Donsker-Type Theorem for BSDEs: Rate of Convergence

Philippe Briand, Christel Geiss, Stefan Geiss, Celine Labart

Donsker-Type Theorem for BSDEs: Rate of Convergence

Philippe Briand [*†] Christel Geiss Stefan Geiss Céline Labart [†]

This version: 2019-06-24

Abstract

In this paper, we study in the Markovian case the rate of convergence in the Wasserstein distance of an approximation of the solution to a BSDE given by a BSDE which is driven by a scaled random walk as introduced in Briand, Delyon and Mémin (Electron. Comm. Probab. **6** (2001), 1–14).

1. Introduction.

In this paper, we are concerned with the discretization of solutions to BSDEs of the form

$$Y_t = G(B) + \int_t^T f(B_s, Y_s, Z_s)\, ds - \int_t^T Z_s\, dB_s, \quad 0 \le t \le T,$$

where B is a standard Brownian motion. These equations have been introduced by Jean-Michel Bismut for linear generators in [2] and by Étienne Pardoux and Shige Peng for Lipschitz generators in [14].

In one of the first studies on this topic, in the case where the generator f may depend on z as well, Philippe Briand, Bernard Delyon and Jean Mémin [5] proposed an approximation based on Donsker's theorem. They showed that the solution (Y, Z) to the previous BSDE can be approximated by the solution (Y^n, Z^n) to the BSDE

$$Y_t^n = G(B^n) + \int_{]t,T]} f(B_{s-}^n, Y_{s-}^n, Z_s^n)\, d\langle B^n\rangle_s + \int_{]t,T]} Z_s^n\, dB_s^n, \quad 0 \le t \le T,$$

where B^n is the scaled random walk

$$B_t^n = \sqrt{T/n} \sum_{k=1}^{[nt/T]} \xi_k, \quad 0 \le t \le T,$$

and $(\xi_k)_{k\ge1}$ is an i.i.d. sequence of symmetric Bernoulli random variables. They proved, in full generality, meaning that $G(B)$ is only required to be a square integrable random variable, that (Y^n, Z^n) converges to (Y, Z). However, the question of the rate of convergence was left open. Right now it seems to be hopeless to get a result in this direction for such a general path-dependent terminal condition $G(B)$. But in the Markovian case, meaning that $G(B) = g(B_T)$,

[*] In memory of Jean Mémin from whom I learned lots of mathematics.

[†] Many thanks to Pierre Baras for very fruitful discussions about the heat equation.

this problem seems to be tractable, in particular due to the PDE structure behind. Indeed, (Y, Z) is related to the semilinear heat equation

$$\partial_t u(t,x) + \frac{1}{2}\Delta u(t,x) + f(t,x,u(t,x),\nabla u(t,x)) = 0, \quad (t,x) \in [0,T[\times\mathbb{R}, \qquad u(T,\cdot) = g, \quad (1)$$

where, under certain regularity conditions, we can choose

$$Y_s = u(s,B_s) \quad \text{and} \quad Z_s = \nabla u(s,B_s).$$

In the case where B^n is the discretized Brownian motion, the link to PDEs was exploited in [18, 3] to get the rate of convergence, in the Markovian case, of the classical scheme for BSDEs. The convergence of this scheme was already proved in [6, Proposition 13] for a general terminal condition and a generator that is Lipschitz in its spatial coordinates but without any rate of convergence.

Even though the link with PDEs was pointed out in [5], the rate of convergence of the approximation of BSDEs given by scaled random walks was completely open. In two recent papers, Christel Geiss, Céline Labart and Antti Luoto [10, 9] give a first answer to this question. They showed that the error between (Y^n, Z^n) and (Y, Z) is of order $n^{-\varepsilon/4}$ when g is assumed to be ε-Hölder continuous and $f(t,\cdot)$ Lipschitz continuous. One of the main arguments in these papers consists in constructing the random walk from the Brownian motion B using the Skorohod embedding (see [17]) together with generalizations of the pioneering work of Jin Ma and Jianfeng Zhang [13] on representation theorems for BSDEs. This approach allows to work with convergence in the L^2-sense even if the problem naturally arises in the weak sense. The drawback is that the rate of convergence $n^{-\varepsilon/4}$ obtained in these papers is not optimal as one can expect $n^{-\varepsilon/2}$.

The objective of our study is to improve the previous rate of convergence by using a weak limit approach, the error being measured in the Wasserstein distance. Our starting point is a result of Emmanuel Rio [15] who proved that, when $T = 1$, for all $r \geq 1$, there exists a constant C_r such that, for all $n \geq 1$, $W_r(B_1^n, G) \leq C_r\, n^{-1/2}$, where W_r is the L^r-Wasserstein distance and G a standard normal random variable (see Section 3). Firstly, we generalize this result to cover the case where $f \equiv 0$ which corresponds to the heat equation. Then, using the associated PDE, in particular representation formulas in the spirit of [13], we are able to prove that

$$W_r(Y_t^n, Y_t) \leq C_r\, n^{-\varepsilon/2}, \quad \text{and} \quad W_r(Z_t^n, Z_t) \leq \frac{C_r}{\sqrt{T-t}}\, n^{-\varepsilon/2} \text{ for all } t \in [0,T[,$$

when g and $f(t,\cdot,y,z)$ are ε-Hölder continuous, which gives the correct rate of convergence when the error is measured in the Wasserstein-distance. We refer to Theorem 10 in Section 5 for the precise statement.

2. Notation.

In all the sequel, $T > 0$ is a fixed positive real number. We work on a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ carrying a standard real Brownian motion $\{B_t\}_{0\leq t\leq T}$, and $\{\mathcal{F}_t\}_{0\leq t\leq T}$ stands for the augmented filtration of B which is right continuous and complete.

We consider the following BSDE

$$Y_t = g(B_T) + \int_t^T f(s,B_s,Y_s,Z_s)\,ds - \int_t^T Z_s\,dB_s, \quad 0 \leq t \leq T. \quad (2)$$

Throughout this article, we will assume for the function g defining the terminal condition and the generator f the following:

Assumption (A1). There exist $0 < \varepsilon \leq 1$ and $0 < \alpha \leq 1$ such that it holds:

(i) The function $g : \mathbb{R} \longrightarrow \mathbb{R}$ is ε-Hölder continuous: for all $(x, x') \in \mathbb{R}^2$ one has

$$|g(x) - g(x')| \leq \|g\|_\varepsilon \, |x - x'|^\varepsilon .$$

(ii) The function $f : [0,T] \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \longrightarrow \mathbb{R}$ is α-Hölder continuous in time, ε-Hölder continuous in space and Lipschitz continuous with respect to (y, z): for all (t, x, y, z) and (t', x', y', z') in $[0,T] \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}$ one has

$$\begin{aligned} &|f(t,x,y,z) - f(t',x',y',z')| \\ &\qquad \leq \|f_t\|_\alpha \, |t - t'|^\alpha + \|f_x\|_\varepsilon \, |x - x'|^\varepsilon + \|f_y\|_{\mathrm{Lip}} \, |y - y'| + \|f_z\|_{\mathrm{Lip}} \, |z - z'| . \end{aligned} \tag{3}$$

Most of the time, we do not need to distinguish between $\|f_y\|_{\mathrm{Lip}}$ and $\|f_z\|_{\mathrm{Lip}}$ and we let $\|f\|_{\mathrm{Lip}} := \max(\|f_y\|_{\mathrm{Lip}}, \|f_z\|_{\mathrm{Lip}})$.

Convention: Later the phrase that *a constant $C > 0$ depends on (T, ε, f, g)* stands for the fact that C can be expressed in terms of $(T, \varepsilon, \|f_x\|_\varepsilon, \|f_y\|_{\mathrm{Lip}}, \|f_z\|_{\mathrm{Lip}}, K_f, \|g\|_\varepsilon, g(0))$ where

$$K_f := \sup_{t \in [0,T]} |f(t,0,0,0)|.$$

Similarly, a dependence on $(T, \alpha, \varepsilon, f, g)$ means an additional dependence on $(\alpha, \|f_t\|_\alpha)$.

From [4, Theorem 4.2] it is known that under (A1), the BSDE (2) has a unique L^p-solution (Y, Z) for any $p \in]1, \infty[$. So for $(t, x) \in [0, T[\times \mathbb{R}$ we let $\left(Y_s^{t,x}, Z_s^{t,x}\right)_{s \in [t,T]}$ be the square integrable solution to the BSDE

$$Y_s^{t,x} = g\left(B_T^{t,x}\right) + \int_s^T f\left(r, B_r^{t,x}, Y_r^{t,x}, Z_r^{t,x}\right) dr - \int_s^T Z_r^{t,x} \, dB_r, \quad t \leq s \leq T, \tag{4}$$

where $B_r^{t,x} := x + B_r - B_t$, and set, as usual, for $x \in \mathbb{R}$, $u(T, x) := g(x)$, and, for $(t, x) \in [0, T[\times \mathbb{R}$,

$$u(t,x) := Y_t^{t,x} = \mathbb{E}\left[g(B_T^{t,x}) + \int_t^T f\left(r, B_r^{t,x}, Y_r^{t,x}, Z_r^{t,x}\right) dr \right].$$

It is well known that the function u is continuous on $[0,T] \times \mathbb{R}$ (see also Lemma 6 below) and under Lipschitz assumptions in (x, y, z) and for $\alpha \geq \frac{1}{2}$ it is the viscosity solution to (1), see [20, Theorem 5.5.8]. Moreover, in this Markovian setting, for $(t, x) \in [0,T] \times \mathbb{R}$, we have $Y_s^{t,x} = u(s, B_s^{t,x})$ a.s. for all $s \in [t, T]$. In [19, Theorem 3.2], for a generator which is Lipschitz continuous in all space variables and a measurable g with polynomial growth, J. Zhang proved that u belongs to $\mathcal{C}^{0,1}([0, T[\times \mathbb{R})$ and that $Z_s^{t,x} = \nabla u(s, B_s^{t,x})$ a.e. on $[t, T[\times \Omega$. Moreover, the following representation holds

$$\nabla u(t,x) = \mathbb{E}\left[g(B_T^{t,x}) \frac{B_T - B_t}{T - t} + \int_t^T f\left(r, B_r^{t,x}, Y_r^{t,x}, Z_r^{t,x}\right) \frac{B_r - B_t}{r - t} \, dr \right], \quad (t,x) \in [0, T[\times \mathbb{R}.$$

If F is the function given by

$$F(s,x) := f(s, x, u(s,x), \nabla u(s,x)) \quad \text{for} \quad (s,x) \in [0, T[\times \mathbb{R}, \tag{5}$$

we thus have

$$u(t,x) = \mathbb{E}\left[g(B_T^{t,x}) + \int_t^T F(r, B_r^{t,x}) \, dr \right], \quad (t,x) \in [0, T[\times \mathbb{R}, \tag{6}$$

together with

$$\nabla u(t,x) = \mathbb{E}\left[g(B_T^{t,x}) \frac{B_T - B_t}{T-t} + \int_t^T F(r, B_r^{t,x}) \frac{B_r - B_t}{r-t}\, dr \right], \quad (t,x) \in [0,T[\times\mathbb{R}. \tag{7}$$

These formulas play an important role in the sequel.

In Section 4 and in the appendix, we extend these results to the case where $f(t,\cdot,y,z)$ is ε-Hölder continuous and make the regularity of u and ∇u precise.

As mentioned before, we are concerned with the approximation of the solution $(Y^{t,x}, Z^{t,x})$ to (4) by a solution to the BSDE driven by a scaled random walk. To do this, let us consider, on some probability space, not necessarily $(\Omega, \mathcal{F}, \mathbb{P})$, an i.i.d. sequence $(\xi_k)_{k\geq 1}$ of symmetric Bernoulli random variables. For $n \in \mathbb{N}^* := \{1,2,3,...\}$ we set $h := T/n$ and we consider the scaled random walk

$$B_t^n := \sqrt{h} \sum_{k=1}^{[t/h]} \xi_k, \quad 0 \leq t \leq T,$$

where $[x] := \max\{r \in \mathbb{Z} : r \leq x\}$ for any real number x. As we did for the Brownian motion, for $x \in \mathbb{R}$ and $0 \leq t \leq s \leq T$ we put

$$B_s^{n,t,x} := x + B_s^n - B_t^n.$$

Let us introduce some further notation. We denote the ceiling function by $\lceil x \rceil := \min\{r \in \mathbb{Z} : r \geq x\}$ for $x \in \mathbb{R}$. Moreover, we set

$$n_t := [t/h], \quad \underline{t} := h[t/h] = hn_t \quad \text{and} \quad \bar{t} := h\lceil t/h \rceil, \quad t \in [0,T].$$

For $n \in \mathbb{N}^*$ let us consider the following BSDE driven by B^n:

$$Y_t^n = g(B_T^n) + \int_{]t,T]} f(s, B_{s^-}^n, Y_{s^-}^n, Z_s^n)\, d\langle B^n \rangle_s - \int_{]t,T]} Z_s^n\, dB_s^n, \quad t \in [0,T].$$

It was shown in [5] that, as soon as $h\,\|f\|_{\mathrm{Lip}} < 1$, this BSDE has a unique square integrable solution (Y^n, Z^n), Y^n being adapted and Z^n being predictable with respect to the filtration generated by B^n. By construction, Y^n is a piecewise constant càdlàg process with $Y_t^n = Y_{\underline{t}}^n$. The process Z^n is defined as an element of $\mathrm{L}^2(\Omega \times [0,T], d\mathbb{P} \otimes d\langle B^n \rangle)$, where we start with a Z^n defined only on the points $\{kh : k = 1, \ldots, n\}$ and extend it to $]0,T]$ as a càglàd process $(Z_t^n)_{t\in]0,T]}$ by setting $Z_t^n = Z_{\bar{t}}^n$. The previous BSDE is actually a discrete BSDE that can be solved by hand since, for $k = 0, \cdots, n-1$, we have

$$Y_{kh}^n = Y_{(k+1)h}^n + h\, f\left((k+1)h, B_{kh}^n, Y_{kh}^n, Z_{(k+1)h}^n\right) - \sqrt{h}\, Z_{(k+1)h}^n \xi_{k+1}, \quad Y_{nh}^n = g(B_T^n).$$

Thus, if $Y_{(k+1)h}^n$ is given,

$$Z_{(k+1)h}^n = h^{-1/2}\, \mathbb{E}\left(Y_{(k+1)h}^n \xi_{k+1} \,|\, \mathcal{F}_{kh}^n\right), \tag{8}$$

$$\begin{aligned} Y_{kh}^n &= Y_{(k+1)h}^n + h\, f\left((k+1)h, B_{kh}^n, Y_{kh}^n, Z_{(k+1)h}^n\right) - \sqrt{h}\, Z_{(k+1)h}^n \xi_{k+1} \\ &= \mathbb{E}\left(Y_{(k+1)h}^n \,|\, \mathcal{F}_{kh}^n\right) + h\, f\left((k+1)h, B_{kh}^n, Y_{kh}^n, Z_{(k+1)h}^n\right), \end{aligned} \tag{9}$$

where the last equality follows by taking the conditional equation w.r.t. $\mathcal{F}_{kh}^n$ of the second line.

Since we are in a Markovian setting, there is also an analog of the Feynman-Kac formula. If u is a given function we set

$$D_+^n u(x) := \frac{1}{2}\left(u(x+\sqrt{h}) + u(x-\sqrt{h})\right), \qquad D_-^n u(x) := \frac{1}{2}\left(u(x+\sqrt{h}) - u(x-\sqrt{h})\right),$$

and

$$\nabla^n u(x) := h^{-1/2}\, D_-^n u(x). \tag{10}$$

Remark 1. From the definition of D_+^n and D_-^n, we get that if u is ε-Hölder, $D_+^n u$ and $D_-^n u$ are also ε-Hölder with constant $\|u\|_\varepsilon$.

Let U^n be the solution to the finite difference equation, where for $x \in \mathbb{R}$ and $k = 0, \dots, n-1$ we require

$$\begin{cases} U^n(kh,x) = D_+^n U^n((k+1)h,x) + hf((k+1)h, x, U^n(kh,x), h^{-1/2} D_-^n U^n((k+1)h,x)), \\ U^n(nh,x) = g(x). \end{cases} \tag{11}$$

Then, we obtain from (8) and (9) (cf [5, Proposition 5.1]) that

$$Y_{kh}^n = U^n\left(kh, \sqrt{h}\sum_{i=1}^{k}\xi_i\right), \quad k = 0, \dots, n,$$
$$Z_{kh}^n = \nabla^n U^n\left(kh, \sqrt{h}\sum_{i=1}^{k-1}\xi_i\right), \quad k = 1, \dots, n.$$

These formulas rewrite in continuous time to

$$Y_t^n = Y_{\underline{t}}^n = U^n(\underline{t}, B_t^n), \quad t \in [0,T], \quad \text{and} \quad Z_t^n = Z_{\overline{t}}^n = \nabla^n U^n\left(\overline{t}, B_{t-}^n\right), \quad t \in]0,T].$$

If we set, for $0 \le t \le T$ and $x \in \mathbb{R}$, $U^n(t,x) := U^n(\underline{t},x)$, we have $Y_t^n = U^n(t, B_t^n)$.

More generally, for $0 \le t < T$, we define $(Y^{n,t,x}, Z^{n,t,x})$ as the solution $Y^{n,\underline{t},x} = (Y_s^{n,\underline{t},x})_{s\in[\underline{t},T]}$ and $Z^{n,\underline{t},x} = (Z_s^{n,\underline{t},x})_{s\in]\underline{t},T]}$ to the BSDE

$$Y_s^{n,\underline{t},x} = g(B_T^{n,t,x}) + \int_{]s,T]} f(r, B_{r^-}^{n,t,x}, Y_{r^-}^{n,\underline{t},x}, Z_r^{n,\underline{t},x}) d\langle B^n\rangle_r - \int_{]s,T]} Z_r^{n,\underline{t},x} dB_r^n, \quad s \in [\underline{t},T]. \tag{12}$$

We set $Y_T^{n,T,x} = g(x)$. Then,

$$Y_s^{n,t,x} = Y_{\underline{s}}^{n,\underline{t},x} = U^n(s, B_s^{n,t,x}), \quad 0 \le t \le s \le T. \tag{13}$$

Let us observe that $Z^{n,\underline{t},x}$ is first defined at the points $t = kh$, $k = n_t + 1, \dots, n$. As before we let $Z_s^{n,\underline{t},x} := Z_{\overline{s}}^{n,\underline{t},x}$ for $s \in]\underline{t},T]$. We have

$$Z_s^{n,t,x} = Z_{\overline{s}}^{n,\underline{t},x} = \nabla^n U^n(\overline{s}, B_{s^-}^{n,t,x}) \quad \text{for } s \in]\underline{t},T].$$

In particular,

$$Z_s^{n,t,x} = \nabla^n U^n(\underline{s}+h, B_s^{n,t,x}) \quad \text{whenever } s \in]\underline{t},T]\backslash\{kh : k = n_t+1, \dots, n\}. \tag{14}$$

Of course, we have

$$U^n(t,x) = Y_t^{n,t,x} = Y_{\underline{t}}^{n,\underline{t},x} \quad \text{for } t \in [0,T].$$

Similarly, we define, for $(t,x) \in [0,T[\times\mathbb{R}$,

$$\Delta^n(t,x) := \nabla^n U^n(\underline{t}+h,x) = Z_{\underline{t}+h}^{n,\underline{t},x}. \tag{15}$$

With this notation, (14) rewrites as

$$Z_s^{n,t,x} = \Delta^n(s, B_s^{n,t,x}) \quad \text{whenever } s \in]\underline{t},T]\backslash\{kh : k = n_t+1,\dots,n\}. \tag{16}$$

It follows that

$$\begin{aligned} U^n(t,x) &= \mathbb{E}\left[g\left(B_T^{n,t,x}\right) + h\sum_{k=n_t+1}^{n} f\left(kh, B_{(k-1)h}^{n,t,x}, Y_{(k-1)h}^{n,t,x}, Z_{kh}^{n,t,x}\right)\right] \\ &= \mathbb{E}\left[g\left(B_T^{n,t,x}\right) + \int_{\underline{t}}^{T} f\left(\overline{s}, B_s^{n,t,x}, Y_s^{n,t,x}, Z_s^{n,t,x}\right) ds\right], \end{aligned}$$

which rewrites, taking into account (13) and (16), to

$$\begin{aligned} U^n(t,x) &= \mathbb{E}\left[g\left(B_T^{n,t,x}\right) + \int_{\underline{t}}^{T} f\left(\overline{s}, B_s^{n,t,x}, U^n(s, B_s^{n,t,x}), \Delta^n(s, B_s^{n,t,x})\right) ds\right] \\ &= \mathbb{E}\left[g\left(B_T^{n,t,x}\right) + \int_{\underline{t}}^{T} f\left(\overline{s}, \Theta_s^{n,t,x}\right) ds\right], \end{aligned} \tag{17}$$

where

$$\Theta_s^{n,t,x} := \left(B_s^{n,t,x}, U^n(s, B_s^{n,t,x}), \Delta^n(s, B_s^{n,t,x})\right). \tag{18}$$

We will prove in Section 5 that (U^n, Δ^n) converges to $(u, \nabla u)$.

From now on we assume that $n \geq n_0(T, \|f\|_{\mathrm{Lip}})$ where $n_0(T, \|f\|_{\mathrm{Lip}}) \in \mathbb{N}^*$ is the integer given in Lemma 12 in the appendix and which automatically implies also existence and uniqueness of solutions because $n_0 > T\|f\|_{\mathrm{Lip}}$.

3. Scaled random walk and Wasserstein distance.

One starting point of our paper is the following result of Emmanuel Rio [15] (Theorem 2.1); see also [16]. This result covers, up to a generalization, the case where the generator vanishes, i.e. $f \equiv 0$.

Let ψ be the convex function defined by $\psi(x) = e^{|x|} - 1$. The Orlicz norm associated to this function ψ of any real random variable X is given by

$$\|X\|_\psi := \inf\{a > 0 : \mathbb{E}\left[\psi(X/a)\right] \leq 1\}, \qquad \inf\emptyset := +\infty.$$

Let us recall that, for any $r \geq 1$,

$$\sup_{x>0}\left\{\tfrac{x^r}{\psi(x)}\right\} < +\infty, \qquad \|X\|_{\mathrm{L}^r} \leq \left(\sup_{x>0}\left\{\tfrac{x^r}{\psi(x)}\right\}\right)^{1/r} \|X\|_\psi. \tag{19}$$

Let X and Y be two random variables end let us denote by μ the law of X and by ν the law of Y. With the usual abuse of notation, the Wasserstein distance associated to ψ is defined by

$$W_\psi(\mu,\nu) = W_\psi(X,Y) := \inf\left\{\|X-Y\|_\psi : \mathrm{law}(X) = \mu,\ \mathrm{law}(Y) = \nu\right\}.$$

Let $(X_k)_{k\geq 1}$ be an i.i.d. sequence of random variables with $\mathbb{E}[X]=0$, $\mathbb{E}[X^2]=1$, and such that, for some $\sigma>0$, $\mathbb{E}\left[e^{\sigma|X|}\right]<+\infty$. Let G be a standard normal random variable. In [15, Theorem 2.1], Emmanuel Rio proved that there exists a constant $C>0$ such that, for $n\geq 1$,

$$W_\psi\left(n^{-1/2}S_n, G\right)\leq C\, n^{-1/2}, \quad \text{where} \quad S_n = X_1+\ldots+X_n.$$

As a byproduct, for any $r\geq 1$, there exists a constant $c_r>0$ such that

$$W_r\left(n^{-1/2}S_n, G\right)\leq c_r\, n^{-1/2},$$

where W_r stands for the L^r-Wasserstein distance

$$W_r(\mu,\nu)=W_r(X,Y):=\inf\left\{\mathbb{E}\left[|X-Y|^r\right]^{1/r} : \mathrm{law}(X)=\mu,\ \mathrm{law}(Y)=\nu\right\}. \tag{20}$$

We have also the result of Kantorovich-Rubinstein, i.e.

$$W_1(\mu,\nu)=W_1(X,Y)=\sup\{\mathbb{E}[f(X)]-\mathbb{E}[f(Y)] : \|f\|_{\mathrm{Lip}}\leq 1\}. \tag{21}$$

Remark 2. We could also consider the case where $0<r<1$ by using the fact that, in this case, $\mathbb{E}(|X-Y|^r)$ is a distance (see the arguments in [1, Section 7.1]). In general, we have $W_p(\mu,\nu)=W_q(\mu,\nu)$ for $0<p<q<\infty$.

Let us start with a straightforward generalization of Rio's result.

Proposition 3. *There exists a $C>0$ such that, for all $x\in\mathbb{R}$ and all $0\leq t\leq s\leq T$,*

$$W_\psi\left(B_s^{n,t,x}, B_s^{t,x}\right)\leq C\left(\frac{T}{n}\right)^{1/2}.$$

As a byproduct, taking into account (19), for any $r\geq 1$, there exists a $c_r>0$ such that, for all $x\in\mathbb{R}$ and all $0\leq t\leq s\leq T$,

$$W_r\left(B_s^{n,t,x}, B_s^{t,x}\right)\leq c_r\left(\frac{T}{n}\right)^{1/2}. \tag{22}$$

Proof of Proposition 3. We have, for any $x\in\mathbb{R}$ and all $0\leq t\leq s\leq T$,

$$W_\psi\left(x+B_s^n-B_t^n, x+B_s-B_t\right)=W_\psi\left(B_s^n-B_t^n, B_s-B_t\right).$$

If $\underline{s}=\underline{t}$, then $B_t^n-B_s^n=0$, and we have

$$W_\psi\left(B_s^n-B_t^n, B_s-B_t\right)=\|B_s-B_t\|_\psi=\sqrt{s-t}\,\|G\|_\psi\leq\sqrt{h}\,\|G\|_\psi.$$

Let us assume that $\underline{t}<\underline{s}$ and let us write

$$\begin{aligned}W_\psi\left(B_s^n-B_t^n, B_s-B_t\right)&=W_\psi\left(B_{\underline{s}}^n-B_{\underline{t}}^n, B_s-B_t\right)\\&\leq W_\psi\left(B_{\underline{s}}^n-B_{\underline{t}}^n, B_{\underline{s}}-B_{\underline{t}}\right)+W_\psi\left(B_{\underline{s}}-B_{\underline{t}}, B_s-B_t\right).\end{aligned} \tag{23}$$

Let us treat each term separately. For the first one, Rio's result gives

$$W_\psi\left(\frac{1}{\sqrt{n_s-n_t}}\sum_{k=n_t+1}^{n_s}\xi_k, G\right)\leq C\,(n_s-n_t)^{-1/2},$$

and multiplying by $\sqrt{n_s - n_t}\sqrt{h}$, we get, since $\sqrt{h(n_s - n_t)}G$ is equal to $B_{\underline{s}} - B_{\underline{t}}$ in distribution,

$$W_\psi \left(B_s^n - B_t^n, B_{\underline{s}} - B_{\underline{t}}\right) \leq C\sqrt{h}.$$

Let us deal with the second term of (23). Let $\beta(s,t) := \min(s-t, \underline{s} - \underline{t})$. Then

$$\begin{aligned} W_\psi \left(B_{\underline{s}} - B_{\underline{t}}, B_s - B_t\right) &= W_\psi \left(\mathcal{N}(0, \underline{s} - \underline{t}), \mathcal{N}(0, s-t)\right) \\ &= W_\psi \left(\mathcal{N}(0, \beta(s,t)), \mathcal{N}(0, \beta(s,t)) * \mathcal{N}(0, |s - t - (\underline{s} - \underline{t})|)\right) \\ &\leq W_\psi \left(0, \mathcal{N}(0, |s - t - (\underline{s} - \underline{t})|)\right) \\ &= \sqrt{|s - t - (\underline{s} - \underline{t})|}\, \|G\|_\psi. \end{aligned}$$

But $|s - t - (\underline{s} - \underline{t})| \leq h$, and this concludes the proof. □

Let us finish with a simple consequence of this result that we will use in the sequel.

Corollary 4. *Let $0 < \varepsilon \leq 1$ and let $g : \mathbb{R} \longrightarrow \mathbb{R}$ be an ε-Hölder continuous function. Then there exists a $C > 0$ depending on T such that, for all $x \in \mathbb{R}$ and all $0 \leq t \leq s \leq T$,*

$$\left|\mathbb{E}\left[g\left(B_s^{n,t,x}\right)\right] - \mathbb{E}\left[g\left(B_s^{t,x}\right)\right]\right| \leq C\, \|g\|_\varepsilon\, n^{-\varepsilon/2},$$

and, setting $\delta(t,s) := \max(s-t, \underline{s} - \underline{t})$,

$$\left|\mathbb{E}\left[g\left(B_s^{n,t,x}\right)\left(B_s^{n,t,x} - x\right)\right] - \mathbb{E}\left[g\left(B_s^{t,x}\right)\left(B_s^{t,x} - x\right)\right]\right| \leq C\, \|g\|_\varepsilon\, \delta(t,s)^{1/2}\, n^{-\varepsilon/2}.$$

Proof. Let $x \in \mathbb{R}$ and $0 \leq t \leq s \leq T$. For any coupling (X, Y) of $B_s^{n,t,x}$ and $B_s^{t,x}$, using Hölder's inequality when $0 < \varepsilon < 1$,

$$W_1\left(g\left(B_s^{n,t,x}\right), g\left(B_s^{t,x}\right)\right) \leq \mathbb{E}\left[|g(X) - g(Y)|\right] \leq \|g\|_\varepsilon\, \mathbb{E}\left[|X - Y|^\varepsilon\right] \leq \|g\|_\varepsilon\, \mathbb{E}\left[|X - Y|\right]^\varepsilon.$$

Thus, we have, by (22) for $r = 1$,

$$W_1\left(g\left(B_s^{n,t,x}\right), g\left(B_s^{t,x}\right)\right) \leq \|g\|_\varepsilon\, W_1\left(B_s^{n,t,x}, B_s^{t,x}\right)^\varepsilon \leq \|g\|_\varepsilon\, c_1^\varepsilon \left(\frac{T}{n}\right)^{\varepsilon/2}.$$

Choosing $f(x) = x$ in (21), this implies the first result.

Let us prove the second assertion. We start by observing that, since $B_s - B_t$ and $B_s^n - B_t^n$ are centered random variables, we have, setting $h(y) := (g(x+y) - g(x))y$,

$$\begin{aligned} \mathbb{E}\left[g\left(B_s^{t,x}\right)\left(B_s^{t,x} - x\right)\right] &= \mathbb{E}\left[\left(g\left(x + B_s - B_t\right) - g(x)\right)\left(B_s - B_t\right)\right] = \mathbb{E}\left[h(B_s - B_t)\right], \\ \mathbb{E}\left[g\left(B_s^{n,t,x}\right)\left(B_s^{n,t,x} - x\right)\right] &= \mathbb{E}\left[\left(g\left(x + B_s^n - B_t^n\right) - g(x)\right)\left(B_s^n - B_t^n\right)\right] = \mathbb{E}\left[h(B_s^n - B_t^n)\right]. \end{aligned}$$

Let us remark that, for any real numbers y and z, $|h(y)| \leq \|g\|_\varepsilon\, |y|^{1+\varepsilon}$, and using the fact that $|y - z|^{1-\varepsilon} \leq |y|^{1-\varepsilon} + |z|^{1-\varepsilon}$,

$$\begin{aligned} |h(y) - h(z)| &\leq |g(x+y) - g(x)|\, |y - z| + |(g(x+y) - g(x)) - (g(x+z) - g(x))|\, |z| \\ &\leq \|g\|_\varepsilon \left(|y - z|\, |y|^\varepsilon + |y - z|^\varepsilon\, |z|\right) \\ &\leq \|g\|_\varepsilon \left(|y - z|^\varepsilon |y - z|^{1-\varepsilon}\, |y|^\varepsilon + |y - z|^\varepsilon\, |z|\right) \\ &\leq \|g\|_\varepsilon |y - z|^\varepsilon \left(|y| + |y|^\varepsilon\, |z|^{1-\varepsilon} + |z|\right). \end{aligned}$$

Young's inequality, $|y|^\varepsilon\,|z|^{1-\varepsilon} \le \varepsilon\,|y| + (1-\varepsilon)\,|z| \le |y| + |z|$, leads to

$$|h(y) - h(z)| \le 2\,\|g\|_\varepsilon\,|y-z|^\varepsilon\,(|y| + |z|)\,. \tag{24}$$

In the case where $\underline{s} = \underline{t}$ we have

$$\begin{aligned} W_1\left(h(B^n_s - B^n_t), h(B_s - B_t)\right) &\le \mathbb{E}\left[|h(B_s - B_t)|\right] \le \|g\|_\varepsilon\,\mathbb{E}\left[|B_t - B_s|^{(1+\varepsilon)}\right] \\ &\le \|g\|_\varepsilon\,(s-t)^{(1+\varepsilon)/2}\,\mathbb{E}\left[|G|^{(1+\varepsilon)}\right], \end{aligned}$$

where $\mathrm{law}(G) = \mathcal{N}(0,1)$. Since $\underline{s} = \underline{t}$ implies $(s-t)^{1/2} = \delta(t,s)^{1/2}$ and $(s-t)^{\varepsilon/2} \le h^{\varepsilon/2}$, we have

$$W_1\left(h(B^n_s - B^n_t), h(B_s - B_t)\right) \le \|g\|_\varepsilon\,\delta(t,s)^{1/2}\,(\tfrac{T}{n})^{\varepsilon/2}$$

using the fact that $\mathbb{E}\left[|G|^{(1+\varepsilon)}\right] \le 1$.

Let us turn to the case $\underline{t} < \underline{s}$. For any coupling (X, Y) of $B^n_s - B^n_t$ and $B_s - B_t$, using (20) and (24),

$$\begin{aligned} W_1\left(h(B^n_s - B^n_t), h(B_s - B_t)\right) &\le \mathbb{E}\left[|h(X) - h(Y)|\right] \\ &\le 2\,\|g\|_\varepsilon\,\mathbb{E}\left[|X-Y|^\varepsilon\,(|X| + |Y|)\right], \end{aligned}$$

and, by Hölder's inequality with $p = 2/\varepsilon$ and $q = 2/(2-\varepsilon)$,

$$\begin{aligned} W_1\left(h(B^n_s - B^n_t), h(B_s - B_t)\right) &\le 2\,\|g\|_\varepsilon\,\mathbb{E}\left[|X-Y|^2\right]^{\varepsilon/2}\,\left(\mathbb{E}[|X|^{2/(2-\varepsilon)}]^{1-\varepsilon/2} + \mathbb{E}[|Y|^{2/(2-\varepsilon)}]^{1-\varepsilon/2}\right) \\ &\le 2\,\|g\|_\varepsilon\,\mathbb{E}\left[|X-Y|^2\right]^{\varepsilon/2}\,\left((\underline{s}-\underline{t})^{1/2} + (s-t)^{1/2}\right). \end{aligned}$$

From (20) it follows that

$$\begin{aligned} W_1\left(h(B^n_s - B^n_t), h(B_s - B_t)\right) &\le 2\,\|g\|_\varepsilon\,W_2\left(B^n_s - B^n_t, B_s - B_t\right)^\varepsilon\,\left((\underline{s}-\underline{t})^{1/2} + (s-t)^{1/2}\right) \\ &\le 4\,c_2^\varepsilon\,\|g\|_\varepsilon\,\left(\frac{T}{n}\right)^{\varepsilon/2}\,\delta(t,s)^{1/2}, \end{aligned}$$

where we have used (22) for $r = 2$.

Thus, for $0 \le t \le s \le T$,

$$W_1\left(h(B^n_s - B^n_t), h(B_s - B_t)\right) \le C\,\|g\|_\varepsilon\,n^{-\varepsilon/2}\,\delta(t,s)^{1/2},$$

and the result follows as before by choosing $f(x) = x$ in (21). □

4. Regularity results on u, U^n, ∇u and Δ^n.

Let us start by known regularity properties of the function u that follow from classical a priori estimates for BSDEs.

Lemma 5. *Under Assumption (A1) there exists a constant $C > 0$ depending on (T, ε, f, g) such that, for all $(t,x) \in [0,T] \times \mathbb{R}$,*

$$|u(t,x)| \le C\,(1+|x|)^\varepsilon, \qquad \|u(t,\cdot)\|_\varepsilon \le C, \qquad \|u(\cdot,x)\|_{\varepsilon/2} \le C\,(1+|x|)^\varepsilon.$$

Proof. The first two results follow directly from classical a priori estimates for BSDEs, see e.g. [8, Proposition 2.1]. The last one ensues from the following upper bound: for any real x and for $0 \le r \le t \le T$,

$$\begin{aligned}
|u(r,x) - u(t,x)| &= |\mathbb{E}[Y_r^{r,x}] - \mathbb{E}[Y_t^{t,x}]| \\
&\le |\mathbb{E}[Y_r^{r,x} - Y_t^{r,x}]| + |\mathbb{E}[Y_t^{r,x} - Y_t^{t,x}]| \\
&\le \int_r^t \mathbb{E}[|f(s, B_s^{r,x}, Y_s^{r,x}, Z_s^{r,x})|]\, ds + \mathbb{E}[|Y_t^{r,x} - Y_t^{t,x}|] \\
&\le C \int_r^t \mathbb{E}[1 + |B_s^{r,x}|^\varepsilon + |Y_s^{r,x}| + |Z_s^{r,x}|]\, ds + \mathbb{E}[|Y_t^{r,x} - Y_t^{t,x}|].
\end{aligned}$$

Since the norm in $\mathcal{S}^2 \times \mathcal{H}^2$ of $(Y^{r,x}, Z^{r,x})$ is of order $(1+|x|)^\varepsilon$, we use Cauchy-Schwarz inequality to bound the first term and a priori estimates enable (similarly as in the proof of [8, Proposition 4.1]) to bound the second term. □

Next we extend [19, Theorem 3.2] to the case where $f(t,\cdot,y,z)$ is Hölder continuous.

Lemma 6. *Recall the notation (5) and let Assumption (A1) hold.*

(a) *The function u belongs to $\mathcal{C}^{0,1}([0,T[\times\mathbb{R})$ and, for all $(t,x) \in [0,T[\times\mathbb{R}$, we have,*

$$Z_s^{t,x} = \nabla u(s, B_s^{t,x}) \quad \text{for a.e. } (s,\omega) \in [t,T[\times\Omega, \tag{25}$$

as well as (7) i.e.

$$\nabla u(t,x) = \mathbb{E}\left(g(B_T^{t,x})\frac{B_T - B_t}{T-t}\right) + \mathbb{E}\left(\int_t^T F(s, B_s^{t,x})\frac{B_s - B_t}{s-t} ds\right).$$

(b) *Moreover, there exists a constant $C > 0$ depending on (T, ε, f, g) such that,*

(i) $\|\nabla u(t,\cdot)\|_\varepsilon \le \frac{C}{\sqrt{T-t}}$ *for all* $t \in [0,T[$,

(ii) $|\nabla u(t,x)| \le \frac{C}{(T-t)^{(1-\varepsilon)/2}}$ *for all* $(t,x) \in [0,T[\times\mathbb{R}$.

Consequently, for $\mathbb{E}_r := \mathbb{E}[\,\cdot\,|\mathcal{F}_r]$,

$$\nabla u(r, B_r^{t,x}) = \mathbb{E}_r\left(g(B_T^{t,x})\frac{B_T - B_r}{T-r}\right) + \mathbb{E}_r\left(\int_r^T F(s, B_s^{t,x})\frac{B_s - B_r}{s-r} ds\right) \quad \text{a.s. for } r \in [t,T[. \tag{26}$$

Proof of Lemma 6. The proof is divided into two steps.

Step 1. We assume in addition that f is Lipschitz continuous w.r.t. x. Then according to [19], we have only the second point to prove and we know that, for some constant C,

$$|\nabla u(t,x)| \le \frac{C(1+|x|)}{\sqrt{T-t}} \quad \text{on } [0,T[\times\mathbb{R}. \tag{27}$$

(bi) The representation (26) yields to

$$\begin{aligned}
\|\nabla u(r, B_r^{t,x}) - \nabla u(r, B_r^{t,y})\|_{L^2} \le{}& \left\|\mathbb{E}_r\left(\left(g(B_T^{t,x}) - g(B_T^{t,y})\right)\frac{B_T - B_r}{T-r}\right)\right\|_{L^2} \\
&+ \int_r^T \left\|\mathbb{E}_r\left(\left(F(s, B_s^{t,x}) - F(s, B_s^{t,y})\right)\frac{B_s - B_r}{s-r}\right)\right\|_{L^2} ds.
\end{aligned} \tag{28}$$

Since g is ε-Hölder continuous we get

$$\left|\mathbb{E}_r\left((g(B_T^{t,x})-g(B_T^{t,y}))\frac{B_T-B_r}{T-r}\right)\right| \le \|g\|_\varepsilon |x-y|^\varepsilon \frac{1}{\sqrt{T-r}}.$$

Similarly, we obtain by the conditional Cauchy-Schwarz inequality the estimate

$$\left\|\mathbb{E}_r\left((F(s,B_s^{t,x})-F(s,B_s^{t,y}))\frac{B_s-B_r}{s-r}\right)\right\|_{L^2} \le \left\|F(s,B_s^{t,x})-F(s,B_s^{t,y})\right\|_{L^2}\frac{1}{\sqrt{s-r}}.$$

Using (3) for f, we have

$$\begin{aligned}&|F(s,B_s^{t,x})-F(s,B_s^{t,y})|\\ &\quad\le \|f_x\|_\varepsilon |x-y|^\varepsilon + \|f_y\|_{\mathrm{Lip}}|u(s,B_s^{t,x})-u(s,B_s^{t,y})| + \|f_z\|_{\mathrm{Lip}}|\nabla u(s,B_s^{t,x})-\nabla u(s,B_s^{t,y})|,\end{aligned}$$

and the Hölder continuity of u stated in Lemma 5 yields

$$\begin{aligned}&|F(s,B_s^{t,x})-F(s,B_s^{t,y})|\\ &\qquad\le (\|f_x\|_\varepsilon + C\|f_y\|_{\mathrm{Lip}})|x-y|^\varepsilon + \|f_z\|_{\mathrm{Lip}}|\nabla u(s,B_s^{t,x})-\nabla u(s,B_s^{t,y})|. \quad (29)\end{aligned}$$

By combining the above estimates we conclude from (28) that

$$\begin{aligned}\|\nabla u(r,B_r^{t,x})-\nabla u(r,B_r^{t,y})\|_{L^2} &\le \|g\|_\varepsilon\frac{|x-y|^\varepsilon}{\sqrt{T-r}} + 2\left(\|f_x\|_\varepsilon + C\,\|f_y\|_{\mathrm{Lip}}\right)|x-y|^\varepsilon\sqrt{T-r}\\ &\quad + \int_r^T \left\|\nabla u(s,B_s^{t,x})-\nabla u(s,B_s^{t,y})\right\|_{L^2}\frac{\|f_z\|_{\mathrm{Lip}}}{\sqrt{s-r}}ds\\ &\le C_1\frac{|x-y|^\varepsilon}{\sqrt{T-r}} + \int_r^T \left\|\nabla u(s,B_s^{t,x})-\nabla u(s,B_s^{t,y})\right\|_{L^2}\frac{\|f_z\|_{\mathrm{Lip}}}{\sqrt{s-r}}ds\end{aligned}$$

for $C_1 := \|g\|_\varepsilon + 2T(\|f_x\|_\varepsilon + C\,\|f_y\|_{\mathrm{Lip}})$. Because of (27) we have

$$\left\|\nabla u(s,B_s^{t,x})-\nabla u(s,B_s^{t,y})\right\|_{L^2} \le C_2\frac{1+|x|+|y|}{(T-s)^{1/2}}$$

with $C_2 = C_2(C,T) > 0$. Hence we may apply Gronwall's lemma (Lemma 14) and get

$$\|\nabla u(r,B_r^{t,x})-\nabla u(r,B_r^{t,y})\|_{L^2} \le C_1 c_0\frac{|x-y|^\varepsilon}{\sqrt{T-r}},$$

for some $c_0 = c_0(T,\|f_z\|_{\mathrm{Lip}}) > 0$. Especially, for $r = t$ this implies

$$|\nabla u(t,x)-\nabla u(t,y)| \le C\frac{|x-y|^\varepsilon}{\sqrt{T-t}}$$

for some $C = C(T,\varepsilon,f,g) > 0$.

(bii) We first notice that for any ε-Hölder continuous function k and for all $0 \le t < s \le T$ we have

$$\left|\mathbb{E}\left[k(B_s^{t,x})\frac{B_s-B_t}{s-t}\right]\right| = \left|\mathbb{E}\left[(k(B_s^{t,x})-k(x))\frac{B_s-B_t}{s-t}\right]\right| \le \frac{\|k\|_\varepsilon}{(s-t)^{(1-\varepsilon)/2}}. \quad (30)$$

Therefore, we obtain from (7) that

$$\begin{aligned}|\nabla u(t,x)| &\le \left|\mathbb{E}\left(g(B_T^{t,x})\frac{B_T-B_t}{T-t}\right)\right| + \left|\mathbb{E}\left(\int_t^T F(s,B_s^{t,x})\frac{B_s-B_t}{s-t}ds\right)\right|\\ &\le \frac{\|g\|_\varepsilon}{(T-t)^{(1-\varepsilon)/2}} + \int_t^T \frac{\|F(s,\cdot)\|_\varepsilon}{(s-t)^{(1-\varepsilon)/2}}ds.\end{aligned}$$

Using (29) for $s = t$ and taking into account that ∇u satisfies (bi) we get

$$\|F(s,\cdot)\|_\varepsilon \leq \frac{C}{\sqrt{T-s}} \tag{31}$$

for some $C = C(T,\varepsilon,f,g) > 0$. This finishes the proof of the first step.

Step 2. General case. The proof relies on a regularization procedure and is postponed to appendix A.3. □

Remark 7. From now on we will always use the continuous version of $Z_s^{t,x}$ given by $\nabla u(s, B_s^{t,x})$.

Lemma 8. *For all $(t,x) \in [0,T[\times\mathbb{R}$ and for $n \geq n_0 \in \mathbb{N}^*$, with n_0 defined as in Lemma 12, we have*

(i) $|U^n(t,x)| \leq C(1+|x|)^\varepsilon$,

(ii) $|\Delta^n(t,x)| \leq \frac{C_n}{(T-t)^{(1-\varepsilon)/2}}$,

where $C > 0$ depends on (T,ε,f,g) and $C_n > 0$ depends on (T,ε,f,g,n).

Proof. The result on U^n ensues from Lemma 12, by choosing $\overline{f} = 0$ and $\overline{g} = 0$. Let us prove the result on Δ^n. By (15) and (10) we have that

$$\begin{aligned}\Delta^n(t,x) &= [\nabla^n U^n](\underline{t}+h,x)\\ &= \tfrac{1}{2\sqrt{h}}(U^n(\underline{t}+h,x+\sqrt{h}) - U^n(\underline{t}+h,x-\sqrt{h}))\\ &= \mathbb{E}\left[U^n(\underline{t}+h, B^{n,\underline{t},x}_{\underline{t}+h})\tfrac{B^{n,\underline{t},x}_{\underline{t}+h}-x}{h}\right].\end{aligned}$$

We want to use (17), where we realize that

$$\mathbb{E}\left[g\left(B_T^{n,\underline{t}+h,B^{n,\underline{t},x}_{\underline{t}+h}}\right)\tfrac{B^{n,\underline{t},x}_{\underline{t}+h}-x}{h}\right] = \mathbb{E}\left[g\left(B_T^{n,t,x}\right)\tfrac{B^{n,\underline{t},x}_{\underline{t}+h}-x}{h}\right] = \mathbb{E}\left[g(B_T^{n,t,x})\frac{B_T^{n,t,x}-x}{T-\underline{t}}\right].$$

A similar argument can be used for the integral expression so that we get

$$\Delta^n(t,x) = \mathbb{E}\left[g(B_T^{n,t,x})\frac{B_T^{n,t,x}-x}{T-\underline{t}}\right] + \mathbb{E}\left[\int_{\underline{t}+h}^T f\left(\overline{s},\Theta_s^{n,t,x}\right)\frac{B_s^{n,t,x}-x}{\underline{s}-\underline{t}}\,ds\right]. \tag{32}$$

Then

$$\begin{aligned}\Delta^n(t,x) =& \mathbb{E}\left((g(B_T^{n,t,x})-g(x))\frac{B_T^n-B_t^n}{T-\underline{t}}\right)\\ &+\mathbb{E}\left(\int_{\underline{t}+h}^T (f(\overline{s},\Theta_s^{n,t,x}) - f(\overline{s},x,U^n(s,x),\Delta^n(s,x)))\frac{B_s^n-B_t^n}{\underline{s}-\underline{t}}ds\right)\\ =:& G+F.\end{aligned}$$

Since g is ε-Hölder, $|G|$ is bounded by $\frac{\|g\|_\varepsilon}{(T-t)^{(1-\varepsilon)/2}}$. Concerning the second term, we get, since f satisfies (3),

$$\begin{aligned}&f(\overline{s},\Theta_s^{n,t,x}) - f(\overline{s},x,U^n(s,x),\Delta^n(s,x))\\ &\leq \|f_x\|_\varepsilon |B_s^n - B_t^n|^\varepsilon + \|f_y\|_{\mathrm{Lip}}|U^n(s,B_s^{n,t,x}) - U^n(x,s)| + \|f_z\|_{\mathrm{Lip}}|\Delta^n(s,B_s^{n,t,x}) - \Delta^n(s,x)|.\end{aligned}$$

We will use that $U^n(t,x)$ and $\Delta^n(t,x)$ are ε-Hölder continuous in x, i.e.

$$|U^n(t,x) - U^n(t,y)| + |\Delta^n(t,x) - \Delta^n(t,y)| \leq c(h)|x-y|^\varepsilon,$$

where $c(h)$ tends to infinity when h tends to 0. For U^n, Lemma 12 with $(\bar{x},\bar{g},\bar{f}) = (y,g,f)$ gives

$$|U^n(t,x) - U^n(t,y)| \leq c_0|x-y|^\varepsilon \tag{33}$$

while for Δ^n this is an immediate consequence of Remark 1 and (33) with $c(h) = c_0 + \frac{c_0}{\sqrt{h}}$. Then

$$\begin{aligned}|F| &\leq (\|f_x\|_\varepsilon + c(h)\|f_y\|_{\mathrm{Lip}} + c(h)\|f_z\|_{\mathrm{Lip}}) \int_{\underline{t}+h}^T \mathbb{E}\left(\frac{|B_s^n - B_{\underline{t}}^n|^{1+\varepsilon}}{\underline{s}-\underline{t}}\right) ds \\ &\leq (\|f_x\|_\varepsilon + c(h)\|f_y\|_{\mathrm{Lip}} + c(h)\|f_z\|_{\mathrm{Lip}}) \int_{\underline{t}+h}^T \frac{1}{(\underline{s}-\underline{t})^{(1-\varepsilon)/2}} ds.\end{aligned}$$

Since $\frac{1}{(\underline{s}-\underline{t})^{(1-\varepsilon)/2}} \leq \frac{1}{(s-(\underline{t}+h))^{(1-\varepsilon)/2}}$ for $s \in]\underline{t}+h, T]$ we get that $|F| \leq C_n$.

□

Proposition 9. *Under (A1), there exists a constant $C > 0$ depending on (T,ε,f,g) such that, for all $x \in \mathbb{R}$,*

$$|\nabla u(t,x) - \nabla u(r,x)| \leq C\frac{(t-r)^{\varepsilon/2}}{\sqrt{T-t}} \quad \text{for all } \ 0 \leq r < t < T. \tag{34}$$

Proof. From Lemma 6, we know that, for $(t,x) \in [0,T[\times\mathbb{R}$,

$$\nabla u(t,x) = \mathbb{E}\left[g\left(B_T^{t,x}\right) H(t,T) + \int_t^T F\left(s, B_s^{t,x}\right) H(t,s)\, ds\right], \tag{35}$$

where we have set

$$H(t,s) := \frac{B_s - B_t}{s-t}, \quad \text{for } \ 0 \leq t < s \leq T.$$

It holds $\mathbb{E}[H(t,s)] = 0$ and $\|H(t,s)\|_{\mathrm{L}^2} = \frac{1}{\sqrt{s-t}}$. We also have, for $0 \leq r \leq t < s$,

$$\begin{aligned}\mathbb{E}|H(r,s) - H(t,s)|^2 &= \frac{1}{s-r} - \frac{2}{(s-t)(s-r)}\mathbb{E}[(B_s - B_r)(B_s - B_t)] + \frac{1}{s-t} \\ &= \frac{1}{s-t} - \frac{1}{s-r} = \frac{t-r}{(s-r)(s-t)}.\end{aligned}$$

Let us observe that, for $0 \leq r < t < s \leq T$ and any ε-Hölder continuous function h, it holds

$$\left|\mathbb{E}\left[h(B_s^{t,x})H(t,s)\right] - \mathbb{E}\left[h(B_s^{r,x})H(r,s)\right]\right| \leq 2\,\|h\|_\varepsilon \frac{(t-r)^{\varepsilon/2}}{(s-t)^{1/2}}.$$

Indeed, we have

$$\begin{aligned}&\left|\mathbb{E}\left[h(B_s^{t,x})H(t,s)\right] - \mathbb{E}\left[h(B_s^{r,x})H(r,s)\right]\right| \\ &\quad\leq \mathbb{E}\left|\left[h(B_s^{t,x}) - h(B_s^{r,x})\right] H(t,s)\right| + \mathbb{E}\left|\left[h(B_s^{r,x}) - h(x)\right]\left[H(t,s) - H(r,s)\right]\right| \\ &\quad\leq \|h\|_\varepsilon \left(\mathbb{E}\left[|B_t - B_r|^\varepsilon |H(t,s)|\right] + \mathbb{E}\left[|B_s - B_r|^\varepsilon |H(t,s) - H(r,s)|\right]\right),\end{aligned}$$

and, from Cauchy-Schwarz inequality, we deduce that

$$\begin{aligned}\left|\mathbb{E}\left[h(B_s^{t,x})H(t,s)\right]-\mathbb{E}\left[h(B_s^{r,x})H(r,s)\right]\right| &\leq \|h\|_\varepsilon\left[\frac{(t-r)^{\varepsilon/2}}{(s-t)^{1/2}}+(s-r)^{\varepsilon/2}\frac{(t-r)^{1/2}}{(s-r)^{1/2}(s-t)^{1/2}}\right]\\ &= \|h\|_\varepsilon\frac{(t-r)^{\varepsilon/2}}{(s-t)^{1/2}}\left[1+(t-r)^{\frac{1}{2}-\frac{\varepsilon}{2}}(s-r)^{\frac{\varepsilon}{2}-\frac{1}{2}}\right]\\ &\leq 2\|h\|_\varepsilon\frac{(t-r)^{\varepsilon/2}}{(s-t)^{1/2}}.\end{aligned}$$

Coming back to (35), we write, for $0\leq r\leq t<T$,

$$\begin{aligned}\nabla u(r,x)=\mathbb{E}\left[g\left(B_T^{r,x}\right)H(r,T)+\int_t^T F\left(s,B_s^{r,x}\right)H(r,s)\,ds\right]\\ +\mathbb{E}\left[\int_r^t\left(F\left(s,B_s^{r,x}\right)-F(s,x)\right)H(r,s)\,ds\right],\end{aligned}$$

to have, taking into account the fact that $\|F(s,\cdot)\|_\varepsilon\leq C(T-s)^{-1/2}$ by (31),

$$|\nabla u(t,x)-\nabla u(r,x)|\leq C\left[(t-r)^{\varepsilon/2}\left(\frac{\|g\|_\varepsilon}{(T-t)^{1/2}}+\int_t^T\frac{\|F(s,\cdot)\|_\varepsilon}{(s-t)^{1/2}}ds\right)+\int_r^t\frac{\|F(s,\cdot)\|_\varepsilon}{(s-r)^{(1-\varepsilon)/2}}ds\right],$$

$$\begin{aligned}&\leq C\left[(t-r)^{\varepsilon/2}\left(\frac{\|g\|_\varepsilon}{(T-t)^{1/2}}+\int_t^T\frac{ds}{(T-s)^{1/2}(s-t)^{1/2}}\right)\right.\\ &\quad\left.+\int_r^t\frac{ds}{(T-s)^{1/2}(s-r)^{(1-\varepsilon)/2}}\right]\\ &\leq C\left[(t-r)^{\varepsilon/2}\left(\frac{\|g\|_\varepsilon}{(T-t)^{1/2}}+B(\tfrac{1}{2},\tfrac{1}{2})\right)+\frac{(t-r)^{(\varepsilon+1)/2}}{(T-t)^{1/2}}\right].\end{aligned}$$

□

5. Main results.

In this section, we state the main result of this paper which gives the rate of convergence in the Wasserstein distance between the solution to the BSDE (4) and the solution to the BSDE driven by the scaled random walk (12). For the following we want to remind the reader of Remark 7.

Theorem 10. *Under (A1), for any $r\in[1,\infty[$, there exists a constant $C_r>0$ depending at most on $(T,\alpha,\varepsilon,f,g,r)$ such that for all $x\in\mathbb{R}$,*

(i) $W_r\left(Y_s^{n,t,x},Y_s^{t,x}\right)\leq C_r\left(1+|x|\right)^\varepsilon n^{-(\alpha\wedge\frac{\varepsilon}{2})}$ *for all* $0\leq t\leq s\leq T$,

(ii) $W_r\left(Z_s^{n,t,x},Z_s^{t,x}\right)\leq C_r\frac{(1+|x|)^\varepsilon}{\sqrt{T-s}}n^{-(\alpha\wedge\frac{\varepsilon}{2})}$ *for all* $s\in[t,T[$.

This result is a consequence of the following proposition which gives the rate of the point-wise convergence of U^n, solution to (11), towards the solution u of the semilinear heat equation (1).

Proposition 11. *Under (A1) there exists a constant $C>0$ depending at most on $(T,\alpha,\varepsilon,f,g)$ such that*

(i) $|u(t,x) - U^n(t,x)| \le C\,(1+|x|)^{\varepsilon}\, n^{-(\alpha\wedge\frac{\varepsilon}{2})}$ *for all* $(t,x) \in \mathbb{R}\times[0,T]$,

(ii) $|\nabla u(t,x) - \Delta^n(t,x)| \le C\,\frac{(1+|x|)^{\varepsilon}}{\sqrt{T-t}}\, n^{-(\alpha\wedge\frac{\varepsilon}{2})}$ *for all* $(t,x) \in \mathbb{R}\times[0,T[$.

Proof. We split the proof into three parts. We begin by studying $|u - U^n|$, we proceed by obtaining an estimate for $\nabla u - \Delta^n$, and then we conclude with a Gronwall argument.

Estimate for $|u - U^n|$. From (6) and (17) we conclude that

$$\begin{aligned} |u(t,x) - U^n(t,x)| \le{}& \left|\mathbb{E}\left[g\left(B_T^{t,x}\right)\right] - \mathbb{E}\left[g\left(B_T^{n,t,x}\right)\right]\right| \\ &+ \left|\mathbb{E}\left[\int_t^T F\left(s, B_s^{t,x}\right) ds\right] - \mathbb{E}\left[\int_{\underline{t}}^T f\left(\overline{s}, \Theta_s^{n,t,x}\right) ds\right]\right|. \end{aligned} \tag{36}$$

Let F_n be the function given by

$$F_n(s,x) := f(s,x,U^n(s,x),\Delta^n(s,x)) \quad \text{for} \quad (s,x) \in [0,T[\times\mathbb{R}. \tag{37}$$

Using the notation (18) we also have that $F_n(s, B_s^{n,t,x}) = f(s, \Theta_s^{n,t,x})$. With this notation in hand, we have, taking into account (13) and (16),

$$\begin{aligned} \mathbb{E}\left[\int_{\underline{t}}^T f\left(\overline{s}, \Theta_s^{n,t,x}\right) ds\right] = \mathbb{E}\left[\int_{\underline{t}}^T \left(f\left(\overline{s}, \Theta_s^{n,t,x}\right) - f\left(s, \Theta_s^{n,t,x}\right)\right)\right] + \mathbb{E}\left[\int_t^T F_n\left(s, B_s^{n,t,x}\right) ds\right] \\ + \mathbb{E}\left[\int_{\underline{t}}^t f\left(s, B_s^{n,t,x}, Y_s^{n,t,x}, Z_s^{n,t,x}\right) ds\right]. \end{aligned}$$

In view of the regularity of f in time, we have

$$\left|\mathbb{E}\left[\int_{\underline{t}}^T \left(f\left(\overline{s}, \Theta_s^{n,t,x}\right) - f\left(s, \Theta_s^{n,t,x}\right)\right) ds\right]\right| \le C\, n^{-\alpha}.$$

Moreover, the Cauchy-Schwarz inequality leads to

$$\left|\mathbb{E}\left[\int_{\underline{t}}^t f\left(s, B_s^{n,t,x}, Y_s^{n,t,x}, Z_s^{n,t,x}\right) ds\right]\right| \le (t-\underline{t})^{1/2}\, \mathbb{E}\left[\int_{\underline{t}}^t \left|f\left(s, B_s^{n,t,x}, Y_s^{n,t,x}, Z_s^{n,t,x}\right)\right|^2 ds\right]^{1/2},$$

and, taking into account the growth of f, we have

$$\begin{aligned} &\left|\mathbb{E}\left[\int_{\underline{t}}^t f\left(s, B_s^{n,t,x}, Y_s^{n,t,x}, Z_s^{n,t,x}\right) ds\right]\right| \\ &\quad\le C\, h^{1/2}\, \mathbb{E}\left[\int_{\underline{t}}^T \left(1 + |B_s^{n,t,x}|^{2\varepsilon} + |Y_s^{n,t,x}|^2 + |Z_s^{n,t,x}|^2\right) ds\right]^{1/2} \\ &\quad\le C\, n^{-1/2}\, (1+|x|)^{\varepsilon}, \end{aligned}$$

where we have used Lemma 12 to get

$$\begin{aligned} \mathbb{E}\left(\sup_{\underline{t}\le s\le T} |Y_s^{n,t,x}|^2 + \int_{\underline{t}}^T |Z_s^{n,t,x}|^2 ds\right) \le{}& C\mathbb{E}\left(|g(B_T^{n,t,x})|^2 + \int_{\underline{t}}^T |f(s, B_{s-}^{n,t,x}, 0, 0)|^2 ds\right) \\ \le{}& C(1+|x|)^{2\varepsilon}. \end{aligned}$$

Coming back to (36), we derive the following inequality

$$|u(t,x)-U^n(t,x)| \le \left|\mathbb{E}\left[g\left(B_T^{t,x}\right)\right]-\mathbb{E}\left[g\left(B_T^{n,t,x}\right)\right]\right| + \left|\int_t^T \mathbb{E}\left[F\left(s,B_s^{t,x}\right)\right]ds - \int_t^T \mathbb{E}\left[F_n\left(s,B_s^{n,t,x}\right)\right]ds\right| + C\,(1+|x|)^{\varepsilon}\,n^{-(\alpha\wedge\frac{1}{2})}. \tag{38}$$

From Corollary 4 we get

$$\left|\mathbb{E}\left[g\left(B_T^{t,x}\right)\right]-\mathbb{E}\left[g\left(B_T^{n,t,x}\right)\right]\right| \le C\,\|g\|_{\varepsilon}\,n^{-\frac{\varepsilon}{2}}.$$

We split the second term on the RHS of (38) into two parts

$$\left|\int_t^T \mathbb{E}\left[F\left(s,B_s^{t,x}\right)\right]ds - \int_t^T \mathbb{E}\left[F_n\left(s,B_s^{n,t,x}\right)\right]ds\right| \le \int_t^T \left|\mathbb{E}\left[F\left(s,B_s^{t,x}\right)\right]-\mathbb{E}\left[F\left(s,B_s^{n,t,x}\right)\right]\right|ds + \int_t^T \mathbb{E}\left[|F-F_n|\left(s,B_s^{n,t,x}\right)\right]ds.$$

Since F has the regularity (31), Corollary 4 gives

$$\int_t^T \left|\mathbb{E}\left[F\left(s,B_s^{t,x}\right)\right]-\mathbb{E}\left[F\left(s,B_s^{n,t,x}\right)\right]\right|ds \le C\int_t^T \frac{ds}{\sqrt{T-s}}\,ds\,n^{-\frac{\varepsilon}{2}} \le C\,n^{-\frac{\varepsilon}{2}}.$$

By the above estimates we derive from (38) the inequality

$$|u(t,x)-U^n(t,x)| \le C\,(1+|x|)^{\varepsilon}\,n^{-(\alpha\wedge\frac{\varepsilon}{2})} + \int_t^T \mathbb{E}\left[|F-F_n|\left(s,B_s^{n,t,x}\right)\right]ds. \tag{39}$$

Coming back to the definition of F and F_n (see (5) and (37)) and using the Lipschitz continuity of f with respect to (y,z), we have

$$|F-F_n|\left(s,B_s^{n,t,x}\right) \le \|f\|_{\mathrm{Lip}}\left[|u-U^n|(s,B_s^{n,t,x})+|\nabla u-\Delta^n|(s,B_s^{n,t,x})\right].$$

Setting for simplicity, for $s\in[0,T]$,

$$\beta_n(s) := \sup_{x\in\mathbb{R}}\left\{\frac{|u-U^n|(s,x)}{(1+|x|)^{\varepsilon}}\right\} \quad \text{and} \quad \gamma_n(s) := \sup_{x\in\mathbb{R}}\left\{\frac{|\nabla u-\Delta^n|(s,x)}{(1+|x|)^{\varepsilon}}\right\}$$

for $s\in[0,T[$, Lemma 5, Lemma 6 and Lemma 8 imply that, for some $C>0$ and $C_n>0$,

$$\beta_n(s)\le C \quad \text{for } s\in[0,T] \quad \text{and} \quad \gamma_n(s)\le \frac{C_n}{(T-s)^{(1-\varepsilon)/2}} \quad \text{for } s\in[0,T[,$$

respectively. We deduce the following estimate

$$|F-F_n|\left(s,B_s^{n,t,x}\right) \le \|f\|_{\mathrm{Lip}}(1+|B_s^{n,t,x}|)^{\varepsilon}\,(\beta_n(s)+\gamma_n(s)) \tag{40}$$

and get, coming back to (39), for $0\le t\le T$ and for any $x\in\mathbb{R}$,

$$|u(t,x)-U^n(t,x)| \le C\,(1+|x|)^{\varepsilon}\left(n^{-(\alpha\wedge\frac{\varepsilon}{2})}+\int_t^T(\beta_n(s)+\gamma_n(s))\,ds\right).$$

We end up with the inequality

$$\beta_n(t) \le C\left(n^{-(\alpha\wedge\frac{\varepsilon}{2})}+\int_t^T(\beta_n(s)+\gamma_n(s))\,ds\right), \quad t\in[0,T],$$

and since γ_n belongs to $\mathrm{L}^1[0,T]$, Gronwall's inequality (Lemma 13) gives

$$\beta_n(t) \le C\left(n^{-(\alpha\wedge\frac{\varepsilon}{2})}+\int_t^T \gamma_n(s)ds\right), \quad t\in[0,T]. \tag{41}$$

Estimate for $|\nabla u - \Delta^n|$. In order to take advantage of the previous inequality, we need to estimate $\gamma_n(s)$. To do this, we use the representations (7) and (32). We will divide the study into two parts

$$|\nabla u(t,x) - \Delta^n(t,x)| \leq |\ g \text{ difference } | + |\ f \text{ difference } |.$$

Study of the g difference. We have

$$\begin{aligned}
&\mathbb{E}\left[g\left(B_T^{t,x}\right)\frac{B_T^{t,x}-x}{T-t}\right] - \mathbb{E}\left[g\left(B_T^{n,t,x}\right)\frac{B_T^{n,t,x}-x}{T-\underline{t}}\right] \\
&\quad = \mathbb{E}\left[g(x+B_T-B_t)(B_T-B_t)\right]\left(\frac{1}{T-t}-\frac{1}{T-\underline{t}}\right) \\
&\qquad + \frac{1}{T-\underline{t}}\left(\mathbb{E}\left[g(x+B_T-B_t)(B_T-B_t)\right] - \mathbb{E}\left[g(x+B_T^n-B_t^n)(B_T^n-B_t^n)\right]\right). \qquad (42)
\end{aligned}$$

For the first term, since

$$\mathbb{E}\left[g(x+B_T-B_t)(B_T-B_t)\right] = \mathbb{E}\left[\left(g(x+B_T-B_t)-g(x)\right)(B_T-B_t)\right],$$

we have, using the fact that g is ε-Hölder continuous,

$$\begin{aligned}
\left|\mathbb{E}\left[g(x+B_T-B_t)(B_T-B_t)\right]\left(\frac{1}{T-t}-\frac{1}{T-\underline{t}}\right)\right| &\leq \|g\|_\varepsilon\,\mathbb{E}\left[|B_T-B_t|^{1+\varepsilon}\right]\frac{t-\underline{t}}{(T-t)(T-\underline{t})} \\
&\leq \|g\|_\varepsilon\,\frac{t-\underline{t}}{(T-t)^{\frac{1-\varepsilon}{2}}(T-\underline{t})}.
\end{aligned}$$

But, exploiting that $(t-\underline{t})^{1-\frac{\varepsilon}{2}} \leq (T-\underline{t})^{1-\frac{\varepsilon}{2}}$ and $\frac{1}{(T-\underline{t})^{\frac{\varepsilon}{2}}} \leq \frac{1}{(T-t)^{\frac{\varepsilon}{2}}}$, we obtain

$$\frac{t-\underline{t}}{(T-t)^{\frac{1-\varepsilon}{2}}(T-\underline{t})} = \frac{(t-\underline{t})^{\frac{\varepsilon}{2}}(t-\underline{t})^{1-\frac{\varepsilon}{2}}}{(T-t)^{\frac{1-\varepsilon}{2}}(T-\underline{t})^{1-\frac{\varepsilon}{2}}(T-\underline{t})^{\frac{\varepsilon}{2}}} \leq \frac{(t-\underline{t})^{\frac{\varepsilon}{2}}}{\sqrt{T-t}} \leq \frac{h^{\frac{\varepsilon}{2}}}{\sqrt{T-t}},$$

from which we deduce that

$$\left|\mathbb{E}\left[g(x+B_T-B_t)(B_T-B_t)\right]\left(\frac{1}{T-t}-\frac{1}{T-\underline{t}}\right)\right| \leq T^{\frac{\varepsilon}{2}}\,\|g\|_\varepsilon\frac{n^{-\frac{\varepsilon}{2}}}{\sqrt{T-t}}.$$

Since $\delta(t,T) = T-\underline{t}$, from Corollary 4, the absolute value of the second term on the RHS of (42) is bounded by

$$C\,\|g\|_\varepsilon\,n^{-\frac{\varepsilon}{2}}\,\frac{(T-\underline{t})^{1/2}}{T-\underline{t}} = C\,\|g\|_\varepsilon\,n^{-\frac{\varepsilon}{2}}\,\frac{1}{\sqrt{T-\underline{t}}} \leq C\,\|g\|_\varepsilon\,n^{-\frac{\varepsilon}{2}}\,\frac{1}{\sqrt{T-t}}.$$

Then we get

$$\left|\mathbb{E}\left[g\left(B_T^{t,x}\right)\frac{B_T^{t,x}-x}{T-t}\right] - \mathbb{E}\left[g\left(B_T^{n,t,x}\right)\frac{B_T^{n,t,x}-x}{T-\underline{t}}\right]\right| \leq C\,\|g\|_\varepsilon\,\frac{n^{-\frac{\varepsilon}{2}}}{\sqrt{T-t}}. \qquad (43)$$

Study of the f difference. Here we have to estimate for $t \in [0, T[$,

$$H(t) = \left| \int_t^T \mathbb{E}\left[F\left(s, B_s^{t,x}\right) \frac{B_s^{t,x} - x}{s-t} \right] ds - \int_{\underline{t}+h}^T \mathbb{E}\left[f\left(\overline{s}, \Theta_s^{n,t,x}\right) \frac{B_s^{n,t,x} - x}{\underline{s} - \underline{t}} \right] ds \right|.$$

When $T - h \le t < s < T$, we observe that $B_s^{n,t,x} - x = 0$, and combining the regularity (31) of F with the estimate (30) we obtain

$$H(t) \le \int_t^T \mathbb{E}\left| F\left(s, B_s^{t,x}\right) \frac{B_s^{t,x} - x}{s-t} \right| ds \le C \int_t^T \frac{ds}{(s-t)^{(1-\varepsilon)/2}\sqrt{T-s}}\, ds = C(T-t)^{\frac{\varepsilon}{2}}\, B\left(\tfrac{1+\varepsilon}{2}, \tfrac{1}{2}\right).$$

Thus, for $T - h \le t \le T$, $H(t) \le C\, n^{-\frac{\varepsilon}{2}}$.

Let us now consider the case where $0 \le t < T - h$ i.e. $\underline{t} + h \le T - h$. We first write

$$\int_t^T \mathbb{E}\left[F\left(s, B_s^{t,x}\right) \frac{B_s^{t,x} - x}{s-t} \right] ds = \int_{\underline{t}+h}^T \mathbb{E}\left[F\left(s, B_s^{t,x}\right) \frac{B_s^{t,x} - x}{s-t} \right] ds + \int_t^{\underline{t}+h} \mathbb{E}\left[F\left(s, B_s^{t,x}\right) \frac{B_s^{t,x} - x}{s-t} \right] ds.$$

For the second term of the RHS of this equality, we proceed as above and get

$$\left| \int_t^{\underline{t}+h} \mathbb{E}\left[F\left(s, B_s^{t,x}\right) \frac{B_s^{t,x} - x}{s-t} \right] ds \right| \le C \int_t^{\underline{t}+h} \frac{ds}{(s-t)^{(1-\varepsilon)/2}\sqrt{T-s}}\, ds.$$

But, since $\underline{t} + h \le T - h$, for $t < s < \underline{t} + h$, $\dfrac{1}{\sqrt{T-s}} \le \dfrac{1}{\sqrt{h}}$ and, since $0 < \underline{t} + h - t \le h$,

$$\int_t^{\underline{t}+h} \frac{ds}{(s-t)^{(1-\varepsilon)/2}\sqrt{T-s}}\, ds \le \frac{1}{\sqrt{h}} \int_t^{\underline{t}+h} \frac{ds}{(s-t)^{(1-\varepsilon)/2}} = \frac{2}{1+\varepsilon} \frac{(\underline{t}+h-t)^{(1+\varepsilon)/2}}{\sqrt{h}} \le \frac{2h^{\frac{\varepsilon}{2}}}{1+\varepsilon}.$$

Secondly, we split the term

$$\int_{\underline{t}+h}^T \mathbb{E}\left[f\left(\overline{s}, \Theta_s^{n,t,x}\right) \frac{B_s^{n,t,x} - x}{\underline{s} - \underline{t}} \right] ds$$

into two parts:

$$\int_{\underline{t}+h}^T \mathbb{E}\left[f\left(s, \Theta_s^{n,t,x}\right) \frac{B_s^{n,t,x} - x}{\underline{s} - \underline{t}} \right] ds = \int_{\underline{t}+h}^T \mathbb{E}\left[F_n\left(s, B_s^{n,t,x}\right) \frac{B_s^{n,t,x} - x}{\underline{s} - \underline{t}} \right] ds$$

and, the remaining term

$$R = \int_{\underline{t}+h}^T \mathbb{E}\left[\left(f\left(\overline{s}, \Theta_s^{n,t,x}\right) - f\left(s, \Theta_s^{n,t,x}\right) \right) \frac{B_s^{n,t,x} - x}{\underline{s} - \underline{t}} \right] ds.$$

But, due to the uniform regularity of f in time, we have, since $\underline{s} - \underline{t} = (\underline{s}+h) - (\underline{t}+h) \ge s - (\underline{t}+h)$,

$$\begin{aligned} |R| &\le \|f_t\|_\alpha \int_{\underline{t}+h}^T (\overline{s} - s)^\alpha \frac{\mathbb{E}\left[|B_s^{n,t,x} - x| \right]}{\underline{s} - \underline{t}}\, ds \le \|f_t\|_\alpha\, h^\alpha \int_{\underline{t}+h}^T \frac{ds}{\sqrt{\underline{s} - \underline{t}}}\, ds \\ &\le \|f_t\|_\alpha\, h^\alpha \int_{\underline{t}+h}^T \frac{ds}{\sqrt{s - (\underline{t}+h)}}\, ds \le C\, h^\alpha. \end{aligned}$$

Thus, for $0 \le t < T - h$,

$$H(t) \le \int_{\underline{t}+h}^{T} \left| \mathbb{E}\left[F\left(s, B_s^{t,x}\right) \frac{B_s^{t,x} - x}{s-t} \right] - \mathbb{E}\left[F_n\left(s, B_s^{n,t,x}\right) \frac{B_s^{n,t,x} - x}{\underline{s}-\underline{t}} \right] \right| ds + C\, n^{-(\alpha \wedge \frac{\varepsilon}{2})}.$$

We split the integrand of the first term on the RHS of the inequality into three parts,

$$\begin{aligned}
&\left| \mathbb{E}\left[F\left(s, B_s^{t,x}\right) \frac{B_s^{t,x} - x}{s-t} \right] - \mathbb{E}\left[F_n\left(s, B_s^{n,t,x}\right) \frac{B_s^{n,t,x} - x}{\underline{s}-\underline{t}} \right] \right| \\
&\quad \le \left| \mathbb{E}\left[F\left(s, B_s^{t,x}\right) \frac{B_s^{t,x} - x}{s-t} \right] - \mathbb{E}\left[F\left(s, B_s^{t,x}\right) \frac{B_s^{t,x} - x}{\underline{s}-\underline{t}} \right] \right| \\
&\qquad + \left| \mathbb{E}\left[F\left(s, B_s^{t,x}\right) \frac{B_s^{t,x} - x}{\underline{s}-\underline{t}} \right] - \mathbb{E}\left[F\left(s, B_s^{n,t,x}\right) \frac{B_s^{n,t,x} - x}{\underline{s}-\underline{t}} \right] \right| \\
&\qquad + \left| \mathbb{E}\left[F\left(s, B_s^{n,t,x}\right) \frac{B_s^{n,t,x} - x}{\underline{s}-\underline{t}} \right] - \mathbb{E}\left[F_n\left(s, B_s^{n,t,x}\right) \frac{B_s^{n,t,x} - x}{\underline{s}-\underline{t}} \right] \right| \\
&\quad =: H_1(s) + H_2(s) + H_3(s),
\end{aligned}$$

so that

$$H(t) \le \int_{\underline{t}+h}^{T} \left(H_1(s) + H_2(s) + H_3(s)\right) ds + C\, n^{-(\alpha \wedge \frac{\varepsilon}{2})}.$$

The term H_1. Since $B_s^{t,x} - x$ has mean zero,

$$H_1(s) \le \|F(s,\cdot)\|_\varepsilon \mathbb{E}(|B_s^{t,x} - x|^{1+\varepsilon}) \left| \frac{1}{s-t} - \frac{1}{\underline{s}-\underline{t}} \right|$$

and the regularity (31) of F gives

$$H_1(s) \le C\, \frac{(s-t)^{(1+\varepsilon)/2}}{\sqrt{T-s}}\, \frac{(t-\underline{t}) + (s-\underline{s})}{(s-t)(\underline{s}-\underline{t})} = \frac{C}{\sqrt{T-s}}\, \frac{(t-\underline{t}) + (s-\underline{s})}{(s-t)^{(1-\varepsilon)/2}(\underline{s}-\underline{t})}.$$

Since $(t-\underline{t}) = (t-\underline{t})^{\frac{\varepsilon}{2}}(t-\underline{t})^{1-\frac{\varepsilon}{2}} \le h^{\frac{\varepsilon}{2}}(\underline{s}-\underline{t})^{1-\frac{\varepsilon}{2}}$ and the same upper bound holds for $s-\underline{s}$ (since $s - \underline{s} \le h \le \underline{s} - \underline{t}$), we get

$$H_1(s) \le \frac{C}{\sqrt{T-s}}\, \frac{h^{\frac{\varepsilon}{2}}}{(s-t)^{(1-\varepsilon)/2}(\underline{s}-\underline{t})^{\frac{\varepsilon}{2}}}.$$

Finally, we remark that $\underline{s} - \underline{t} = (\underline{s}+h) - (\underline{t}+h) \ge s - (\underline{t}+h)$ and $s - t \ge s - (\underline{t}+h)$, to obtain

$$H_1(s) \le \frac{C}{\sqrt{T-s}}\, \frac{h^{\frac{\varepsilon}{2}}}{(s-(\underline{t}+h))^{(1-\varepsilon)/2}(s-(\underline{t}+h))^{\frac{\varepsilon}{2}}} \le \frac{C\, h^{\frac{\varepsilon}{2}}}{\sqrt{T-s}\sqrt{s-(\underline{t}+h)}},$$

and, as a consequence,

$$\int_{\underline{t}+h}^{T} H_1(s)\, ds \le C\, h^{\frac{\varepsilon}{2}}\, B(\tfrac{1}{2}, \tfrac{1}{2}).$$

The term H_2. We use once again the regularity (31) of F together with Corollary 4 to obtain, for $\underline{t}+h<s<T$,

$$H_2(s) \le C\, n^{-\frac{\varepsilon}{2}}\, \delta(t,s)^{1/2}\, \frac{1}{\sqrt{T-s}}\, \frac{1}{\underline{s}-\underline{t}}.$$

We first use the fact that $\delta(s,t)=\max(s-t,\underline{s}-\underline{t}) \le s-\underline{t} \le s-\underline{s}+\underline{s}-\underline{t} \le 2(\underline{s}-\underline{t})$ to get

$$H_2(s) \le C\, \frac{n^{-\frac{\varepsilon}{2}}}{\sqrt{T-s}\sqrt{\underline{s}-\underline{t}}},$$

and, since $\underline{s}-\underline{t} \ge s-(\underline{t}+h)$, we have

$$H_2(s) \le C\, \frac{n^{-\frac{\varepsilon}{2}}}{\sqrt{T-s}\sqrt{s-(\underline{t}+h)}},$$

from which we deduce the estimate

$$\int_{\underline{t}+h}^{T} H_2(s)\, ds \le C\, n^{-\frac{\varepsilon}{2}}\, B(\tfrac{1}{2},\tfrac{1}{2}).$$

The term H_3. For this last term, we come back to the definitions (5) and (37) of F and F_n, respectively. By (40) we have, for $\underline{t}+h<s<T$,

$$\begin{aligned} H_3(s) &\le \mathbb{E}\left[|F-F_n|(s, B_s^{n,t,x}) \left| \frac{B_s^{n,t,x}-x}{\underline{s}-\underline{t}} \right| \right] \\ &\le \|f\|_{\mathrm{Lip}}\, (\beta_n(s)+\gamma_n(s))\, \mathbb{E}\left[\left(1+|B_s^{n,t,x}|\right)^{\varepsilon} \left| \frac{B_s^{n,t,x}-x}{\underline{s}-\underline{t}} \right| \right], \end{aligned}$$

and, by the Cauchy-Schwarz inequality, we derive the estimate

$$H_3(s) \le C\, (\beta_n(s)+\gamma_n(s))\, (1+|x|)^{\varepsilon}\, \frac{1}{\sqrt{\underline{s}-\underline{t}}}.$$

Summary for H. Let us summarize the estimates we got for H. For $T-h \le t \le T$, we have $H(t) \le C\, n^{-\frac{\varepsilon}{2}}$. For $0 \le t < T-h$ we obtained the upper bound

$$H(t) \le C\, n^{-(\alpha\wedge\frac{\varepsilon}{2})} + C\, (1+|x|)^{\varepsilon} \int_{\underline{t}+h}^{T} (\beta_n(s)+\gamma_n(s))\, \frac{ds}{\sqrt{\underline{s}-\underline{t}}}.$$

Hence we have, for $t \in [0,T]$,

$$H(t) \le C\, (1+|x|)^{\varepsilon} \left(n^{-(\alpha\wedge\frac{\varepsilon}{2})} + \int_{(\underline{t}+h)\wedge T}^{T} (\beta_n(s)+\gamma_n(s))\, \frac{ds}{\sqrt{\underline{s}-\underline{t}}} \right).$$

Coming back to (43), we have, for any $x \in \mathbb{R}$ and $t \in [0,T[$,

$$|\nabla u(t,x) - \Delta^n(t,x)| \le C\, (1+|x|)^{\varepsilon} \left(\frac{n^{-(\alpha\wedge\frac{\varepsilon}{2})}}{\sqrt{T-t}} + \int_{(\underline{t}+h)\wedge T}^{T} (\beta_n(s)+\gamma_n(s))\, \frac{ds}{\sqrt{\underline{s}-\underline{t}}} \right),$$

and, as a byproduct,

$$\gamma_n(t) \le C \left(\frac{n^{-(\alpha\wedge\frac{\varepsilon}{2})}}{\sqrt{T-t}} + \int_{(\underline{t}+h)\wedge T}^{T} (\beta_n(s)+\gamma_n(s))\, \frac{ds}{\sqrt{\underline{s}-\underline{t}}} \right), \quad t \in [0,T[. \tag{44}$$

Global estimate. Plugging (41) into (44), we get, for $t \in [0, T[$,

$$\gamma_n(t) \le C \left(\frac{n^{-(\alpha \wedge \frac{\varepsilon}{2})}}{\sqrt{T-t}} + \int_{(\underline{t}+h)\wedge T}^{T} \left(n^{-(\alpha \wedge \frac{\varepsilon}{2})} + \int_s^T \gamma_n(r)\, dr + \gamma_n(s) \right) \frac{ds}{\sqrt{\underline{s}-\underline{t}}} \right).$$

For $t < T-h$, since $\underline{s}-\underline{t} = (\underline{s}+h) - (\underline{t}+h) \ge s - (\underline{t}+h)$, we have

$$n^{-(\alpha \wedge \frac{\varepsilon}{2})} \int_{\underline{t}+h}^{T} \frac{ds}{\sqrt{\underline{s}-\underline{t}}} \le n^{-(\alpha \wedge \frac{\varepsilon}{2})} \int_{\underline{t}+h}^{T} \frac{ds}{\sqrt{s-(\underline{t}+h)}} \le 2\, n^{-(\alpha \wedge \frac{\varepsilon}{2})} \sqrt{T-(\underline{t}+h)} \le 2\sqrt{T}\, n^{-(\alpha \wedge \frac{\varepsilon}{2})}.$$

Again, we have,

$$\int_{\underline{t}+h}^{T} \int_s^T \gamma_n(r)\, dr \frac{ds}{\sqrt{\underline{s}-\underline{t}}} \le \int_{\underline{t}+h}^{T} \int_s^T \gamma_n(r)\, dr \frac{ds}{\sqrt{s-(\underline{t}+h)}},$$

from which we deduce

$$\begin{aligned} \int_{\underline{t}+h}^{T} \int_s^T \gamma_n(r)\, dr \frac{ds}{\sqrt{s-(\underline{t}+h)}} &= \int_{\underline{t}+h}^{T} \gamma_n(r) \int_{\underline{t}+h}^{T} \mathbf{1}_{r>s} \frac{ds}{\sqrt{s-(\underline{t}+h)}} dr \\ &= 2 \int_{\underline{t}+h}^{T} \gamma_n(r) \sqrt{r-(\underline{t}+h)}\, dr \\ &\le 2T \int_{\underline{t}+h}^{T} \gamma_n(s) \frac{ds}{\sqrt{\underline{s}-\underline{t}}}. \end{aligned}$$

It follows that, for $t \in [0, T[$,

$$\gamma_n(t) \le C \left(\frac{n^{-(\alpha \wedge \frac{\varepsilon}{2})}}{\sqrt{T-t}} + \int_{(\underline{t}+h)\wedge T}^{T} \gamma_n(s) \frac{ds}{\sqrt{\underline{s}-\underline{t}}} \right).$$

Thus,

$$\begin{aligned} \gamma_n(t) \le\ & C \frac{n^{-(\alpha \wedge \frac{\varepsilon}{2})}}{\sqrt{T-t}} \\ & + C^2 \left(\int_{(\underline{t}+h)\wedge T}^{T} \left(\frac{n^{-(\alpha \wedge \frac{\varepsilon}{2})}}{\sqrt{T-s}} + \int_{(\underline{s}+h)\wedge T}^{T} \gamma_n(r) \frac{dr}{\sqrt{\underline{r}-\underline{s}}} \right) \frac{ds}{\sqrt{\underline{s}-\underline{t}}} \right). \end{aligned}$$

But we have

$$\int_{(\underline{t}+h)\wedge T}^{T} \frac{n^{-(\alpha \wedge \frac{\varepsilon}{2})}}{\sqrt{T-s}} \frac{ds}{\sqrt{\underline{s}-\underline{t}}} \le \int_{(\underline{t}+h)\wedge T}^{T} \frac{n^{-(\alpha \wedge \frac{\varepsilon}{2})}}{\sqrt{T-s}} \frac{ds}{\sqrt{s-(\underline{t}+h)}} = n^{-(\alpha \wedge \frac{\varepsilon}{2})}\, B(\tfrac{1}{2}, \tfrac{1}{2}),$$

and, for $t < T-h$, since $\underline{s}+h \le r$ if and only if $s < \underline{r}$,

$$\begin{aligned} \int_{(\underline{t}+h)\wedge T}^{T} \int_{(\underline{s}+h)\wedge T}^{T} \gamma_n(r) \frac{dr}{\sqrt{\underline{r}-\underline{s}}} \frac{ds}{\sqrt{\underline{s}-\underline{t}}} &= \int_{(\underline{t}+2h)\wedge T}^{T} \gamma_n(r) \int_{(\underline{t}+h)\wedge T}^{T} \mathbf{1}_{s<\underline{r}} \frac{ds}{\sqrt{\underline{r}-\underline{s}}\sqrt{\underline{s}-\underline{t}}}\, dr \\ &= \int_{(\underline{t}+2h)\wedge T}^{T} \gamma_n(r) \int_{(\underline{t}+h)\wedge T}^{\underline{r}} \frac{ds}{\sqrt{\underline{r}-\underline{s}}\sqrt{\underline{s}-\underline{t}}}\, dr. \end{aligned}$$

But $\underline{r}-\underline{s} \ge \underline{r}-s$ and $\underline{s}-\underline{t} \ge s-(\underline{t}+h)$, so we get

$$\begin{aligned} \int_{(\underline{t}+h)\wedge T}^{T} \int_{(\underline{s}+h)\wedge T}^{T} \gamma_n(r) \frac{dr}{\sqrt{\underline{r}-\underline{s}}} \frac{ds}{\sqrt{\underline{s}-\underline{t}}} &\le \int_{(\underline{t}+2h)\wedge T}^{T} \gamma_n(r) \int_{(\underline{t}+h)\wedge T}^{\underline{r}} \frac{ds}{\sqrt{\underline{r}-s}\sqrt{s-(\underline{t}+h)}}\, dr \\ &= B(\tfrac{1}{2}, \tfrac{1}{2}) \int_{(\underline{t}+2h)\wedge T}^{T} \gamma_n(r)\, dr. \end{aligned}$$

Finally, we have

$$\gamma_n(t) \le C\left(\frac{n^{-(\alpha\wedge\frac{\varepsilon}{2})}}{\sqrt{T-t}} + \int_t^T \gamma_n(s)\,ds\right) \quad \text{for} \quad t \in [0,T[,$$

and from Gronwall's inequality (Lemma 13)

$$\gamma_n(t) \le C\,\frac{n^{-(\alpha\wedge\frac{\varepsilon}{2})}}{\sqrt{T-t}} \quad \text{for} \quad t \in [0,T[.$$

Coming back to (41), we have also,

$$\beta_n(t) \le C\,n^{-(\alpha\wedge\frac{\varepsilon}{2})} \quad \text{for} \quad t \in [0,T].$$

The proof of Proposition 11 is complete. □

Proof of Theorem 10. Theorem 10 is mainly a corollary of Proposition 11.

Let us begin with the convergence of the $(Y^n)_n$ processes. Let us fix $r \ge 1$, $x \in \mathbb{R}$ and $0 \le t \le s \le T$. We have

$$\begin{aligned} W_r\left(Y_s^{t,x}, Y_s^{n,t,x}\right) &= W_r\left(u(s,B_s^{t,x}), U^n(s,B_s^{n,t,x})\right) \\ &\le W_r\left(u(s,B_s^{t,x}), u(s,B_s^{n,t,x})\right) + W_r\left(u(s,B_s^{n,t,x}), U^n(s,B_s^{n,t,x})\right). \end{aligned}$$

Since, by Lemma 5, u is ε-Hölder continuous in space, uniformly in time, we have, by Hölder's inequality,

$$W_r\left(u(s,B_s^{t,x}), u(s,B_s^{n,t,x})\right) \le C\,W_r\left(B_s^{t,x}, B_s^{n,t,x}\right)^{\varepsilon} \le C c_r^{\varepsilon} T^{\frac{\varepsilon}{2}}\, n^{-\frac{\varepsilon}{2}},$$

where we have used Proposition 3 (see (22)) for the last inequality. Moreover, by Proposition 11,

$$W_r\left(u(s,B_s^{n,t,x}), U^n(s,B_s^{n,t,x})\right) \le \left\|u(s,B_s^{n,t,x}) - U^n(s,B_s^{n,t,x})\right\|_{\mathrm{L}^r} \le C\,(1+|x|)^{\varepsilon}\, n^{-(\alpha\wedge\frac{\varepsilon}{2})}.$$

This gives the first part of the result.

Let us continue with the convergence of the $(Z^n)_n$ processes. The proof is almost the same except for the grid points. Let $0 \le t < s \le T$ with $s \ne \underline{s}$ i.e. $s \notin \{kh, k = n_t+1, \dots, n\}$. We have as before

$$\begin{aligned} &W_r\left(\nabla u(s,B_s^{t,x}), \Delta^n(s,B_s^{n,t,x})\right) \\ &\qquad \le W_r\left(\nabla u(s,B_s^{t,x}), \nabla u(s,B_s^{n,t,x})\right) + W_r\left(\nabla u(s,B_s^{n,t,x}), \Delta^n(s,B_s^{n,t,x})\right). \end{aligned}$$

Since, by Lemma 6, $\nabla u(s,\cdot)$ is ε-Hölder continuous, we have, by Hölder's inequality,

$$W_r\left(\nabla u(s,B_s^{t,x}), \nabla u(s,B_s^{n,t,x})\right) \le \frac{C}{\sqrt{T-s}}\,W_r\left(B_s^{t,x}, B_s^{n,t,x}\right)^{\varepsilon} \le \frac{C c_r^{\varepsilon} T^{\frac{\varepsilon}{2}}}{\sqrt{T-s}}\, n^{-\frac{\varepsilon}{2}},$$

where the last inequality follows from (22). Moreover, by Proposition 11,

$$W_r\left(\nabla u(s,B_s^{n,t,x}), \Delta^n(s,B_s^{n,t,x})\right) \le \left\|\nabla u(s,B_s^{n,t,x}) - \Delta^n(s,B_s^{n,t,x})\right\|_{\mathrm{L}^r} \le C\,\frac{(1+|x|)^{\varepsilon}}{\sqrt{T-s}} n^{-(\alpha\wedge\frac{\varepsilon}{2})},$$

and the result follows, in this case, from equalities (25) and (16) together with Remark 7.

Let us now consider the case where $s \in \{kh, k = n_t + 1, \dots, n-1\}$. In this case we have

$$Z_s^{n,t,x} = \nabla^n U^n \left(s, B_{s-}^{n,t,x}\right) = \Delta^n \left(s - h/2, B_{s-h/2}^{n,t,x}\right)$$

which is not equal to $\Delta^n \left(s, B_s^{n,t,x}\right)$ in general. We first write

$$W_r \left(Z_s^{t,x}, Z_s^{n,t,x}\right) = W_r \left(Z_s^{t,x}, Z_{s-h/2}^{n,t,x}\right) \leq W_r \left(Z_s^{t,x}, Z_{s-h/2}^{t,x}\right) + W_r \left(Z_{s-h/2}^{t,x}, Z_{s-h/2}^{n,t,x}\right).$$

The second term on the RHS can be bounded by using our previous result, namely

$$W_r \left(Z_{s-h/2}^{t,x}, Z_{s-h/2}^{n,t,x}\right) \leq C \frac{n^{-(\alpha \wedge \frac{\varepsilon}{2})} \left(1 + |x|\right)^{\varepsilon}}{\sqrt{T - (s - h/2)}} \leq C \frac{n^{-(\alpha \wedge \frac{\varepsilon}{2})} \left(1 + |x|\right)^{\varepsilon}}{\sqrt{T - s}}.$$

For the first term, one can write

$$\begin{aligned} W_r \left(Z_s^{t,x}, Z_{s-h/2}^{t,x}\right) &= W_r \left(\nabla u(s, B_s^{t,x}), \nabla u(s - h/2, B_{s-h/2}^{t,x})\right) \\ &\leq W_r \left(\nabla u(s - h/2, B_{s-h/2}^{t,x}), \nabla u(s - h/2, B_s^{t,x})\right) + W_r \left(\nabla u(s - h/2, B_s^{t,x}), \nabla u(s, B_s^{t,x})\right). \end{aligned}$$

From Lemma 6, $\nabla u(s - h/2, \cdot)$ is ε-Hölder continuous and we have

$$\begin{aligned} W_r \left(\nabla u(s - h/2, B_{s-h/2}^{t,x}), \nabla u(s - h/2, B_s^{t,x})\right) &\leq \frac{C}{\sqrt{T - (s - h/2)}} W_r \left(B_{s-h/2}^{t,x}, B_s^{t,x}\right)^{\varepsilon} \\ &\leq \frac{C \|G\|_{L^r}^{\varepsilon}}{\sqrt{T - s}} \left(\frac{h}{2}\right)^{\frac{\varepsilon}{2}}, \end{aligned}$$

where G is a standard normal random variable. Finally, for the last term, we use Proposition 9 for the time regularity of ∇u. We have

$$W_r \left(\nabla u(s - h/2, B_s^{t,x}), \nabla u(s, B_s^{t,x})\right) \leq \left\| \nabla u(s - h/2, B_s^{t,x}) - \nabla u(s, B_s^{t,x}) \right\|_{\mathrm{L}^r} \leq \frac{C\, h^{\frac{\varepsilon}{2}}}{\sqrt{T - s}},$$

and the estimate for $W_r \left(Z_s^{n,t,x}, Z_s^{t,x}\right)$ follows.

This ends the proof. □

A. Appendix.

A.1. A priori estimate for discrete BSDEs. For the convenience of the reader, we prove a generalization of an a priori estimate for BSDEs driven by random walks given in [6, Proposition 7] (see also the appendix in [5]). This generalization allows to consider two different generators.

Lemma 12. *There exists an integer $n_0 \in \mathbb{N}^*$ and a constant $C > 0$ both depending only on T and $\|f\|_{\mathrm{Lip}}$ such that for any couple of functions $(\bar{g}, \bar{f})$ satisfying (A1) and for all $n \geq n_0$ with $n > T\|\bar{f}\|_{\mathrm{Lip}}$ and all $(t, x, \bar{x}) \in [0, T] \times \mathbb{R}^2$,*

$$\begin{aligned} &\mathbb{E}\left[\sup_{t \leq s \leq T} \left|Y_s^{n,t,x} - \bar{Y}_s^{n,t,\bar{x}}\right|^2 + \int_{]t,T]} \left|Z_r^{n,t,x} - \bar{Z}_r^{n,t,\bar{x}}\right|^2 d\langle B^n \rangle_r\right] \\ &\qquad \leq C\, \mathbb{E}\left[\left|g\left(B_T^{n,t,x}\right) - \bar{g}\left(B_T^{n,t,\bar{x}}\right)\right|^2 + \int_{]t,T]} \left|\delta f\left(r, x, \bar{x}, \bar{Y}_{r^-}^{n,t,\bar{x}}, \bar{Z}_r^{n,t,\bar{x}}\right)\right|^2 d\langle B^n \rangle_r\right], \end{aligned}$$

where $\delta f(r, x, \bar{x}, y, z) = f\left(r, B_{r^-}^{n,t,x}, y, z\right) - \bar{f}\left(r, B_{r^-}^{n,t,\bar{x}}, y, z\right)$ and, for all (t, x), $\left(\bar{Y}^{n,t,x}, \bar{Z}^{n,t,x}\right)$ denotes the solution to (12) where (g, f) is replaced by $(\bar{g}, \bar{f})$.

Proof. Let n be such that $T\|f\|_{\mathrm{Lip}}/n < 1$ and $T\|\bar{f}\|_{\mathrm{Lip}}/n < 1$. Since, $\langle B^n\rangle_t - \langle B^n\rangle_s \leq (t-s) + T/n$, doing exactly the same computation as in the proof of Proposition 7 in [6], we get, for a universal constant $c \geq 1$,

$$\begin{aligned}\mathbb{E}\left[\sup_{\sigma\leq s\leq\tau}|\delta Y_s^n|^2 + \int_{]\sigma,\tau]}|\delta Z_r^n|^2 d\langle B^n\rangle_r\right] \\ \leq c\,\mathbb{E}\left[|\delta Y_\tau^n|^2\right] + C(\tau-\sigma,n)\,\mathbb{E}\left[\int_{]\sigma,\tau]}\left|\delta f\left(r,x,\bar{x},\bar{Y}_{r^-}^{n,t,\bar{x}},\bar{Z}_r^{n,t,\bar{x}}\right)\right|^2 d\langle B^n\rangle_r\right] \\ + C(\tau-\sigma,n)\,\mathbb{E}\left[\sup_{\sigma\leq s\leq\tau}|\delta Y_s^n|^2 + \int_{]\sigma,\tau]}|\delta Z_r^n|^2\, d\langle B^n\rangle_r\right]\end{aligned}$$

for all deterministic $0 \leq \sigma < \tau \leq T$, where

$$C(\alpha,n) = c\,\max\left(\|f\|_{\mathrm{Lip}}^2, 1\right)\,\max\left((\alpha+T/n)^2, (\alpha+T/n)\right) \qquad \text{for } \alpha \geq 0.$$

We choose an integer $m \in \mathbb{N}^*$ such that, with $\alpha = T/m$, $C(\alpha,+\infty) \leq 1/3$. Then, there exists an $n_0 \in \mathbb{N}^*$ such that, for $n \geq n_0$ it holds $C(\alpha,n) \leq 1/2$ and $T\|f\|_{\mathrm{Lip}}/n < 1$. As soon as $\tau - \sigma \leq \alpha$ and $n \geq n_0$, we have

$$\begin{aligned}\mathbb{E}\left[\sup_{\sigma\leq s\leq\tau}|\delta Y_s^n|^2 + \int_{]\sigma,\tau]}|\delta Z_r^n|^2 d\langle B^n\rangle_r\right] \\ \leq 2c\,\mathbb{E}\left[|\delta Y_\tau^n|^2 + \int_{]\sigma,\tau]}\left|\delta f\left(r,x,\bar{x},\bar{Y}_{r^-}^{n,t,\bar{x}},\bar{Z}_r^{n,t,\bar{x}}\right)\right|^2 d\langle B^n\rangle_r\right]. \quad (45)\end{aligned}$$

We set, for $0 \leq k \leq m-1$, $I_k = \underline{t} +]k(T-\underline{t})/m, (k+1)(T-\underline{t})/m]$ and we introduce the following norm on $\mathcal{S}^2 \times \mathcal{M}^2$:

$$\|(Y^n,Z^n)\|_s^2 = \sum_{k=0}^{m-1}(2\times 2c)^k\,\mathbb{E}\left[\sup_{t\in I_k}|Y_t^n|^2 + \int_{I_k}|Z_r^n|^2 d\langle B^n\rangle_r\right].$$

Considering $I_k =]\sigma,\tau]$ and summing up (45) over k yields

$$\begin{aligned}\|(\delta Y^n,\delta Z^n)\|_s^2 \leq \frac{1}{2}\,\|(\delta Y^n,\delta Z^n)\|_s^2 + (4c)^{m-1}(2c)\,\mathbb{E}\left[\left|g\left(B_T^{n,t,x}\right) - \bar{g}\left(B_T^{n,t,\bar{x}}\right)\right|^2\right] \\ + (4c)^{m-1}\,(2c)\,\mathbb{E}\left[\int_{]\underline{t},T]}\left|\delta f\left(r,x,\bar{x},\bar{Y}_{r^-}^{n,\underline{t},x},\bar{Z}_r^{n,\underline{t},x}\right)\right|^2 d\langle B^n\rangle_r\right],\end{aligned}$$

from which we get

$$\|(\delta Y^n,\delta Z^n)\|_s^2 \leq (4c)^m\,\mathbb{E}\left[\left|g\left(B_T^{n,t,x}\right) - \bar{g}\left(B_T^{n,t,\bar{x}}\right)\right|^2 + \int_{]\underline{t},T]}\left|\delta f\left(r,x,\bar{x},\bar{Y}_{r^-}^{n,t,\bar{x}},\bar{Z}_r^{n,t,\bar{x}}\right)\right|^2 d\langle B^n\rangle_r\right].$$

This finishes the proof since $\|(\delta Y^n,\delta Z^n)\|_s^2$ upper bounds the LHS of the inequality stated in the Lemma. □

A.2. Gronwall lemmas. We recall the Gronwall lemmas used in this article.

Lemma 13. *Suppose that $g(t), \alpha(t) : [0,T[\to [0,\infty[$ are integrable functions, and $\beta > 0$. For $0 \le t < T$, if*

$$g(t) \le \alpha(t) + \beta \int_t^T g(s)ds,$$

then

$$g(t) \le \alpha(t) + \beta \int_t^T \alpha(s) e^{\beta(s-t)} ds.$$

The second lemma is of Volterra type. It can be either proved directly by a convolution argument or one can use [12, Exercise 4, page 190].

Lemma 14. *Assume a measurable $g : [0,T[\to [0,\infty[\in L_1([0,T[)$ and $\alpha, \beta > 0$ such that*

$$g(t) \le \frac{\alpha}{\sqrt{T-t}} + \beta \int_t^T \frac{g(s)}{\sqrt{s-t}} ds$$

for all $t \in [0,T[$. Then $g(t) \le C \frac{\alpha}{\sqrt{T-t}}$ for $t \in [0,T[$ for a constant $C = C(T,\beta) > 0$.

A.3. Proof of Lemma 6: Step 2. We assume that $0 < \varepsilon < 1$; the case $\varepsilon = 1$ was treated in Step 1. For $\eta > 0$, let us consider the function

$$f^\eta(t,x,y,z) = \inf\{f(t,p,y,z) + \eta|x-p| : p \in \mathbb{R}\} = \inf\{f(t,x-q,y,z) + \eta|q| : q \in \mathbb{R}\}.$$

When f satisfies (3), then f^η does as well and it is η-Lipschitz continuous w.r.t. x. Moreover

$$|f^\eta(t,x,y,z) - f(t,x,y,z)| \le c(\|f_x\|_\varepsilon)\, \eta^{-\varepsilon/(1-\varepsilon)}, \tag{46}$$

with $c(t) = (\varepsilon t)^{\varepsilon/(1-\varepsilon)} t(1-\varepsilon)$. Indeed,

$$\begin{aligned} f(t,x,y,z) \ge f^\eta(t,x,y,z) &= f(t,x,y,z) + \inf\{f(t,x-q,y,z) - f(t,x,y,z) + \eta|q| : q \in \mathbb{R}\} \\ &\ge f(t,x,y,z) + \inf\{-\|f_x\|_\varepsilon |q|^\varepsilon + \eta|q| : q \in \mathbb{R}\} \\ &= f(t,x,y,z) - c(\|f_x\|_\varepsilon)\, \eta^{-\varepsilon/(1-\varepsilon)}. \end{aligned}$$

In particular, $|K_f - K_{f^\eta}| \le c(\|f_x\|_\varepsilon)\, \eta^{-\varepsilon/(1-\varepsilon)}$. Let $(Y^{t,x,\eta}, Z^{t,x,\eta})$ be the solution to the BSDE (4) with data (g, f^η) and u^η be the function $u^\eta(t,x) := Y_t^{t,x,\eta}$. By the usual classical estimate for BSDEs (see, for instance, [8, Proposition 2.1 and remarks] or [11, Lemma 5.26]), there exists a $C > 0$ such that, for all $(t,x) \in [0,T] \times \mathbb{R}$,

$$\begin{aligned} \mathbb{E}\left[\sup_{s\in[t,T]} |Y_s^{t,x,\eta} - Y_s^{t,x}|^2 + \int_t^T |Z_s^{t,x,\eta} - Z_s^{t,x}|^2 ds \right] &\le C\mathbb{E} \int_t^T |(f^\eta - f)(s, B_s^{t,x}, Y_s^{t,x}, Z_s^{t,x})|^2 ds \\ &\le C\,T\, c(\|f_x\|_\varepsilon)^2\, \eta^{-2\varepsilon/(1-\varepsilon)}. \end{aligned} \tag{47}$$

In particular, $(u^\eta)_\eta$ converges to u, as $\eta \to +\infty$, uniformly on $[0,T] \times \mathbb{R}$.

Proof of (a), (bi), and (bii). Since f^η is Lipschitz continuous w.r.t. (x,y,z) and satisfies (3) (uniformly in η), by Step 1, we know that $u^\eta \in \mathcal{C}^{0,1}([0,T[\times\mathbb{R})$ and $Z_s^{t,x,\eta} = \nabla u^\eta(s, B_s^{t,x})$ for a.e. $(s,\omega) \in [t,T[\times\Omega$ with

$$|\nabla u^\eta(s,x)| \le \frac{C}{(T-s)^{(1-\varepsilon)/2}} \quad \text{for} \quad (s,x) \in [0,T[\times\mathbb{R}. \tag{48}$$

Taking into account the convergence of $(Z^{t,x,\eta})_\eta$ to $Z^{t,x}$ in $\mathrm{L}^2([t,T[\times\Omega)$, we have

$$|Z_s^{t,x}| \le \frac{C}{(T-s)^{(1-\varepsilon)/2}}, \quad \text{for a.e. } (s,\omega) \in [t,T[\times\Omega. \tag{49}$$

We define the function

$$v(t,x) := \mathbb{E}\left(g(B_T^{t,x})\frac{B_T-B_t}{T-t} + \int_t^T f(s,B_s^{t,x},Y_s^{t,x},Z_s^{t,x})\frac{B_s-B_t}{s-t}ds\right) \quad \text{for } (t,x)\in[0,T[\times\mathbb{R}.$$

First we show that

$$\nabla u^\eta(t,x) \to v(t,x) \quad \text{for} \quad (t,x)\in[0,T[\times\mathbb{R}$$

which also implies that $v:[0,T[\times\mathbb{R}\to\mathbb{R}$ is measurable. For this we denote

$$\widehat{Z}_r^{t,x} := \mathbb{E}_r\left(g(B_T^{t,x})\frac{B_T-B_r}{T-r} + \int_r^T f(s,B_s^{t,x},Y_s^{t,x},Z_s^{t,x})\frac{B_s-B_r}{s-r}ds\right) \quad \text{for } r\in[t,T[.$$

We also use (26) for (g,f^η) so that

$$\nabla u^\eta(r,B_r^{t,x}) = \mathbb{E}_r\left(g(B_T^{t,x})\frac{B_T-B_r}{T-r} + \int_r^T f^\eta(s,B_s^{t,x},Y_s^{t,x,\eta},Z_s^{t,x,\eta})\frac{B_s-B_r}{s-r}ds\right).$$

Then we apply the conditional Cauchy-Schwarz inequality to $\frac{(f^\eta-f)}{(s-r)^{\varepsilon/4}}\frac{B_s-B_r}{(s-r)^{1-\varepsilon/4}}$ and use (46) to get

$$\begin{aligned}
&\mathbb{E}|\nabla u^\eta(r,B_r^{t,x}) - \widehat{Z}_r^{t,x}|^2 \\
&\le \mathbb{E}\left|\mathbb{E}_r\left(\int_r^T (f^\eta(s,B_s^{t,x},Y_s^{t,x,\eta},Z_s^{t,x,\eta}) - f(s,B_s^{t,x},Y_s^{t,x},Z_s^{t,x}))\frac{B_s-B_r}{s-r}ds\right)\right|^2 \\
&\le \frac{2}{\varepsilon}T^{\frac{\varepsilon}{2}}\,\mathbb{E}\int_r^T \left(c(\|f_x\|_\varepsilon)\,\eta^{-\varepsilon/(1-\varepsilon)} + \|f_y\|_{\mathrm{Lip}}|Y_s^{t,x,\eta}-Y_s^{t,x}| + \|f_z\|_{\mathrm{Lip}}|Z_s^{t,x,\eta}-Z_s^{t,x}|\right)^2 \frac{1}{(s-r)^{\varepsilon/2}}ds.
\end{aligned}$$

Taking into account the bound for Y given in (47), we have

$$\mathbb{E}|\nabla u^\eta(r,B_r^{t,x}) - \widehat{Z}_r^{t,x}|^2 \le C\left(\eta^{-2\varepsilon/(1-\varepsilon)} + \mathbb{E}\int_r^T |Z_s^{t,x,\eta}-Z_s^{t,x}|^2\frac{1}{(s-r)^{\varepsilon/2}}\,ds\right), \tag{50}$$

where C depends on T, ε, $\|f_x\|_\varepsilon$ and $\|f\|_{\mathrm{Lip}}$.

Now we find a sequence $\eta_k\uparrow+\infty$ such that the RHS tends to 0. Indeed, this follows by dominated convergence as (47) guarantees a sequence $\eta_k\uparrow+\infty$ such that

$$Z_s^{t,x,\eta_k} \to Z_s^{t,x} \text{ a.e.} \quad \text{on} \quad [t,T[\times\Omega,$$

and because of the equations (48) and (49). For $r\ge t$, $\nabla u^\eta(r,B_r^{t,x})$ converges to $\widehat{Z}_r^{t,x}$ in L^2 and, in particular, for $r=t$ we obtain the desired convergence $\nabla u^\eta(t,x)\to v(t,x)$. Because of $Z_s^{t,x,\eta} = \nabla u^\eta(s,B_s^{t,x})$ for a.e. $(s,\omega)\in[t,T[\times\Omega$ this also gives that

$$Z_s^{t,x} = v(s,B_s^{t,x}) = \widehat{Z}_s^{t,x} \text{ a.e.} \quad \text{on} \quad [t,T[\times\Omega. \tag{51}$$

Coming back to (50), Gronwall's lemma (Lemma 14) gives,

$$\mathbb{E}|Z_r^{t,x,\eta} - Z_r^{t,x}|^2 \le C\eta^{-2\varepsilon/(1-\varepsilon)}.$$

Especially, $|\nabla u^\eta(t,x)-v(t,x)|\le C\eta^{-\varepsilon/(1-\varepsilon)}$, so that $\lim_{\eta\to\infty}\nabla u^\eta(t,x) = v(t,x)$.

Moreover, we conclude from this estimate and from the continuity of ∇u^η on $[0,T[\times\mathbb{R}$ that also v is continuous. Finally, $v(t,x)=\nabla u(t,x)$ follows from taking the limit $\eta\uparrow+\infty$ in

$$u^\eta(t,x) = u^\eta(t,0) + \int_0^x \nabla u^\eta(t,y)dy,$$

where we use dominated convergence based on the inequality (48).

References.

[1] L. Ambrosio, N. Gigli and G. Savaré, *Gradient Flows in Metric Spaces and in the Space of Probability Measures*, Birkhäuser, Basel, Boston, Berlin (2005).

[2] J.M. Bismut, *Théorie probabiliste du contrôle des diffusions*, Mem. AMS **176** (1973).

[3] B. Bouchard and N. Touzi, *Discrete-time approximation and Monte-Carlo simulation of backward stochastic differential equations*, Stochastic Process. Appl. **111** (2004), no. 2, 175–206.

[4] P. Briand, B. Delyon, Y. Hu, E. Pardoux and L. Stoica, *Lp solutions of backward stochastic differential equations*, Stochastic Process. Appl. **108** (2003) , 109–129.

[5] Ph. Briand, B. Delyon, and J. Mémin, *Donsker–type theorem for BSDEs*, Electron. Comm. Probab. **6** (2001), 1–14, (electronic).

[6] ———, *On the robustness of backward stochastic differential equations*, Stochastic Process. Appl. **97** (2002), no. 2, 229–253.

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

[8] N. El Karoui, S. Peng, and M.C. Quenez, *Backward Stochastic Differential Equations in Finance*, Math. Finance **7** (1997), 1–71.

[9] C. Geiss, C. Labart, and A. Luoto, *L^2-Approximation rate of forward -backward SDEs using random walk*, arXiv:1807.05889 (2018).

[10] ———, *Random walk approximation of BSDEs with Hölder continuous terminal condition*, arXiv:1806.07674, to appear in Bernoulli, (2019).

[11] S. Geiss and J. Ylinen, *Decoupling on the Wiener Space, Related Besov Spaces, and Applications to BSDEs.* arXiv:1409.5322, to appear Memoirs AMS, (2019).

[12] D. Henry, *Geometric theory of semilinear parabolic equations*, Lecture Notes in Mathematics 840, Springer, 1981.

[13] J. Ma and J. Zhang, *Representation theorems for backward stochastic differential equations*, Ann. Appl. Probab. **12** (2002), no. 4, 1390–1418.

[14] É. Pardoux and S. Peng, *Adapted solution of a backward stochastic differential equation*, Systems Control Lett. **14** (1990), no. 1, 55–61.

[15] E. Rio, *Upper bounds for minimal distances in the central limit theorem*, Ann. Inst. Henri Poincaré Probab. Stat. **45** (2009), no. 3, 802–817.

[16] ———, *Asymptotic constants for minimal distance in the central limit theorem*, Electron. Commun. Probab. **16** (2011), 96–103.

[17] John B. Walsh, *The rate of convergence of the binomial tree scheme*, Finance Stoch. **7** (2003), no. 3, 337–361.

[18] J. Zhang, *A numerical scheme for BSDEs*, Ann. Appl. Probab. **14** (2004), no. 1, 459–488.

[19] ———, *Representation of solutions to BSDEs associated with a degenerate FSDE*, Ann. Appl. Probab. **15** (2005), no. 3, 1798–1831.

[20] ———, *Backward Stochastic Differential Equations*, Springer, New York (2017).