat{\Psi}(x) + \hat{\Psi}(\bar{x}))|x - \bar{x}|^\alpha, \\
|\partial_x \mathcal{D}^n_{m+1} u^n(t_{m+1}, X^{n,t_m,x}_{t_{m+1}})| &\leq C(1 + \Psi^2(x)), \qquad (60)
\end{aligned}$$

uniformly in $m = 0, 1, \ldots, n-1$, where

$$\hat{\Psi}(x) := 1 + |x|^{6p_0+8}. \qquad (61)$$

In addition, for $\partial_x F^n(t_{m+1}, x) := \partial_x f(t_{m+1}, x, u^n(t_m, x), \mathcal{D}^n_{m+1} u^n(t_{m+1}, X^{n,t_m,x}_{t_{m+1}}))$ we have

$$|\partial_x F^n(t_{m+1}, x) - \partial_x F^n(t_{m+1}, \bar{x})| \leq C(1 + \hat{\Psi}(x) + \hat{\Psi}(\bar{x}))|x - \bar{x}|^\alpha \qquad (62)$$

uniformly in $m = 0, 1, \ldots, n-1$. The latter inequality follows from the assumption that the partial derivatives of f are bounded and Lipschitz continuous w.r.t. the spatial variables, from estimates proved in Proposition 4.2 and from those stated in (60) above.

From the calculations it can be seen that in general Assumption 2.3 can not be weakened if one needs $\partial_x F^n(t_{m+1}, x)$ to be locally α-Hölder continuous.

□

5. Technical results and estimates

In this section we collect some facts which are needed for the proofs of our results. We start with properties of the stopping times used to construct a random walk.

Lemma 5.1. (Proposition 11.1 [39], Lemma A.1 [21].) *For all $0 \leq k \leq m \leq n$ and $p > 0$, it holds for $h = \frac{T}{n}$ and τ_k defined in (23) that*

(i) $\mathbb{E}\tau_k = kh$,

(ii) $\mathbb{E}|\tau_1|^p \leq C(p)h^p$,

(iii) $\mathbb{E}|B_{\tau_k} - B_{t_k}|^{2p} \leq C(p)\mathbb{E}|\tau_k - t_k|^p \leq C(p)(t_k h)^{\frac{p}{2}}$.

The next lemma lists some estimates concerning the diffusion X defined by (27) and its discretization (25), where we assume that B and $\tilde{B}$ are connected as in (26).

Lemma 5.2. *Under Assumption 2.1 on b and σ it holds for $p \geq 2$ that there exists a constant $C = C(b, \sigma, T, p) > 0$ such that*

(i) $\mathbb{E}\big|X_T^{s,y} - X_T^{t,x}\big|^p \leq C(|y-x|^p + |s-t|^{\frac{p}{2}}), \quad x, y \in \mathbb{R},\ s, t \in [0,T],$

(ii) $\tilde{\mathbb{E}} \sup_{\tilde{\tau}_l \wedge t_m \leq r \leq \tilde{\tau}_{l+1} \wedge t_m} |\tilde{X}_{t_k+r}^{t_k,x} - \tilde{X}_{t_k+\tilde{\tau}_l \wedge t_m}^{t_k,x}|^p \leq C h^{\frac{p}{4}},\ 0 \leq k \leq n,\ 0 \leq l \leq n-k-1,\ 0 \leq m \leq n-k,$

(iii) $\mathbb{E}|\nabla X_T^{s,y} - \nabla X_T^{t,x}|^p \leq C(|y-x|^p + |s-t|^{\frac{p}{2}}), \quad x, y \in \mathbb{R},\ s, t \in [0,T],$

(iv) $\mathbb{E} \sup_{0 \leq l \leq m} \big|\nabla X_{t_k+t_l}^{n,t_k,x}\big|^p \leq C, \quad 0 \leq k \leq n,\ 0 \leq m \leq n-k,$

(v) $\tilde{\mathbb{E}}\big|\tilde{X}_{t_k+t_m}^{t_k,x} - \tilde{\mathcal{X}}_{\tau_k+\tilde{\tau}_m}^{\tau_k,y}\big|^p \leq C(|x-y|^p + h^{\frac{p}{4}}), \quad 0 \leq k \leq n,\ 0 \leq m \leq n-k,$

(vi) $\tilde{\mathbb{E}}|\nabla \tilde{X}_{t_k+t_m}^{t_k,x} - \nabla \tilde{\mathcal{X}}_{\tau_k+\tilde{\tau}_m}^{\tau_k,y}|^p \leq C(|x-y|^p + h^{\frac{p}{4}}), \quad 0 \leq k \leq n,\ 0 \leq m \leq n-k.$

Proof. (i): This estimate is well-known.

(ii): For the stochastic integral we use the inequality of BDG and then, since b and σ are bounded, we get by Lemma 5.1 (ii) that

$$\begin{aligned}\tilde{\mathbb{E}} \sup_{\tilde{\tau}_l \wedge t_m \leq r \leq \tilde{\tau}_{l+1} \wedge t_m} & |\tilde{X}_{t_k+r}^{t_k,x} - \tilde{X}_{t_k+\tilde{\tau}_l \wedge t_m}^{t_k,x}|^p \\ & \leq \quad C(p)(\|b\|_\infty^p \tilde{\mathbb{E}}|\tilde{\tau}_{l+1} - \tilde{\tau}_l|^p + \|\sigma\|_\infty^p \mathbb{E}|\tilde{\tau}_{l+1} - \tilde{\tau}_l|^{\frac{p}{2}}) \leq C(b, \sigma, T, p)\, h^{\frac{p}{2}}.\end{aligned}$$

(iii): This can be easily seen because the process $(\nabla X_r^{s,y})_{r \in [s,T]}$ solves the linear SDE (12) with bounded coefficients.

(iv): The process solves (64). The estimate follows from the inequality of BDG and Gronwall's lemma.

(v): Recall that from (4) and (25) we have

$$\tilde{\mathcal{X}}_{\tau_k+\tilde{\tau}_m}^{\tau_k,y} = \tilde{X}_{t_k+t_m}^{n,t_k,y} = y + \int_{(0,t_m]} b(t_k + r, \tilde{X}_{t_k+r-}^{n,t_k,y}) d[\tilde{B}^n, \tilde{B}^n]_r + \int_{(0,t_m]} \sigma(t_k + r, \tilde{X}_{t_k+r-}^{n,t_k,y}) d\tilde{B}_r^n,$$

and $\tilde{X}_{t_k+t_m}^{t_k,x}$ is given by

$$\tilde{X}_{t_k+t_m}^{t_k,x} = x + \int_0^{t_m} b(t_k + r, \tilde{X}_{t_k+r}^{t_k,x}) dr + \int_0^{t_m} \sigma(t_k + r, \tilde{X}_{t_k+r}^{t_k,y}) d\tilde{B}_r.$$

To compare the stochastic integrals of the previous two equations we use the relation

$$\int_{(0,t_m]} \sigma(t_k + r, \tilde{X}_{t_k+r-}^{n,t_k,y}) d\tilde{B}_r^n = \int_0^\infty \sum_{l=0}^{m-1} \sigma(t_{k+l+1}, \tilde{X}_{t_{k+l}}^{n,t_k,y}) \mathbf{1}_{(\tilde{\tau}_l, \tilde{\tau}_{l+1}]}(r) d\tilde{B}_r.$$

We define an 'increasing' map $i(r) := t_{l+1}$ for r in $(t_l, t_{l+1}]$ and a 'decreasing' map $d(r) := t_l$ for $(t_l, t_{l+1}]$ and split the differences as follows (using Assumption 2.1-(iii) for the coefficient b)

$$\begin{aligned}
&\tilde{\mathbb{E}}\big|\tilde{X}^{t_k,x}_{t_k+t_m} - \tilde{X}^{n,t_k,y}_{t_k+t_m}\big|^p \\
&\quad\leq\ C(b,p)\left(|x-y|^p + \tilde{\mathbb{E}}\int_0^{t_m} |r-i(r)|^{\frac{p}{2}} + |\tilde{X}^{t_k,x}_{t_k+r} - \tilde{X}^{t_k,x}_{t_k+d(r)}|^p + |\tilde{X}^{t_k,x}_{t_k+d(r)} - \tilde{X}^{n,t_k,y}_{t_k+d(r)}|^p dr\right) \\
&\qquad +C(p)\tilde{\mathbb{E}}|\int_{t_m\wedge\tilde{\tau}_m}^{t_m} \sigma(t_k+r, \tilde{X}^{t_k,x}_{t_k+r})d\tilde{B}_r|^p \\
&\qquad +C(p)\tilde{\mathbb{E}}|\int_{t_m\wedge\tilde{\tau}_m}^{\tilde{\tau}_m} \sum_{l=0}^{m-1} \sigma(t_{k+l+1}, \tilde{X}^{n,t_k,y}_{t_{k+l}})\mathbf{1}_{(\tilde{\tau}_l,\tilde{\tau}_{l+1}]}(r)d\tilde{B}_r|^p \\
&\qquad +C(p)\tilde{\mathbb{E}}|\int_0^{t_m\wedge\tilde{\tau}_m} \sigma(t_k+r, \tilde{X}^{t_k,x}_{t_k+r}) - \sum_{l=0}^{m-1} \sigma(t_{k+l+1}, \tilde{X}^{n,t_k,y}_{t_{k+l}})\mathbf{1}_{(\tilde{\tau}_l,\tilde{\tau}_{l+1}]}(r)d\tilde{B}_r|^p. \qquad (63)
\end{aligned}$$

We estimate the terms on the r.h.s as follows: by standard estimates for SDEs with bounded coefficients one has that

$$\tilde{\mathbb{E}}\int_0^{t_m} |r-i(r)|^{\frac{p}{2}} + |\tilde{X}^{t_k,x}_{t_k+r} - \tilde{X}^{t_k,x}_{t_k+d(r)}|^p dr \leq C(b,\sigma,T,p)h^{\frac{p}{2}}.$$

By the BDG inequality, the fact that σ is bounded and Lemma 5.1 we conclude that

$$\begin{aligned}
&\tilde{\mathbb{E}}\left|\int_{t_m\wedge\tilde{\tau}_m}^{t_m} \sigma(t_k+r, \tilde{X}^{t_k,x}_{t_k+r})d\tilde{B}_r\right|^p + \tilde{\mathbb{E}}\left|\int_{t_m\wedge\tilde{\tau}_m}^{\tilde{\tau}_m} \sum_{l=0}^{m-1} \sigma(t_{k+l+1}, \tilde{X}^{n,t_k,y}_{t_{k+l}})\mathbf{1}_{(\tilde{\tau}_l,\tilde{\tau}_{l+1}]}(r)d\tilde{B}_r\right|^p \\
&\qquad \leq C(\sigma,p)\|\sigma\|_\infty^p \tilde{\mathbb{E}}|\tilde{\tau}_m - t_m|^{\frac{p}{2}} \leq C(\sigma,p)(t_m h)^{\frac{p}{4}}.
\end{aligned}$$

Finally, by the BDG inequality

$$\begin{aligned}
&\tilde{\mathbb{E}}\left|\int_0^{t_m\wedge\tilde{\tau}_m} \sigma(t_k+r, \tilde{X}^{t_k,x}_{t_k+r}) - \sum_{l=0}^{m-1} \sigma(t_{k+l+1}, \tilde{X}^{n,t_k,y}_{t_{k+l}})\mathbf{1}_{(\tilde{\tau}_l,\tilde{\tau}_{l+1}]}(r)d\tilde{B}_r\right|^p \\
&\quad\leq\ C(p)\tilde{\mathbb{E}}\Big(\int_0^{t_m} \sum_{l=0}^{m-1} |\sigma(t_k+r, \tilde{X}^{t_k,x}_{t_k+r}) - \sigma(t_{k+l+1}, \tilde{X}^{n,t_k,y}_{t_{k+l}})|^2 \mathbf{1}_{(\tilde{\tau}_l,\tilde{\tau}_{l+1}]}(r)dr\Big)^{\frac{p}{2}} \\
&\quad\leq\ C(\sigma,p)\tilde{\mathbb{E}}\Big(\sum_{l=0}^{m-1} \int_{\tilde{\tau}_l\wedge t_m}^{\tilde{\tau}_{l+1}\wedge t_m} |\tilde{\tau}_{l+1} - t_{l+1}|^{\frac{p}{2}} + |\tilde{\tau}_l - t_{l+1}|^{\frac{p}{2}} + |\tilde{X}^{t_k,x}_{t_k+r} - \tilde{X}^{t_k,x}_{t_k+\tilde{\tau}_l\wedge t_m}|^p \\
&\qquad\qquad +|\tilde{X}^{t_k,x}_{t_k+\tilde{\tau}_l\wedge t_m} - \tilde{X}^{n,t_k,y}_{t_{k+l}}|^p dr\Big) \\
&\quad\leq\ C(\sigma,T,p)\Big(h^{\frac{p}{2}} + \max_{1\leq l<m}(\tilde{\mathbb{E}}|\tilde{\tau}_l - t_l|^p)^{\frac{1}{2}} + \max_{0\leq l<m}(\tilde{\mathbb{E}} \sup_{\tilde{\tau}_l\wedge t_m\leq r\leq\tilde{\tau}_{l+1}\wedge t_m} |\tilde{X}^{t_k,x}_{t_k+r} - \tilde{X}^{t_k,x}_{t_k+\tilde{\tau}_l\wedge t_m}|^{2p})^{\frac{1}{2}} \\
&\qquad +\tilde{\mathbb{E}}\sum_{l=0}^{m-1} |\tilde{X}^{t_k,x}_{t_k+\tilde{\tau}_l\wedge t_m} - \tilde{X}^{n,t_k,y}_{t_{k+l}}|^p(\tilde{\tau}_{l+1} - \tilde{\tau}_l)\Big).
\end{aligned}$$

Moreover, since $\tilde{\tau}_{l+1}-\tilde{\tau}_l$ is independent from $|\tilde{X}^{t_k,x}_{t_k+\tilde{\tau}_l\wedge t_m}-\tilde{X}^{n,t_k,y}_{t_k+t_l}|^p$ we get by Lemma 5.1-(i)

$$\begin{aligned}
&\tilde{\mathbb{E}}\sum_{l=0}^{m-1}|\tilde{X}^{t_k,x}_{t_k+\tilde{\tau}_l\wedge t_m}-\tilde{X}^{n,t_k,y}_{t_k+l}|^p(\tilde{\tau}_{l+1}-\tilde{\tau}_l)\\
&\quad=\quad\tilde{\mathbb{E}}\sum_{l=0}^{m-1}|\tilde{X}^{t_k,x}_{t_k+\tilde{\tau}_l\wedge t_m}-\tilde{X}^{n,t_k,y}_{t_k+l}|^p(t_{l+1}-t_l)\\
&\quad\le\quad C(T,p)\Big(\tilde{\mathbb{E}}\int_0^{t_m}|\tilde{X}^{t_k,x}_{t_k+d(r)}-\tilde{X}^{n,t_k,y}_{t_k+d(r)}|^p dr+\max_{0\le l<m}\tilde{\mathbb{E}}|\tilde{X}^{t_k,x}_{t_k+\tilde{\tau}_l\wedge t_m}-\tilde{X}^{t_k,x}_{t_k+t_l}|^p\Big).
\end{aligned}$$

Using Lemma 5.1-(iii) one concludes similarly as in the proof of (ii) that $\tilde{\mathbb{E}}|\tilde{X}^{t_k,x}_{t_k+\tilde{\tau}_l\wedge t_m}-\tilde{X}^{t_k,x}_{t_k+t_l}|^p\le C(b,\sigma,T,p)h^{\frac{p}{4}}$. Then (63) combined with the above estimates implies that

$$\tilde{\mathbb{E}}\big|\tilde{X}^{t_k,x}_{t_k+t_m}-\tilde{X}^{n,t_k,y}_{t_k+t_m}\big|^p\le C(b,\sigma,T,p)\Big(|x-y|^p+h^{\frac{p}{4}}+\tilde{\mathbb{E}}\int_0^{t_m}|\tilde{X}^{t_k,x}_{t_k+d(r)}-\tilde{X}^{n,t_k,y}_{t_k+d(r)}|^p dr\Big).$$

Then Gronwall's lemma yields

$$\tilde{\mathbb{E}}\big|\tilde{X}^{t_k,x}_{t_k+t_m}-\tilde{X}^{n,t_k,y}_{t_k+t_m}\big|^p\le C(b,\sigma,T,p)(|x-y|^p+h^{\frac{p}{4}}).$$

(vi): We have

$$\begin{aligned}
\nabla\tilde{X}^{n,t_k,y}_{t_k+t_m}\quad=\quad&1+\int_{(0,t_m]}b_x(t_k+r,X^{n,t_k,y}_{t_k+r-})\nabla\tilde{X}^{n,t_k,y}_{t_k+r-}d[\tilde{B}^n,\tilde{B}^n]_r\\
&+\int_{(0,t_m]}\sigma_x(t_k+r,\tilde{X}^{n,t_k,y}_{t_k+r-})\nabla\tilde{X}^{n,t_k,y}_{t_k+r-}d\tilde{B}^n_r
\end{aligned}\tag{64}$$

and

$$\nabla\tilde{X}^{t_k,x}_{t_k+t_m}=1+\int_0^{t_m}b_x(t_k+r,\tilde{X}^{t_k,x}_{t_k+r})\nabla\tilde{X}^{t_k,x}_{t_k+r}dr+\int_0^{t_m}\sigma_x(t_k+r,\tilde{X}^{t_k,x}_{t_k+r})\nabla\tilde{X}^{t_k,x}_{t_k+r}d\tilde{B}_r.\tag{65}$$

We may proceed similarly as in (v) but this time the coefficients are not bounded but have linear growth. Here one uses that the integrands are bounded in any $L_p(\mathbb{P})$. □

Finally, we estimate the difference between the continuous-time Malliavin weight and its discrete-time counterpart.

Lemma 5.3. *Let B and $\tilde{B}$ be connected via (26). Under Assumption 2.1 it holds that*

$$\tilde{\mathbb{E}}|\tilde{N}^{t_k}_{t_m}\sigma(t_k,X_{t_k})-\tilde{N}^{n,\tau_k}_{\tilde{\tau}_m}\sigma(t_{k+1},\mathcal{X}_{\tau_k})|^2\le C(b,\sigma,T,\delta)\frac{|X_{t_k}-\mathcal{X}_{\tau_k}|^2+h^{\frac{1}{2}}}{(t_m-t_k)^{\frac{3}{2}}},\quad m=k+1,...,n.$$

Proof. For $N^{n,\tau_k}_{\tilde{\tau}_m}$ and $N^{t_k}_{t_m}$ given by (11) and (17), respectively, we introduce the notation

$$\tilde{N}^{t_k}_{t_m}\sigma(t_k,X_{t_k})=:\frac{1}{t_{m-k}}\int_0^{t_{m-k}}a_{t_k+s}d\tilde{B}_s\quad\text{and}\quad\tilde{N}^{n,\tau_k}_{\tilde{\tau}_m}\sigma(t_{k+1},\mathcal{X}_{\tau_k})=:\frac{1}{t_{m-k}}\int_0^{\tilde{\tau}_{m-k}}a^n_{\tau_k+s}d\tilde{B}_s$$

with

$$a_{t_k+s} := \nabla \tilde{X}^{t_k,X_{t_k}}_{t_k+s} \frac{\sigma(t_k, X_{t_k})}{\sigma(t_k+s, \tilde{X}^{t_k,X_{t_k}}_{t_k+s})} \quad \text{and} \quad a^n_{\tau_k+s} := \sum_{\ell=1}^{m-k} \nabla \tilde{\mathcal{X}}^{\tau_k,\mathcal{X}_{\tau_k}}_{\tau_k+\tilde{\tau}_{\ell-1}} \frac{\sigma(t_{k+1}, \mathcal{X}_{\tau_k})}{\sigma(t_{k+\ell}, \tilde{\mathcal{X}}^{\tau_k,\mathcal{X}_{\tau_k}}_{\tau_k+\tilde{\tau}_{\ell-1}})} \mathbf{1}_{s\in(\tilde{\tau}_{\ell-1},\tilde{\tau}_\ell]}.$$

By the inequality of BDG,

$$\begin{aligned}
&(t_m - t_k)^2 \tilde{\mathbb{E}} |\tilde{N}^{t_k}_{t_m} \sigma(t_k, X_{t_k}) - \tilde{N}^{n,\tau_k}_{\tilde{\tau}_m} \sigma(t_{k+1}, \mathcal{X}_{\tau_k})|^2 \\
&\quad = \tilde{\mathbb{E}} \Big| \int_0^{t_{m-k}} a_{t_k+s} d\tilde{B}_s - \int_0^{\tilde{\tau}_{m-k}} a^n_{\tau_k+s} d\tilde{B}_s \Big|^2 \\
&\quad = \tilde{\mathbb{E}} \int_0^{t_{m-k}\wedge\tilde{\tau}_{m-k}} (a_{t_k+s} - a^n_{\tau_k+s})^2 ds + \tilde{\mathbb{E}} \int_0^\infty a^2_{t_k+s} \mathbf{1}_{(\tilde{\tau}_{m-k}, t_{m-k}]}(s) ds \\
&\qquad + \tilde{\mathbb{E}} \int_0^\infty (a^n_{\tau_k+s})^2 \mathbf{1}_{(t_{m-k},\tilde{\tau}_{m-k}]}(s) ds \\
&\quad \le \sum_{\ell=1}^{m-k} \left(\tilde{\mathbb{E}} \sup_{s\in[0,t_{m-k}]\cap(\tilde{\tau}_{\ell-1},\tilde{\tau}_\ell]} \left| a_{t_k+s} - a^n_{\tau_k+\tilde{\tau}_\ell} \right|^4 \right)^{\frac{1}{2}} (\tilde{\mathbb{E}} |\tilde{\tau}_\ell - \tilde{\tau}_{\ell-1}|^2)^{\frac{1}{2}} \\
&\qquad + \left(\tilde{\mathbb{E}} \sup_{s\in[0,t_{m-k}]} |a_{t_k+s}|^4 + \tilde{\mathbb{E}} \max_{1\le\ell\le m-k} |a^n_{\tau_k+\tilde{\tau}_\ell}|^4 \right)^{\frac{1}{2}} (\tilde{\mathbb{E}} |t_{m-k} - \tilde{\tau}_{m-k}|^2)^{\frac{1}{2}}.
\end{aligned}$$

The assertion follows then from Lemma 5.1 and from the estimates

$$\tilde{\mathbb{E}} \sup_{s\in[0,t_{m-k}]\cap[\tilde{\tau}_{\ell-1},\tilde{\tau}_\ell]} |a_{t_k+s} - a^n_{\tau_k+\tilde{\tau}_\ell}|^4 \le C(b,\sigma,T,\delta)(|X_{t_k} - X^n_{t_k}|^4 + h) \tag{66}$$

$$\tilde{\mathbb{E}} \sup_{s\in[0,t_{m-k}]} |a_{t_k+s}|^4 + \tilde{\mathbb{E}} \max_{1\le\ell\le m-k} |a^n_{\tau_k+\tilde{\tau}_\ell}|^4 \le 2\|\sigma\|^4_\infty \delta^{-4}. \tag{67}$$

So it remains to show these inequalities. We put

$$\tilde{K}^{t_k}_{t_k+s} := \frac{\sigma(t_k, X_{t_k})}{\sigma(t_k+s, \tilde{X}^{t_k,X_{t_k}}_{t_k+s})} \quad \text{and} \quad \tilde{K}^{n,\tau_k}_{\tau_k+\tilde{\tau}_{\ell-1}} := \frac{\sigma(t_{k+1}, \mathcal{X}_{\tau_k})}{\sigma(t_{k+\ell}, \tilde{\mathcal{X}}^{\tau_k,\mathcal{X}_{\tau_k}}_{\tau_k+\tilde{\tau}_{\ell-1}})}$$

and notice that by Assumption 2.1 both expressions are bounded by $\|\sigma\|_\infty \delta^{-1}$. To show (66) let us split $a_{t_k+s} - a^n_{\tau_k+\tilde{\tau}_\ell}$ in the following way:

$$\begin{aligned}
a_{t_k+s} - a^n_{\tau_k+\tilde{\tau}_\ell} =& \tilde{K}^{t_k}_{t_k+s} (\nabla \tilde{X}^{t_k,X_{t_k}}_{t_k+s} - \nabla \tilde{X}^{t_k,X_{t_k}}_{t_k+t_{\ell-1}}) + \nabla \tilde{X}^{t_k,X_{t_k}}_{t_k+t_{\ell-1}} (\tilde{K}^{t_k}_{t_k+s} - \tilde{K}^{t_k}_{t_k+t_{\ell-1}}) \\
&+ \tilde{K}^{t_k}_{t_k+t_{\ell-1}} (\nabla \tilde{X}^{t_k,X_{t_k}}_{t_k+t_{\ell-1}} - \nabla \tilde{\mathcal{X}}^{\tau_k,\mathcal{X}_{\tau_k}}_{\tau_k+\tilde{\tau}_{\ell-1}}) + \nabla \tilde{\mathcal{X}}^{\tau_k,\mathcal{X}_{\tau_k}}_{\tau_k+\tilde{\tau}_{\ell-1}} (\tilde{K}^{t_k}_{t_k+t_{\ell-1}} - \tilde{K}^{n,\tau_k}_{\tau_k+\tilde{\tau}_{\ell-1}}).
\end{aligned}$$

Then

$$\begin{aligned}
&\tilde{\mathbb{E}} \sup_{s\in[\tilde{\tau}_{\ell-1}\wedge t_{m-k}, \tilde{\tau}_\ell \wedge t_{m-k}]} |\tilde{K}^{t_k}_{t_k+s} (\nabla \tilde{X}^{t_k,X_{t_k}}_{t_k+s} - \nabla \tilde{X}^{t_k,X_{t_k}}_{t_k+t_{\ell-1}})|^4 \\
&\quad \le \|\sigma\|^4_\infty \delta^{-4} \tilde{\mathbb{E}} \sup_{s\in[\tilde{\tau}_{\ell-1}\wedge t_{m-k}, \tilde{\tau}_\ell \wedge t_{m-k}]} |\nabla \tilde{X}^{t_k,X_{t_k}}_{t_k+s} - \nabla \tilde{X}^{t_k,X_{t_k}}_{t_k+t_{\ell-1}}|^4 \le C(b,\sigma,T,\delta) h
\end{aligned}$$

since one can show similarly to Lemma 5.2-(ii) that

$$\tilde{\mathbb{E}} \sup_{s\in[\tilde\tau_{\ell-1}\wedge t_{m-k},\tilde\tau_\ell\wedge t_{m-k}]} |\nabla\tilde X^{t_k,X_{t_k}}_{t_k+s} - \nabla\tilde X^{t_k,X_{t_k}}_{t_k+t_{\ell-1}}|^4 \le C(b,\sigma,T,\delta)h.$$

Notice that $\nabla\tilde X^{t_k,X_{t_k}}_t$ and $\nabla\tilde{\mathcal{X}}^{\tau_k,\mathcal{X}_{\tau_k}}_{\tau_m}$ solve the linear SDEs (65) and (64), respectively. Therefore,

$$\tilde{\mathbb{E}} \sup_{s\in[0,t_{m-k}]} |\nabla\tilde X^{t_k,X_{t_k}}_{t_k+s}|^p \le C(b,\sigma,T,p) \quad \text{and} \quad \tilde{\mathbb{E}} \max_{0\le\ell\le m-k} |\nabla\tilde{\mathcal{X}}^{\tau_k,\mathcal{X}_{\tau_k}}_{\tilde\tau_\ell+\tau_k}|^p \le C(b,\sigma,T,p). \tag{68}$$

For the second term we get

$$\begin{aligned}
&\tilde{\mathbb{E}} \sup_{s\in[\tilde\tau_{\ell-1}\wedge t_{m-k},\tilde\tau_\ell\wedge t_{m-k}]} |\nabla\tilde X^{t_k,X_{t_k}}_{t_k+t_{\ell-1}}(\tilde K^{t_k}_{t_k+s} - \tilde K^{t_k}_{t_k+t_{\ell-1}})|^4 \\
&\quad\le C(\sigma,\delta)(\tilde{\mathbb{E}}|\nabla\tilde X^{t_k,X_{t_k}}_{t_k+t_{\ell-1}}|^8)^{\frac12}(\tilde{\mathbb{E}} \sup_{s\in[\tilde\tau_{\ell-1}\wedge t_{m-k},\tilde\tau_\ell\wedge t_{m-k}]} (|t_\ell - s|^4 + |\tilde X^{t_k,X_{t_k}}_{t_k+s} - \tilde X^{t_k,X_{t_k}}_{t_k+t_\ell}|^8)^{\frac12} \\
&\quad\le C(b,\sigma,T,\delta)h.
\end{aligned}$$

For the third term Lemma 5.2-(vi) implies that

$$\tilde{\mathbb{E}}|\tilde K^{t_k}_{t_k+t_{\ell-1}}(\nabla\tilde X^{t_k,X_{t_k}}_{t_k+t_{\ell-1}} - \nabla\tilde{\mathcal{X}}^{\tau_k,\mathcal{X}_{\tau_k}}_{\tau_k+\tilde\tau_{\ell-1}})|^4 \le C(b,\sigma,T)\|\sigma\|^4_\infty\delta^{-4}(|X_{t_k} - \mathcal{X}_{\tau_k}|^4 + h).$$

The last term we estimate similarly to the second one,

$$\begin{aligned}
&\tilde{\mathbb{E}}|\nabla\tilde{\mathcal{X}}^{\tau_k,\mathcal{X}_{\tau_k}}_{\tau_k+\tilde\tau_{\ell-1}}(\tilde K^{t_k}_{t_k+t_{\ell-1}} - \tilde K^{n,\tau_k}_{\tau_k+\tilde\tau_{\ell-1}})|^4 \\
&\quad\le C(\sigma,\delta)(\tilde{\mathbb{E}}|\nabla\tilde{\mathcal{X}}^{\tau_k,\mathcal{X}_{\tau_k}}_{\tau_k+\tilde\tau_{\ell-1}}|^8)^{\frac12}(|X_{t_k} - \mathcal{X}_{\tau_k}|^8 + \tilde{\mathbb{E}}|\mathcal{X}^{\tau_k,\mathcal{X}_{\tau_k}}_{\tau_k+\tilde\tau_{\ell-1}} - \tilde X^{t_k,X_{t_k}}_{t_k+t_{\ell-1}}|^8)^{\frac12} \\
&\quad\le C(b,\sigma,T,\delta)(|X_{t_k} - \mathcal{X}_{\tau_k}|^4 + h).
\end{aligned}$$

To see (67) use the estimates (68). □

We close this section with estimates concerning the effect of $T_{m,\pm}$ and the discretized Malliavin derivative $\mathcal{D}^n_k$ (see Definition 2.1) on X^n.

Lemma 5.4. *Under Assumption 2.1, and for $p \ge 2$, we have*

(i) $\mathbb{E}|X^n_{t_l} - T_{m,\pm}X^n_{t_l}|^p \le C(b,\sigma,T,p)h^{\frac p2}, \quad 1 \le l,m \le n,$

(ii) $\mathbb{E}\left|\nabla X^{n,t_k,X^n_{t_k}}_{t_m} - \dfrac{\mathcal{D}^n_{k+1}X^n_{t_m}}{\sigma(t_{k+1},X^n_{t_k})}\right|^p \le C(b,\sigma,T,p)h^{\frac p2}, \quad 0 \le k < m \le n.$

(iii) $\mathbb{E}|\mathcal{D}^n_k X^n_{t_m}|^p \le C(b,\sigma,T,p), \quad 0 \le k \le m \le n.$

Proof. (i) By definition, $T_{m,\pm}X^n_{t_l} = X^n_{t_l}$ for $l \le m-1$, and for $l \ge m$ we have

$$\begin{aligned}
T_{m,\pm}X^n_{t_l} &= X^n_{t_{m-1}} + b(t_m,X^n_{t_{m-1}})h \pm \sigma(t_m,X^n_{t_{m-1}})\sqrt h \\
&\quad + h\sum_{j=m+1}^{l} b(t_j,T_{m,\pm}X^n_{t_{j-1}}) + \sqrt h\sum_{j=m+1}^{l}\sigma(t_j,T_{m,\pm}X^n_{t_{j-1}})\varepsilon_j.
\end{aligned}$$

By the properties of b and σ and thanks to the inequality of Burkholder-Davis-Gundy and Hölder's inequality we see that

$$\begin{aligned}
&\mathbb{E}|X^n_{t_l} - T_{m,\pm}X^n_{t_l}|^p \\
&\leq C(p)\Big(\mathbb{E}\big|\sigma(t_m, X^n_{t_{m-1}})\sqrt{h}(1\pm\varepsilon_m)\big|^p + h^p\mathbb{E}\Big|\sum_{j=m+1}^{l}\big(b(t_j, X^n_{t_{j-1}}) - b(t_j, T_{m,\pm}X^n_{t_{j-1}})\big)\Big|^p \\
&\quad + h^{\frac{p}{2}}\mathbb{E}\Big|\sum_{j=m+1}^{l}\big(\sigma(t_j, X^n_{t_{j-1}}) - \sigma(t_j, T_{m,\pm}X^n_{t_{j-1}})\big)^2\Big|^{\frac{p}{2}}\Big) \\
&\leq C(p)\Big(\|\sigma\|^p_\infty h^{\frac{p}{2}} + h(\|b_x\|^p_\infty t^{p-1}_{l-m} + \|\sigma_x\|^p_\infty t^{\frac{p}{2}-1}_{l-m})\sum_{j=m+1}^{l}\mathbb{E}|X^n_{t_{j-1}} - T_{m,\pm}X^n_{t_{j-1}}|^p\Big).
\end{aligned}$$

It remains to apply Gronwall's lemma.

(ii) By the inequality of Burkholder-Davis-Gundy (BDG) and Hölder's inequality,

$$\begin{aligned}
\mathbb{E}\Big|\nabla X^{n,t_k,X^n_{t_k}}_{t_m} - \frac{\mathcal{D}^n_{k+1}X^n_{t_m}}{\sigma(t_{k+1}, X^n_{t_k})}\Big|^p \leq C(p,T)\Big(&|b_x(t_{k+1}, X^n_{t_k})h + \sigma_x(t_{k+1}, X^n_{t_k})\sqrt{h}\varepsilon_{k+1}|^p \\
&+ h^p\sum_{l=k+2}^{m}\mathbb{E}\Big|b_x(t_l, X^n_{t_{l-1}})\nabla X^{n,t_k,X^n_{t_k}}_{t_{l-1}} - b^{(k+1,l)}_x\frac{\mathcal{D}^n_{k+1}X^n_{t_{l-1}}}{\sigma(t_{k+1}, X^n_{t_k})}\Big|^p \\
&+ h^{\frac{p}{2}}\sum_{l=k+2}^{m}\mathbb{E}\Big|\sigma_x(t_l, X^n_{t_{l-1}})\nabla X^{n,t_k,X^n_{t_k}}_{t_{l-1}} - \sigma^{(k+1,l)}_x\frac{\mathcal{D}^n_{k+1}X^n_{t_{l-1}}}{\sigma(t_{k+1}, X^n_{t_k})}\Big|^p\Big).
\end{aligned}$$

Since by Lemma 5.4 (i) we conclude that

$$\mathbb{E}|b^{(k+1,l)}_x - b_x(t_l, X^n_{t_{l-1}})|^{2p} + \mathbb{E}|\sigma^{(k+1,l)}_x - \sigma_x(t_l, X^n_{t_{l-1}})|^{2p} \leq C(b,\sigma,T,p)h^p,$$

and Lemma 5.2 implies that

$$\mathbb{E}\sup_{k+1\leq l\leq m}\Big|\nabla X^{n,t_k,X^n_{t_k}}_{t_{l-1}}\Big|^{2p} \leq C(b,\sigma,T,p),$$

the assertion follows by Gronwall's lemma.

(iii) This is an immediate consequence of (i).

□

Acknowledgements

Christel Geiss would like to thank the Erwin Schrödinger Institute, Vienna, for hospitality and support, where a part of this work was written.

References

[1] S. Alanko. *Regression-based Monte Carlo methods for solving nonlinear PDEs.* PhD thesis, New York University, 2015.

[2] S. Ankirchner, T. Kruse, and M. Urusov. *Wasserstein convergence rates for coin tossing approximations of continuous Markov processes.* https://arxiv.org/abs/1903.07880, 2019.

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

[4] C. Bender and P. Parczewski. *Discretizing Malliavin calculus.* Stochastic Process. Appl., 128(8):2489–2537, 2018.

[5] P. Briand, B. Delyon, and J. Mémin. *Donsker-Type theorem for BSDEs.* Electron. Comm. Probab., 6(1):1–14, 2001.

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

[7] 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.

[8] C. Bender and J. Zhang. *Time discretization and Markovian iteration for coupled FBSDEs.* Ann. Appl. Probab., 18(1):143–177, 2008.

[9] J.-F. Chassagneux and D. Crisan. *Runge-Kutta schemes for backward stochastic differential equations.* Ann. Appl. Probab., 24(2):679–720, 2014.

[10] J.-F. Chassagneux. *Linear multistep schemes for BSDEs.* SIAM J. Num. Anal., 52(6):2815–2836, 2014.

[11] J.-F. Chassagneux, D. Crisan, and F. Delarue. *Numerical Method for FBSDEs of McKean-Vlasov type.* Ann. Appl. Probab., 29(3):1640–1684, 2019.

[12] J.-F. Chassagneux and A. Richou. *Numerical Stability Analysis of the Euler Scheme for BSDEs.* SIAM J. Num. Anal., 53(2):1172–1193, 2015.

[13] J.-F. Chassagneux and A. Richou. *Rate of convergence for discrete-time approximation of reflected BSDEs arising in switching problems.* Stochastic Process. Appl., 129(11): 4597–4637, 2019.

[14] J.-F. Chassagneux and C. A. Garcia Trillos. *Cubature methods to solve BSDEs: Error expansion and complexity control.* https://arxiv.org/abs/1702.00999, 2017.

[15] P.-E. Chaudru de Raynal and C. A. Garcia Trillos. *A cubature based algorithm to solve decoupled McKean-Vlasov forward-backward stochastic differential equations.* Stochastic Process. Appl., 125(6):2206–2255, 2015.

[16] P. Cheridito and M. Stadje. *BSΔEs and BSDEs with non-Lipschitz drivers: Comparison, convergence and robustness.* Bernoulli 19(3): 1047–1085, 2013.

[17] D. Crisan, K. Manolarakis, and N. Touzi. *On the Monte-Carlo simulation of BSDEs: An improvement on the Malliavin weights.* Stochastic Process. Appl., 120(7):1133–1158, 2010.

[18] F. Delarue and S. Menozzi. *A forward-backward stochastic algorithm for quasi-linear PDEs.* Ann. Appl. Probab., 16(1):140–184, 2006.

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

[20] C. Geiss, S. Geiss, and E. Gobet. *Generalized fractional smoothness and Lp-variation of BSDEs with non-Lipschitz terminal conditions.* Stochastic Process. Appl., 122(5):2078–2116, 2012.

[21] C. Geiss, C. Labart, and A. Luoto. *Random walk approximation of BSDEs with Hölder continuous terminal condition.* Bernoulli, 26(1):159–190, 2020.

[22] C. Geiss and C. Labart. *Simulation of BSDEs with jumps by Wiener chaos expansion.* Stochastic Process. Appl., 126(7):2123–2162, 2016.

[23] 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.

[24] J. Jacod and A. N. Shiryaev. *Limit theorems for stochastic processes.* Springer, 2003.

[25] K. Jańczak-Borkowska. *Discrete approximations of generalized RBSDE with random terminal time.* Discuss. Math. Probab. Stat., 32(1–2):69–85, 2012.

[26] P. Henry-Labordere, 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.

[27] T. Kruse and A. Popier. *BSDEs with monotone generator driven by Brownian and Poisson noises in a general filtration*, Stochastics 88(4):491–539, 2016.

[28] T. Kruse, A. Popier, *L_p-solution for BSDEs with jumps in the case $p < 2$*, Stochastics 89(8):1201–1227,2017.

[29] J. Ma, P. Protter, J. San Martín, and S. Torres. *Numerical method for backward stochastic differential equations.* Ann. Appl. Probab., 12(4):302–316, 2007.

[30] J. Ma and J. Zhang. *Representation theorems for backward stochastic differential equations.* Ann. Appl. Probab., 12(4):1390–1418, 2002.

[31] M. Martínez, J. San Martín, and S. Torres. *Numerical Method for reflected backward stochastic differential equations.* Stoch. Anal. Appl., 29(6):1008–1032, 2011.

[32] J. Mémin, S. Peng and M. Xu. *Convergence of solutions of discrete reflected backward SDE's and simulations.* Acta Math. Appl. Sin. Engl. Ser., 24(1):1–18, 2008.

[33] S. Peng and M. Xu. *Numerical algorithms for backward stochastic differential equations with 1-d Brownian motion: convergence and simulations.* Math. Model. Numer. Anal., 45(2):335–360, 2011.

[34] N. Privault. *Stochastic Analysis in Discrete and Continuous Settings–With Normal Martingales*. LNM 192, Springer Berlin, 2009.

[35] T. Tao. *An Epsilon of Room I: Real Analysis*. AMS, 2010.

[36] S. Toldo. *Stability of solutions of BSDEs with random terminal time*. ESAIM: Probability and Statistics, 10:141–163, 2006.

[37] S. Toldo. *Corrigendum to 'Stability of solutions of BSDEs with random terminal time'*. ESAIM: Probability and Statistics, 11:381–384, 2007.

[38] S. Yao. *Lp Solutions of Backward Stochastic Differential Equations with Jumps*. Stochastic Process. Appl., 127(11): 3465–3511, 2017.

[39] J. B. Walsh. *The rate of convergence of the binomial tree scheme*. Finance Stochast., 7(3):337–36, 2003.

[40] E. Weinan, M. Hutzenthaler, A. Jentzen, and T. Kruse. *On multilevel Picard numerical approximations for high-dimensional nonlinear parabolic partial differential equations and high-dimensional nonlinear backward stochastic differential equations*. J. Sci. Comput., 79(3): 1534–1571, 2019.

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

[42] J. Zhang. *Representation of solutions to BSDEs associated with a degenerate FSDE*. Ann. Appl. Probab., 15(3):1798–1831, 2005.

[43] J. Zhang. *Some fine properties of backward stochastic differential equations, with applications*. PhD thesis, Purdue University, 2001.