UNIVERSITY OF JYVÄSKYLÄ
DEPARTMENT OF MATHEMATICS
AND STATISTICS

REPORT 169

UNIVERSITÄT JYVÄSKYLÄ
INSTITUT FÜR MATHEMATIK
UND STATISTIK

BERICHT 169

APPROXIMATION OF HEAT EQUATION AND BACKWARD SDES USING RANDOM WALK: CONVERGENCE RATES

ANTTI LUOTO

JYVÄSKYLÄ
2018

UNIVERSITY OF JYVÄSKYLÄ
DEPARTMENT OF MATHEMATICS
AND STATISTICS

REPORT 169

UNIVERSITÄT JYVÄSKYLÄ
INSTITUT FÜR MATHEMATIK
UND STATISTIK

BERICHT 169

APPROXIMATION OF HEAT EQUATION AND BACKWARD SDES USING RANDOM WALK: CONVERGENCE RATES

ANTTI LUOTO

To be presented, with permission of the Faculty of Mathematics and Science
of the University of Jyväskylä, for public criticism in Auditorium H320
on November 10th, 2018, at 12 o'clock noon.

JYVÄSKYLÄ
2018

Editor: Pekka Koskela
Department of Mathematics and Statistics
P.O. Box 35 (MaD)
FI–40014 University of Jyväskylä
Finland

Permanent link to this publication: http://urn.fi/URN:ISBN:978-951-39-7576-0

ISBN 978-951-39-7575-3 (print)
ISBN 978-951-39-7576-0 (pdf)
ISSN 1457-8905

University Printing House
Jyväskylä 2018

Acknowledgements

I would like to express my deepest gratitude to my supervisors Christel Geiss and Anni Laitinen for all their help and guidance throughout my doctoral studies. I am also indebted to my former supervisor Dario Gasbarra for the helpful discussions during the early stage of my studies as well as to Stefan Geiss for his valuable comments and suggestions regarding this work. Finally, I would like to thank my family for their support, and also my friends, office mates, and other colleagues for all the math and non-math related discussions and activities during these years.

Jyväskylä, October 2018

Antti Luoto

List of included articles

This dissertation consists of an introductory part and the following articles:

References

[A] A. Luoto. *Time-dependent weak rate of convergence for functions of generalized bounded variation.* Submitted (2017).

[B] C. Geiss, A. Luoto, and P. Salminen. *On first exit times and their means for Brownian bridges.* In revision (2018).

[C] C. Geiss, C. Labart, and A. Luoto. *Random walk approximation of BSDEs with Hölder continuous terminal condition.* Submitted (2018).

[D] C. Geiss, C. Labart, and A. Luoto. *L_2-Approximation rate of forward-backward SDEs using random walk.* Submitted (2018).

In the introduction these articles will be referred to as [A], [B], [C], and [D], whereas other references will be referred as [Abu02], [AKU16],. . .

The author of this dissertation has actively taken part in the research of the joint articles [B], [C], and [D].

Introduction

This thesis addresses questions related to approximation arising from the fields of stochastic analysis and partial differential equations. Theoretical results regarding convergence rates are obtained by using discretization schemes where the limiting process, the Brownian motion, is approximated by a simple discrete-time random walk.

The rate of convergence is derived for a finite-difference approximation of the solution of a terminal value problem for the backward heat equation. This weak approximation result is proved for a terminal function which has bounded variation on compact sets. The sharpness of the according rate is achieved by applying some new results related to the first exit time behavior of Brownian bridges. In addition, convergence rates in the L_2-norm are proved for Markovian forward-backward stochastic differential equations, where the underlying forward process is either Brownian motion or a more general Itô diffusion.

1. Theoretical background

Suppose that $W = (W_t)_{t\geq 0}$ is a 1-dimensional Brownian motion defined on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t\geq 0})$ which satisfies the usual conditions. Under standard Lipschitz and linear growth assumptions imposed on the drift and the diffusion coefficients b and σ, there exists a unique adapted and continuous process $X = (X_t)_{t\geq 0}$ satisfying the stochastic differential equation (SDE)

$$X_t = x_0 + \int_0^t b(s, X_s)ds + \int_0^t \sigma(s, X_s)dW_s, \quad t \geq 0. \tag{1.1}$$

Besides being of theoretical interest, SDEs have several practical applications e.g. in population dynamics, biology, and especially in mathematical finance. It is therefore important to find ways to obtain realizations of the continuous-time process (1.1) in some approximative manner. Perhaps the most well-known approximation method is the Euler (or Euler-Maruyama) scheme. Although this scheme is not the topic of this thesis, it is described briefly in order to introduce the relevant concepts related to approximation in general and also for future reference.

1.1. **The Euler-Maryuama scheme for SDEs.** Given an integer $n \in \mathbb{N}$, choose a partition $\pi = \pi(n)$ of time instants $0 =: t_0 < t_1 < \cdots < t_n := T$, where $T > 0$ is a fixed time horizon. Given the step size in time $\Delta t_i := t_i - t_{i-1}$ and in space $\Delta W_{t_i} := W_{t_i} - W_{t_{i-1}}$, define $X_0^\pi := x_0 \in \mathbb{R}$ and

$$X_{t_i}^\pi := X_{t_{i-1}}^\pi + b(t_i, X_{t_{i-1}}^\pi)\Delta t_i + \sigma(t_i, X_{t_{i-1}}^\pi)\Delta W_{t_i}, \quad 1 \leq i \leq n. \tag{1.2}$$

It is natural to ask whether X_T^π converges to X_T in some sense as the partition becomes more dense and the mesh size $|\pi| := \sup_{1\leq i\leq n} |\Delta t_i|$ tends to zero. In particular, it is of interest to know how large the associated approximation error is when $|\pi|$ is small. Approximation *in the strong sense* means that for a given $\gamma > 0$ there exist constants $C, \delta > 0$ such that

$$\mathbb{E}|X_T - X_T^\pi| \leq C|\pi|^\gamma \quad \text{whenever} \quad |\pi| < \delta. \tag{1.3}$$

Alternatively, one may consider some other L_p-norm with $p > 1$ instead of the L_1-norm.

For the strong approximation it is required that X_T and each X_T^π are defined on the same probability space. This requirement is no longer necessary when considering approximation *in the weak sense* instead. Given a class $\mathfrak{G}$ of test functions, weak approximation is related to the question of finding for a given $\gamma > 0$ and an arbitrary $g \in \mathfrak{G}$ some constants $C, \delta > 0$ such that

$$|\mathbb{E}[g(X_T)] - \mathbb{E}[g(X_T^\pi)]| \leq C|\pi|^\gamma \quad \text{provided that } |\pi| < \delta. \tag{1.4}$$

Note that this is a sufficient condition for the convergence in distribution of X_T^π to X_T when $\mathfrak{G}$ is the class of continuous and bounded functions.

The constant $\gamma > 0$ appearing in (1.3)–(1.4) is called the *order* of convergence while the quantity $|\pi|^\gamma$ (modulo a multiplicative constant) will be referred to as the *rate* of convergence. One often seeks for the optimal, i.e. the largest possible constant γ for which (1.3) or (1.4) holds true. Several factors influence the rate of convergence: the choice of the partitions (equidistant, deterministic or random), regularity of the coefficients b and σ of (1.1) and properties of the test functions.

The strong (resp. weak) order of convergence for the Euler scheme is known to be $\frac{1}{2}$ (resp. 1) under standard Lipschitz assumptions on b and σ, see [KP99, Theorem 10.2.2]. For the weak rate it is typical to assume that the parameters b and σ are at least twice continuously differentiable in space (see [KP99, Theorem 14.1.5] and [BT96]).

1.2. **Approximation based on simple random walk and the Skorokhod representation.** This thesis studies problems of approximation concerning *simple random walk schemes* instead of the Euler scheme. These problems are time-dependent: For a given partition π, the approximation error is considered as a function of t_k where $t_k \in \pi$. Of particular interest is the potential explosion of the error as $t_k \to T$, where T is the terminal time.

Suppose that π is an equidistant partition whose partition points are given by

$$t_k := hk, \quad 0 \leq k \leq n, \qquad \text{where} \qquad h := \frac{T}{n}.$$

Approximation using random walk is based on the idea of replacing the normally distributed increments ΔW_{t_i} in the Euler scheme by i.i.d. random variables U_i attaining only finitely many values – provided that $\mathbb{E}[U_i] = 0$ and $\mathbb{E}[U_i^2] = h$ still hold. A particularly simple choice is to use scaled Rademacher variables $(U_i)_{i\geq 1}$, i.e. define

$$U_i := \sqrt{h}\xi_i, \quad i \geq 1, \tag{1.5}$$

where $\xi_1, \xi_2, \ldots$ are i.i.d. and satisfy $\mathbb{P}(\xi_i = 1) = \frac{1}{2} = \mathbb{P}(\xi_i = -1)$. This leads to the approximation of the Brownian motion W by a *simple random walk* $W^n = (W_{t_k}^n)_{k\geq 0}$ and the solution X of (1.1) by the process $X^n = (X_{t_k}^n)_{k\geq 0}$, defined as

$$W_0^n := 0, \quad W_{t_k}^n := \sqrt{h}\sum_{i=1}^k \xi_i, \tag{1.6}$$

$$X_0^n := x_0, \quad X_{t_k}^n := x_0 + h\sum_{i=1}^k b(t_i, X_{t_{i-1}}^n) + \sqrt{h}\sum_{i=1}^k \sigma(t_i, X_{t_{i-1}}^n)\xi_i, \tag{1.7}$$

where $x_0 \in \mathbb{R}$ and $1 \le k \le n$.

Note that the Rademacher sequence $(\xi_i)_i$ and the Brownian motion W are not necessarily assumed to be defined on the same probability space. Nevertheless, this will be the case in the approach adopted in [A], [C], and [D], where the random variables ξ_i are obtained by sampling the Brownian motion at certain first exit times. Embedding techniques of this type have gained some renewed interest since the late 1990s (e.g. [RS97], [LR00], [Wal03], [SS04], [BJ08], [AKU16]). A version of the original result due to A.V. Skorokhod [Sko65]) is presented below.

Theorem 1.1 (Skorokhod embedding theorem, [Bil79, Theorem 37.7])**.** *Suppose that $(V_i)_{i\ge1}$ is a sequence of i.i.d. random variables with zero mean and finite variance. For each $k \ge 1$, define $S_k := \sum_{i=1}^k V_k$. Then, there exists a Brownian motion W and a sequence of stopping times $\tau_1 \le \tau_2 \le \dots$ such that*

(i) $\tau_1, \tau_2 - \tau_1, \tau_3 - \tau_2, \dots$ *are i.i.d.,*

(ii) $\mathbb{E}[\tau_i - \tau_{i-1}] = \mathbb{E}[V_1^2]$ *for all* $i \ge 1$,

(iii) $W_{\tau_k} \sim S_k$ *for each* $k \ge 1$ *('$\sim$' means 'to be equal in law').*

For $(V_i)_{i\ge1} := (U_i)_{i\ge1}$, the i.i.d. sequence given by (1.5), a solution to the above problem can be constructed as follows: Let $\tau_0 := 0$, and define recursively

$$\tau_k = \tau_k(n) := \inf\left\{t \ge \tau_{k-1} : |W_t - W_{\tau_{k-1}}| = \sqrt{h}\right\} \qquad \text{for} \quad k \ge 1. \quad (1.8)$$

The properties (i) and (iii) are satisfied for the sequence $(\tau_k)_k$ due to the strong Markov property of Brownian motion. For the property (ii), see Proposition 1.2 below. Consequently, for each $n \ge 1$, $(W_{\tau_k})_{k\ge0}$ is a random walk equal in law with $(W^n_{t_k})_{k\ge0}$, called the *Skorokhod version of* W^n.

This particular representation $(W_{\tau_k})_{k\ge0}$ defined on the same probability space as the Brownian motion has certain advantages over some other approximations of W. For example, one can use Itô calculus and properties of the stopping times $(\tau_k)_{k\ge0}$ to show that the sequence $(W_{\tau_n})_{n\ge0}$ converges to W_T in the strong sense with order $\frac{1}{4}$. Several properties of the process $(\tau_k, W_{\tau_k})_{k\ge0}$ are exploited in [A], also in [C]–[D], some of which are collected below for convenience.

Proposition 1.2 (Properties of the first hitting time and of the Skorokhod version of W^n)**.**

(i) *The distribution of the stopping time* $\tau_1 = \inf\{t \ge 0 : |W_t| = \sqrt{h}\}$, $h > 0$, *is absolutely continuous w.r.t. the Lebesgue measure.*

(ii) *For each integer* $m \ge 1$ *there is a constant* $C_m > 0$ *such that* $\mathbb{E}[\tau_1^m] \le C_m h^m$. *In particular,* $\mathbb{E}[\tau_1] = h$ *and* $\mathbb{E}[\tau_1^2] = \frac{5}{3}h^2$.

(iii) *Let* $T > 0$, $h = \frac{T}{n}$, *and* $t_k = \frac{kT}{n}$ *for each integer* $0 \le k \le n$. *Then, for each* $p > 0$ *there exist constants* $c, c' > 0$, *depending at most on* p *and* T, *such that*

$$\mathbb{E}|W_{t_k} - W_{\tau_k}|^p \le c\mathbb{E}|t_k - \tau_k|^{p/2} \le c'(t_k h)^{\frac{p}{4}}. \quad (1.9)$$

In addition, for $p = 2$ *there is the equality* $\mathbb{E}|W_{t_k} - W_{\tau_k}|^2 = \mathbb{E}|t_k - \tau_k|$.

For the items (i)–(ii), see e.g. [Wal03]. Item (iii) is obtained by a slight generalization of [Wal03, Proposition 11.1 (iv)], see [C, Lemma A.1 (iv)].

2. Approximation of the backward heat equation with an irregular terminal condition

Some early research papers concerning the approximation of the solutions to initial and boundary value problems for the *heat equation* by *finite differences* date back to the 1950s. The idea behind the finite-difference approximation is to replace the partial derivatives by difference quotients. One obtains an indexed family of difference equations the solutions of which are expected to give increasingly better approximations of the PDE solution as the mesh size decreases.

The rate of convergence of a finite-difference scheme for the backward heat equation in the presence of an irregular terminal condition is studied in [A]. Earlier results concerning convergence rates for related problems in various settings have been obtained, for instance, in [JY53], [Rey72], [DK05], and [Lin07].

In [A], an upper bound is derived for the resulting approximation error by reinterpreting the problem as a problem of time-dependent weak approximation. The proof follows and adapts the ideas of J. B. Walsh [Wal03]. The article [Wal03] concerns the rate of convergence of option prices implied by the Black-Scholes model, when they are approximated by the prices implied by the Cox-Ross-Rubinstein (or the binomial tree) model. See also [HZ00] and [LR00].

2.1. **The setting and the formulation of the problem.**

Let $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t\geq 0})$ be a stochastic basis where $(\mathcal{F}_t)_{t\geq 0}$ stands for the natural filtration of a standard Brownian motion $W = (W_t)_{t\geq 0}$. Given a finite time horizon $T > 0$ and a constant parameter $\sigma > 0$, consider the *backward heat equation*

$$\begin{cases} u_t(t,x) + \frac{\sigma^2}{2} u_{xx}(t,x) &= 0, \qquad (t,x) \in [0,T) \times \mathbb{R}, \\ u(T,x) &= g(x), \qquad x \in \mathbb{R}, \end{cases} \tag{2.1}$$

where the terminal condition g is assumed to be an exponentially bounded Borel function. Without additional regularity assumptions imposed on g, we cannot expect that equation (2.1) has a *classical* solution $u \in C^{1,2}([0,T]\times\mathbb{R})$. However, there exists a *probabilistic* solution $u \in C^{1,2}([0,T) \times \mathbb{R})$, which is unique in the class of functions having at most exponential growth (see e.g. [SP12]). This solution can be represented in terms of the Brownian motion $(W_t)_{t\geq 0}$ as

$$u(t,x) := \mathbb{E}[g(x + \sigma W_{T-t})], \quad (t,x) \in [0,T] \times \mathbb{R}.$$

The main objective in [A] is to estimate the approximation error $u^n - u$, where u^n is the solution of the finite-difference equation (2.2) introduced below. A setting slightly different to that of Section 1 is considered here. For a given $n \in 2\mathbb{N}$, choose a partition $\pi := \{0 =: t_0 < t_1 < \cdots < t_{n/2} := T\}$ by letting $t_k := 2kh = \frac{2kT}{n}$ for each k. The double step size $\Delta t_k = \frac{2T}{n}$ is opted in order to avoid the so-called 'sawtooth effect' of the error. Proceed by fixing the space increment $\Delta x := 2\sigma\sqrt{h}$. Then, for each fixed level $z_0 \in \mathbb{R}$, define a partition of space by letting $\mathcal{S}_{z_0} := \{z_0 + m\Delta x \,|\, m \in \mathbb{Z}\}$. The finite-difference equation associated to a given grid $\pi \times \mathcal{S}_{z_0} \subset [0,T] \times \mathbb{R}$ is given by

$$\begin{cases} \dfrac{u^n(t_k,x)-u^n(t_{k-1},x)}{\Delta t} + \dfrac{\sigma^2}{2}\dfrac{u^n(t_k,x+\Delta x)-2u^n(t_k,x)+u^n(t_k,x-\Delta x)}{(\Delta x)^2} = 0, \\ u^n(T,x) = g(x). \end{cases} \tag{2.2}$$

Given g and the corresponding set of terminal values on $\{T\} \times \mathcal{S}_{z_0}$, the solution $u^n : \pi \times \mathcal{S}_{z_0} \to \mathbb{R}$ of (2.2) is uniquely determined by backward recursion, and it satisfies

$$u^n(t_k,x) = \mathbb{E}[g(x+\sigma W^n_{T-t_k})], \quad (t_k,x) \in \pi \times \mathcal{S}_{z_0},$$

where W^n is the random walk given by (1.6). Consequently, the connection between the original problem and the problem related to weak approximation of W by W^n is

$$|u^n(t_k,x) - u(t_k,x)| = |\mathbb{E}[g(x+\sigma W^n_{T-t_k})] - \mathbb{E}[g(x+\sigma W_{T-t_k})]|. \tag{2.3}$$

2.2. **Functions of generalized bounded variation.** In [Wal03], a detailed error expansion was derived for the *weak error*

$$u^n(0,0) - u(0,0) = \mathbb{E}[g(\sigma W_{\tau_n})] - \mathbb{E}[g(\sigma W_T)] \tag{2.4}$$

(recall $W_{\tau_n} \sim W^n_T$). The error was shown to converge with order 1 or $\frac{1}{2}$ depending on the position of the possible discontinuities of g. The result was then transferred to the geometric setting of the CRR and the Black-Scholes models by certain transformations including a Girsanov transform. The test functions $g : \mathbb{R} \to \mathbb{R}$ considered in (2.4) were assumed to satisfy the following conditions:

- g is piecewise $C^2(\mathbb{R})$ such that g, g', and g'' are exponentially bounded,
- $g(x) = \frac{1}{2}(g(x-) + g(x+))$ at each point $x \in \mathbb{R}$.

This class excludes e.g. functions with infinitely many jumps and indicator functions of intervals. To obtain rates for such functions, let us first recall the notion of bounded variation.

The total variation V_g of a function $g : \mathbb{R} \to \mathbb{R}$ is defined as

$$V_g(x) := \sup \sum_{i=1}^{N} |g(x_i) - g(x_{i-1})|, \quad x \in \mathbb{R},$$

where the supremum is taken over N and each partition $-\infty < x_1 < \cdots < x_N = x$. The function g is said to be of bounded variation provided that the limit $\lim_{x\to\infty} V_g(x) < \infty$ exists. Each (left-continuous) function g of bounded variation can be represented as a sum of a constant and a distribution function of a finite signed Borel measure on $\mathbb{R}$, see [Rud74]. The requirement $\lim_{x\to\infty} V_g(x) < \infty$ is, however, rather restrictive since it forces the function itself to be bounded. A method to circumvent this restriction was presented in [Avi09], where the concept of bounded variation was generalized to allow functions to have prescribed asymptotic growth by introducing weights.

Definition 2.1 ([Avi09, Definition 3.2]). *Denote by $\mathcal{M}$ the class of all set functions μ on the ring of* bounded *sets of $\mathcal{B}(\mathbb{R})$ for which there exists measures $\mu^1, \mu^2 : \mathcal{B}(\mathbb{R}) \to [0,\infty]$ such that $\mu = \mu_1 - \mu_2$, and $\mu^1(K), \mu^2(K) < \infty$ for each compact set $K \in \mathcal{B}(\mathbb{R})$.*

Even though a set function $\mu \in \mathcal{M}$ is not a signed measure being undefined on unbounded sets, it admits a minimal decomposition (similar to the Hahn-Jordan decomposition) which induces a unique σ-finite measure $|\mu|$ on $\mathcal{B}(\mathbb{R})$, called the total variation of μ.

In [Avi09], classes of functions of *generalized bounded variation* are constructed by assigning a weight function to control the growth of the distribution function associated to μ. The class GBV_{exp} considered in [A] can be seen as an instance of such a class with weight functions having exponential decay.

Definition 2.2 (The class GBV_{exp}). *Denote by GBV_{exp} the class of functions $g : \mathbb{R} \to \mathbb{R}$ which can be represented as*

$$g(x) = c + \mu([0,x)) - \mu([x,0)) + \sum_{i=1}^{\infty} \alpha_i \mathbb{1}_{\{x_i\}}(x), \quad x \in \mathbb{R}, \tag{2.5}$$

where $c \in \mathbb{R}$ is a constant, $\mu \in \mathcal{M}$, and $\mathcal{J} = (\alpha_i, x_i)_{i=1,2,\ldots} \subset \mathbb{R}^2$ is a countable set such that $x_i \neq x_j$ whenever $i \neq j$. In addition, it is required that for some constant $\beta \geq 0$,

$$\int_{\mathbb{R}} e^{-\beta|x|} d|\mu|(x) + \sum_{i=1}^{\infty} |\alpha_i| e^{-\beta|x_i|} < \infty. \tag{2.6}$$

Each polynomial belongs to the class GBV_{exp}. Moreover, indicator functions of intervals (open, closed, or semi-open) and their linear combinations belong to the class GBV_{exp}.

2.3. **Results.** The main result of [A] is given below.

Theorem 2.3 ([A, Theorem 2.4(A)]). *Suppose that $g \in GBV_{exp}$ is given by (2.5) with $\beta \geq 0$ as in (2.6), and that $z_0 \in \mathbb{R}$. Then, for all $(t_k, x) \in \pi^n \times \mathcal{S}^n_{z_0}$ with $t_k \neq T$,*

$$|u^n(t_k,x) - u(t_k,x)| \leq \frac{Ce^{\beta|x|}}{\sqrt{T-t_k}} \frac{1}{\sqrt{n}}, \tag{2.7}$$

where $C = C(\beta, \sigma, T, g) > 0$ is a constant.

Theorem 2.3 implies the rate $n^{-1/2}$ locally on compact sets of $[0,T) \times \mathbb{R}$ for the test function class GBV_{exp}. Moreover, the rate is sharp for this class in the following sense: There exists a function $g \in GBV_{\mathrm{exp}}$ such that

$$0 < \liminf_{n\to\infty} n^{\frac{1}{2}} |u^n(0,0) - u(0,0)| \leq \limsup_{n\to\infty} n^{\frac{1}{2}} |u^n(0,0) - u(0,0)| < \infty. \tag{2.8}$$

Indeed, (2.8) is seen to hold true for the step function $g = \mathbb{1}_{[0,\infty)}$ ([A, Proposition 4.14]).

Functions of bounded variation were considered as a class of test functions by M. L. Juncosa and D. M. Young in [JY53] in the context of an initial value problem associated to the forward heat equation on $[0,\infty) \times [-1,1]$. Using Fourier methods, they obtained the rate $n^{-1/2}$ locally as well, but did not study the explosion in t_k.

The requirement for the terminal condition g to be of bounded variation on compact sets is not necessary in order to obtain upper bounds of the type (2.7). A result similar to Theorem 2.3 is proved for a class of exponentially bounded, locally α-Hölder continuous test functions.

Theorem 2.4 ([A, Theorem 2.4(B)])**.** *Suppose that $g : \mathbb{R} \to \mathbb{R}$ is a function for which there exist constants $A, \beta \geq 0$ and $\alpha \in (0, 1]$ such that for each $R > 0$,*

$$\sup_{x,y\in[-R,R],x\neq y} \frac{|g(x) - g(y)|}{|x - y|^\alpha} \leq Ae^{\beta R}.$$

Then, for all $(t_k, x) \in \pi \times \mathcal{S}_{z_0}$ with $t_k \neq T$, there exists $C = C(\beta, \sigma, T, g) > 0$ such that

$$|u^n(t_k, x) - u(t_k, x)| \leq \frac{Ce^{(\beta+1)|x|}}{(T - t_k)^{\frac{\alpha}{2}}} \frac{1}{n^{\alpha/2}}.$$

2.4. **Method of the proof.** The proof of Theorem 2.3 and 2.4 is based on the Skorokhod representation of W^n. As in [Wal03], the error (2.3) is split into a 'global' and a 'local' part, which take into account different properties of the terminal function g.

Let $\sigma = 1$ for simplicity. To take into account the time-dependence, introduce for a fixed time instant $s_0 \in [0, T)$ the auxiliary variables

$$n_\theta := 2 \left\lceil \tfrac{T-s_0}{\Delta t} \right\rceil \in \{2, 4, \ldots, n\}, \quad \theta_n := \tfrac{n_\theta T}{n} \in \left\{ \tfrac{2T}{n}, \tfrac{4T}{n}, \ldots, T \right\}.$$

Then $\theta_n = T - t_k = \frac{n_\theta T}{n}$ if and only if $s_0 \in [t_k, t_{k+1})$, so that

$$u^n(t_k, x) - u(t_k, x) = \mathbb{E}[g(x + W_{\tau_{n_\theta}})] - \mathbb{E}[g(x + W_{\theta_n})], \; s_0 \in [t_k, t_{k+1}). \quad (2.9)$$

The error (2.9) may thus be considered as a function of the pair (n, θ_n), where the blow-up rate is described in terms of $1/\theta_n$. By introducing the random index $J_n := \sup\{2m \in 2\mathbb{N} : \tau_{2m} > \theta_n\}$ associated to the sequence of stopping times $(\tau_k)_{k\geq 0}$, the error is finally decomposed into the sum of

$$\varepsilon_n^{\text{glob}}(t_k, x) := \mathbb{E}[g(x + W_{\tau_{n_\theta}}) - g(x + \sigma W_{\tau_{J_n}})] \quad \text{and} \quad (2.10)$$

$$\varepsilon_n^{\text{loc}}(t_k, x) := \mathbb{E}[g(x + W_{\tau_{J_n}}) - g(x + \sigma W_{\theta_n})]. \quad (2.11)$$

The main idea behind the estimation of the *global error* (2.10) is to expand it into an infinite series whose terms can be expressed with the help of the moments of J_n and $W_{\tau_{n_\theta}}$. For the analysis of this expansion, technical moment and tail estimates for the random times τ_{n_θ}, J_n, and τ_{J_n} are derived by generalizing estimates stated in [Wal03] to the time-dependent setting. Part of this extension is carried out in the arXiv version [Luo17] of [A]. For instance, for each $p > 0$ one finds a constant $C_p > 0$ such that

$$\mathbb{E}\left(\frac{|J_n - n_\theta|}{\sqrt{n_\theta}}\right)^p \leq C_p \quad , \quad \text{uniformly in } (n, s_0).$$

As opposed to the global error, the *local error* (2.11) is influenced by the smoothness and the local behavior of the terminal condition g. By considering the conditional expectation $\mathbb{E}[g^x(W_{\tau_{J_n}})|\mathcal{F}_{\theta_n}]$ for the shifted function $g^x := g(x + \cdot)$, one may represent (2.11) as

$$\varepsilon_n^{\mathrm{loc}}(t_k, x) = \mathbb{E}\big[\Pi_e g^x(W_{\theta_n}) - g^x(W_{\theta_n})\big] \\ + \mathbb{E}\big[\big(\Pi_o \Pi_e g^x(W_{\theta_n}) - \Pi_e g^x(W_{\theta_n})\big)\mathbb{P}(\tau_{L_n} \text{ is even}|W_{\theta_n})\big] \quad (2.12)$$

in terms of the random index $L_n := \sup\{m \in \mathbb{N} \cup \{0\} : \tau_m < \theta_n\}$. Here, Π_e and Π_o are projection operators mapping functions $f : \mathbb{R} \to \mathbb{R}$ into piecewise linear functions according to the rule

$$\Pi_e f(x) = f(x) \quad \text{on} \quad \mathbb{Z}_e^h := \{2k\sqrt{h} : k \in \mathbb{Z}\},$$
$$\Pi_o f(x) = f(x) \quad \text{on} \quad \mathbb{Z}_o^h := \{(2k+1)\sqrt{h} : k \in \mathbb{Z}\}.$$

Since these operators are linear, they are well-suited for the estimation of functions belonging to $GBV_{\exp}$ due to the representation (2.5).

It turns out to be crucial (e.g. for the proof of the sharpness result (2.8)) to further estimate the conditional probability

$$\mathbb{P}(\tau_{L_n} \text{ is even}|W_{\theta_n} = x), \quad x \in \mathbb{R}, \quad (2.13)$$

appearing in (2.12). The random walk $(W_{\tau_k})_{k\geq 0}$ either moves up or down the amount $\sqrt{h}$ at each step, and thus the condition 'τ_{L_n} is even' means that the final value the process admits before time θ_n is an *even* multiple of $\sqrt{h} = \sqrt{T/n}$. Using the notation of this section, it was pointed out using time reversal in [Wal03, pp. 348–349] that the probability (2.13) is equal to

$$q_n(x) := \mathbb{P}(\text{'}B_t^{x,\theta_n,0} \text{ hits } \mathbb{Z}_e^h \text{ before hitting } \mathbb{Z}_o^h\text{'}),$$

where $(B_t^{x,\theta_n,0}) = (B_t^{x,\theta_n,0})_{t\in[0,\theta_n]}$ denotes the *Brownian bridge from x to 0 of length θ_n*. By comparing this probability to the known hitting probability for the Brownian motion

$$\tilde{q}_n(x) := \mathbb{P}(\text{'}(x+W_t)_{t\geq 0} \text{ hits } \mathbb{Z}_e^h \text{ before hitting } \mathbb{Z}_o^h\text{'}) = \frac{\mathrm{dist}(x, \mathbb{Z}_o^h)}{\sqrt{h}},$$

it was stated that $q_n(x) - \tilde{q}_n(x) = O(n^{-\frac{1}{2}})$ as $n \to \infty$. It can be shown that this convergence holds for any fixed $x \in \mathbb{R}$, but not uniformly in x. Applying results of [B] related to the *expected first exit times* of Brownian bridges, it turns out that more can be said: For a known function ψ, it holds that

$$|q_n(x) - \tilde{q}_n(x)| \leq \psi(x, h, \theta_n, T)\mathbb{E}[\mathcal{T}_{x,\theta_n,0}],$$

where $\mathcal{T}_{0,\theta_n,x}$ denotes the first exit time of the bridge $B^{0,\theta_n,x}$ from a certain interval of length $\sqrt{h}$ containing x. More details can be found in [A, Subsection 4.2] and [B, Section 4].

3. The expected first exit times of Brownian bridges

The research paper [B] studies the first exit time of a Brownian bridge from an open interval. Of particular interest is the expected amount of time this process spends before exiting the interval for the first time in relation to the length of the interval.

For the Brownian motion W, it is known that $\mathbb{E}[W_{\mathcal{T}_{(a,b)}}] = |a|b$, where $\mathcal{T}_{(a,b)}$ denotes the first exit time of from the open interval $0 \in (a,b)$. In particular, $\mathbb{E}[W_{\mathcal{T}_{(-h,h)}}] = h^2$. In [B], the associated expectation for the Brownian bridge $B^{x,T,y} = (B_t^{x,T,y})_{t\in[0,T]}$ is proved to be of the same order for small $h > 0$:

$$\lim_{h\downarrow 0} \frac{\mathbb{E}\big[B^{x,T,y}_{\mathcal{T}_{(-h,h)}}\big]}{h^2} = 1, \qquad T > 0, \quad x \neq y.$$

Similar limiting behavior is found also for the 3-dimensional Bessel bridge.

The bridge $B^{x,T,y}$ can be thought of as a Brownian motion initiated at x and 'conditioned to hit y at time T'. It is a continuous Gaussian process which can be represented in several ways. For example, it can be obtained as a solution to an SDE, or be defined in terms of the Brownian motion by letting

$$B_t^{x,T,y} := W_t - \frac{t}{T}W_T + x + (x-y)\frac{t}{T}, \quad t \in [0,T].$$

To study more general bridges $X^{x,T,y}$, it is typical to associate them to a probability measure on the *canonical space*, i.e. to search for a probability law on a suitable function space under which the random variables $X_t^{x,T,y}$ play the role of coordinate projections. If the bridge is constructed from a linear (i.e. one-dimensional) diffusion, it is common to consider the space of continuous functions $C := C([0,\infty),\mathbb{R}) := \{\omega : [0,\infty) \to \mathbb{R} : \omega \text{ is continuous}\}$.

For some related literature, see [BO99], [PW01], [Abu02], and [SY11].

3.1. **The canonical framework for linear diffusions and bridges.** The space C is equipped with the natural filtration $(\mathcal{C}_t)_{t\geq 0}$ generated by the coordinate process $\hat{\pi} = (\hat{\pi}_t)_{t\geq 0}$ consisting of projection mappings $\hat{\pi}_t : C \to \mathbb{R}$, $\hat{\pi}_t(\omega) := \omega(t)$ for each $t \geq 0$. The smallest σ-algebra containing each $\mathcal{C}_t$ is denoted by $\mathcal{C}$. For each $(\mathcal{C}_t)_{t\geq 0}$-stopping time τ, define also the *random shift operator*

$$\varrho_\tau : \{\omega \in C : \tau(\omega) < \infty\} \to C, \ \varrho_\tau(\omega) := \omega(\,\cdot\, + \tau(\omega)).$$

Roughly speaking, a linear diffusion is a strong Markov process with continuous paths. The following technical definition is taken from [BS15].

Definition 3.1 (Linear diffusion). *Let $I \subset \mathbb{R}$ be an interval. For each $x \in I$, let $\mathbb{P}_x$ denote the probability measure on $(C,\mathcal{C})$ under which a homogeneous Markov process $X = (X_t)_{t\geq 0}$ taking values in I is the canonical process $\hat{\pi}$ started at x. It is called a* linear diffusion *provided that for each $x \in I$,*

- $\mathbb{P}_x(\Lambda \circ \varrho_\tau | \mathcal{C}_{\tau_+}) = \mathbb{P}_{X_\tau}(\Lambda) \quad \mathbb{P}_x$*-a.s. on* $\{\tau < \infty\}$,
- $\mathbb{P}_x(t \mapsto \omega(t)$ *is continuous on* $[0,\zeta(\omega))) = 1$,

where Λ is any bounded $\mathcal{C}$-measurable random variable, τ is any $(\mathcal{C}_t)_{t\geq 0}$-stopping time, and where $\zeta(\omega) := \inf\{t \geq 0 : \omega(t) \notin I\} \in [0,\infty]$ is the lifetime.

For later use, let us denote the first hitting time $\mathcal{T}_y$ to $y \in \mathbb{R}$ and the first exit time $\mathcal{T}_{(a,b)}$ from the interval $(a,b) \subset \mathbb{R}$, respectively, by

$$\mathcal{T}_y(\omega) := \inf\{t \geq 0 : \omega(t) = y\}, \quad \mathcal{T}_{(a,b)}(\omega) := \inf\{t \geq 0 : \omega(t) \notin (a,b)\}$$

A linear diffusion $(X_t)_{t\geq 0}$ is *regular* if $\mathbb{P}_x(\mathcal{T}_y < \infty) > 0$ holds true for each $x, y \in I$, i.e. the process X will eventually hit any given point on the interval. Each regular linear diffusion X admits a strictly positive, continuous transition density $p_t^X(x,y)$ which is also *symmetric* (see [IM65, p.149, p.157]). The transition density p^X is a map $[0,\infty) \times I^2 \to [0,\infty)$ s.t.

$$\mathbb{P}_x(X_t \in A) = \int_A p_t^X(x,y) m^X(dy), \qquad A \in \mathcal{B}(I),\ x \in I,\ t > 0$$

holds for some reference measure m^X on $\mathcal{B}(I)$ (the Borel σ-algebra of subsets of I), called the *speed measure* of X.

The key result regarding the construction of bridges is the following: Suppose that $(X_t)_{t\geq 0}$ is a regular linear diffusion, and let $x, y \in I$. Then, there exists a unique probability measure $\mathbb{Q}$ on $(C, \mathcal{C}_T)$ such that $\mathbb{Q}(\omega(0) = x) = \mathbb{Q}(\omega(T-) = y) = 1$, and

$$\mathbb{Q}(A) = \mathbb{Q}\big|_{\mathcal{C}_t}(A) = \mathbb{E}_x\left[\frac{p_{T-t}^X(\omega(t), y)}{p_T^X(x,y)} \mathbb{1}_A\right] \quad \text{for each } A \in \mathcal{C}_t,\ 0 \leq t < T \quad (3.1)$$

([FPY92], cf. [Sal97] and [CU11]). The property (3.1) implies that each restriction $\mathbb{Q}\big|_{\mathcal{C}_t}$ to the sub-σ-algebra $\mathcal{C}_t \subset \mathcal{C}_T$ is absolutely continuous w.r.t. $\mathbb{P}_x$. This measure $\mathbb{Q}$ is denoted by $\mathbb{P}^X_{x,T,y}$ and called the *law of the diffusion bridge* $X^{x,T,y}$.

In the following, $\mathbb{P}_x$ (resp. $\mathbb{P}_x^{(3)}$) denotes the law of the Brownian motion (resp. the 3-dimensional Bessel process) started at x. The associated bridge measures will be denoted by $\mathbb{P}_{x,T,y}$ and $\mathbb{P}^{(3)}_{x,T,y}$, respectively.

3.2. **The main results.** Besides the Brownian bridge, also the three-dimensional Bessel bridge is considered in [B], the latter being constructed by conditioning 3-dimensional Bessel process. The 3-dimensional Bessel process is a regular linear diffusion having the law of the Euclidean norm of the 3-dimensional Brownian motion. It is transient, and never hits zero. The transition densities (w.r.t. to the Lebesgue measure) $p_t(x,y)$ of the Brownian motion and $p_t^{(3)}(x,y)$ of the 3-dimensional Bessel process are given by

$$p_t(x,y) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{(x-y)^2}{2t}}, \quad x, y \in \mathbb{R},$$
$$p_t^{(3)}(x,y) = \frac{y}{x}(p_t(x,y) - p_t(x,-y)), \quad x, y > 0.$$

To state the main result of [B], let us also introduce the function

$$F_K(h) := \sum_{m=-\infty}^{\infty} (-1)^m e^{-2m^2h^2}, \quad h > 0.$$

Theorem 3.2 ([B, Theorem 3.4, 3.8])**.**
(*i*) *For the Brownian bridge* $\mathbb{P}_{0,T,y}$ *with* $|y| \geq h$,

$$\mathbb{E}_{0,T,y}[\mathcal{T}_{(-h,h)}] = h \int_0^T \frac{p_{T-t}(0,y)}{p_T(0,y)} \frac{F_K(h/\sqrt{t})}{\sqrt{2\pi t}} dt. \quad (3.2)$$

(ii) For the 3-D Bessel bridge $\mathbb{P}^{(3)}_{x,T,y}$ with $x > h$ and $y \notin (x-h, x+h)$,

$$\mathbb{E}^{(3)}_{x,T,y}[\mathcal{T}_{(x-h,x+h)}] = h \int_0^T \frac{p^{(3)}_{T-t}(x,y)}{p^{(3)}_T(x,y)} \frac{F_K(h/\sqrt{t})}{\sqrt{2\pi t}} dt. \tag{3.3}$$

(iii) The expectations in (i)–(ii) have the following limiting behavior as $h \to 0$:

$$\lim_{h \downarrow 0} \frac{\mathbb{E}_{0,T,y}[\mathcal{T}_{(-h,h)}]}{\mathbb{E}_0[\mathcal{T}_{(-h,h)}]} = 1, \qquad \lim_{h \downarrow 0} \frac{\mathbb{E}^{(3)}_{x,T,y}[\mathcal{T}_{(x-h,x+h)}]}{\mathbb{E}^{(3)}_x[\mathcal{T}_{(x-h,x+h)}]} = 1.$$

3.3. **Some remarks.** Theorem 3.2 (*iii*) provides an asymptotic relation between the expected first exit times for the diffusions and the diffusion bridges under consideration. Notice that the influence of the conditioning vanishes in the limit. To which extent the result is true for general diffusion bridges remains an open question. See, however, the result [B, Proposition 3.3] concerning first exit times of (unconditional) regular linear diffusions.

The integral representations in Theorem 3.2 are obtained via the following general result for diffusion bridges $X^{x,T,y}$: For each $x \in (a,b) \subset I$ and $y \in I \setminus (a,b)$,

$$\mathbb{E}_{x,T,y}[\mathcal{T}_{(a,b)}] = \int_0^T \int_a^b \mathbb{P}^X_x(\tau_{(a,b)} > t, X_t \in dz) \frac{p^X_{T-t}(z,y)}{p^X_T(x,y)} dt \tag{3.4}$$

(see [B, Proposition 3.2]). The idea is to show that the Laplace transform w.r.t. the variable T of the right-hand side of (3.4) agrees with the Laplace transform of the corresponding single-integral representations (3.2) or (3.3).

The function F_K, the *Kolmogorov distribution function* named after A. N. Kolmogorov, was originally obtained as the limiting distribution of the Kolmogorov-Smirnov test statistics [Kol33]. The connection between the function F_K and the Brownian bridge was later established by J. L. Doob in [Doo49], who identified F_K as the the distribution function of the supremum of the absolute standard Brownian bridge,

$$\mathbb{P}_{0,1,0}\left(\sup_{0 \le s \le 1} |\omega(s)| \le x\right) = F_K(x).$$

There is also the following fact about the distribution function F_K due to F. B. Knight [Kni69]: The function

$$t \mapsto \frac{F_K(h/\sqrt{t})}{h\sqrt{2\pi t}}, \quad t > 0,$$

is the density of the last visit to zero $\lambda_0 := \sup\{t > 0 : \widehat{W}_t = 0\}$ of $(\widehat{W}_t)_{t \ge 0}$, the Brownian motion killed at $\mathcal{T}_{(-h,h)}$ (see [B, Proposition 3.10] for a different proof). For a discussion on special functions such as F_K and their relations to probability theory, see [BPY01].

4. Approximation of forward-backward stochastic differential equations using simple random walk

Recall the setting introduced in Section 1. A stochastic equation of the type

$$Y_t = \xi + \int_t^T f(s, Y_s, Z_s)ds - \int_t^T Z_s dW_s, \quad 0 \le t \le T, \tag{4.1}$$

is called a backward stochastic differential equation (BSDE). The *data* consists of a pair (ξ, f), where the *terminal condition* ξ is an $\mathcal{F}_T$-measurable random variable, and the *generator* f is real-valued, possibly random function. Given such a pair, the *solution* of (4.1) is a pair $(Y, Z) = (Y_t, Z_t)_{t\in[0,T]}$ of $(\mathcal{F}_t)_{t\in[0,T]}$-adapted processes belonging to the L_2-space $\mathbb{S}_2 \times \mathbb{H}_2$ defined by

$$\mathbb{S}_2 := \mathbb{S}_2^{[0,T]} := \{\varphi : [0,T]\times\Omega \to \mathbb{R} \,|\, \varphi \text{ is cádlág, adapted, and } \|\varphi\|_{\mathbb{S}_2} < \infty\},$$
$$\mathbb{H}_2 := \mathbb{H}_2^{[0,T]} := \{\varphi : [0,T]\times\Omega \to \mathbb{R} \,|\, \varphi \text{ is prog. measurable and } \|\varphi\|_{\mathbb{H}_2} < \infty\},$$

where the norms are given by

$$\|\varphi\|_{\mathbb{S}_2}^2 := \mathbb{E} \sup_{t\in[0,T]} |\varphi_t|^2 \quad \text{and} \quad \|\varphi\|_{\mathbb{H}_2}^2 := \mathbb{E}\int_0^T |\varphi_t|^2 dt.$$

Existence and uniqueness of solutions to (4.1) for a non-linear generator f was first proved by E. Pardoux and S. Peng [PP90] in the presence of the following assumptions: ξ is square-integrable, $f(t, \omega, y, z)$ is Lipschitz continuous w.r.t. the spatial variables y, z, and f satisfies $f(\,\cdot\,, \cdot\,, 0, 0) \in \mathbb{H}_2$. Since then, considerable amount of effort has been dedicated to weaken these assumptions and to study generalizations of such equations. BSDEs have applications in mathematical finance, stochastic control theory and stochastic game theory.

4.1. **FBSDEs.** Forward-backward stochastic differential equations (FBSDEs) are backward SDEs, where the solution $(X_t)_{t\ge 0}$ of a forward SDE serves as the source of randomness for the generator and the terminal condition. The FBSDEs considered in this thesis are assumed to be *Markovian*, i.e. it is additionally assumed that $f(t, \omega, y, z) = f(t, X_t(\omega), y, z)$ and $\xi(\omega) = g(X_T(\omega))$.

Given an initial point $(t, x) \in [0, T) \times \mathbb{R}$ together with a quadruplet (b, σ, g, f) of functions $b, \sigma : [0, T] \times \mathbb{R} \to \mathbb{R}$, $g : \mathbb{R} \to \mathbb{R}$, and $f : [0, T] \times \mathbb{R}^3 \to \mathbb{R}$, a FBSDE is a (parametrized) pair of equations

$$\begin{cases} X_s^{t,x} = x + \int_t^s b(r, X_r^{t,x})dr + \int_t^s \sigma(r, X_r^{t,x})dW_r, & t \le s \le T, \\ Y_s^{t,x} = g(X_T^{t,x}) + \int_s^T f(r, X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x})dr - \int_s^T Z_r^{t,x} dW_r, & t \le s \le T. \end{cases} \tag{4.2}$$

A solution of the system (4.2) is a triplet $(X_s^{t,x}, Y_s^{t,x}, Z_s^{t,x})_{s\in[t,T]} \in \mathbb{S}_2^{[t,T]} \times \mathbb{S}_2^{[t,T]} \times \mathbb{H}_2^{[t,T]}$, adapted to the augmentation of the filtration $(\mathcal{F}_s^t)_{s\in[t,T]}$, $\mathcal{F}_s^t := \sigma(W_r - W_t, t \le r \le s)$. The law of the Brownian motion W started at (t, x) is denoted by $\mathbb{P}_{t,x}$. As a convention, let us simply write $(X, Y, Z) = (X^{0,x}, Y^{0,x}, Z^{0,x})$ when $t = 0$.

Existence and uniqueness of solutions to (4.2) was shown by E. Pardoux and S. Peng [PP92]. They also investigated the connection between the FBSDE (4.2) and the quasilinear second-order parabolic PDE

$$\begin{cases} u_t(t,x) + \mathcal{L}u(t,x) + f(t,x,u(t,x),\sigma(t,x)u_x(t,x)) & = 0, \quad (t,x) \in [0,T) \times \mathbb{R}, \\ u(T,x) & = g(x), \quad x \in \mathbb{R}, \end{cases} \tag{4.3}$$

where $\mathcal{L}u(t,x) := b(t,x)u_x(t,x) + \frac{1}{2}\sigma^2(t,x)u_{xx}(t,x)$. Under smoothness assumptions on (b,σ,f,g), it was shown in [PP92] that if $(Y_s^{t,x}, Z_s^{t,x})_{s\in[t,T]}$ is the unique solution to the BSDE (4.2) for each $(t,x) \in [0,T] \times \mathbb{R}$, then $u(t,x) := Y_t^{t,x}$ belongs to $C^{1,2}([0,T] \times \mathbb{R})$ and solves (4.3). In addition,

$$Y_s^{t,x} = u(s, X_s^{t,x}) \quad \text{and} \quad Z_s^{t,x} = \sigma(s, X_s^{t,x})u_x(s, X_s^{t,x}),$$

i.e. it is possible to represent the solution of the FBSDE as a function of the current state of the forward process. Moreover, under less restrictive assumptions (g is only Lipschitz continuous), $u(t,x) := Y_t^{t,x}$ is the unique *viscosity* solution to (4.3), and this solution can be represented in terms of the nonlinear Feynman-Kac formula

$$u(t,x) = \mathbb{E}\left[g(X_T^{t,x}) + \int_t^T f(r, X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x})dr\right], \quad (t,x) \in [0,T] \times \mathbb{R}. \tag{4.4}$$

The results of [C]–[D] rely on the fact that a representation similar to (4.4) can be found also for the process $Z^{t,x}$.

4.2. **Approximation of BSDEs.** First results related to approximation of backward SDEs date back to the late 1990s and are due to V. Bally [Bal97] and D. Chevance [Che97a], [Che97b]. The known methods can be divided into two categories: *time discretization* (or *backward Euler*) *methods* ([Zha04], [BT04], [BD07]) and *random walk methods* ([BDM01], [MPSMT02], [PX11]). The main idea behind the former is as follows. Given a partition $\pi^n : 0 = t_0 < t_1 < \cdots < t_n = T$, let $\Delta t_i = t_i - t_{i-1}$, $\Delta W_{t_i} = W_{t_i} - W_{t_{i-1}}$, and consider the discrete-time equation

$$Y_{t_{i-i}}^\pi = Y_{t_i}^\pi + f(t_i, X_{t_i}^\pi, Y_{t_i}^\pi, Z_{t_i}^\pi)\Delta t_i - Z_{t_i}^\pi \Delta W_{t_i}, \quad 1 \le i \le n, \tag{4.5}$$

where $Y_{t_n}^\pi := g(X_{t_n}^\pi)$ and the discretization X^π of X is given by (1.2). This system can be solved by backward recursion: Given the quantity

$$Z_{t_{i-1}}^\pi := \frac{1}{\Delta t_i}\mathbb{E}_{t_{i-1}}[Y_{t_i}^\pi \Delta W_{t_i} | \mathcal{G}_i^\pi], \quad i \le n, \tag{4.6}$$

where $\mathcal{G}_i^\pi := \sigma(X_{t_j}^\pi, 0 \le j \le i)$, one obtains $Y_{t_{i-1}}^\pi$ by solving the implicit equation

$$Y_{t_{i-1}}^\pi = \mathbb{E}[Y_{t_i}^\pi | \mathcal{G}_{i-1}^\pi] + f(t_{i-1}, X_{t_{i-1}}^\pi, Y_{t_{i-1}}^\pi, Z_{t_{i-1}}^\pi).$$

This method was considered by B. Bouchard and N. Touzi [BT04]. A similar, explicit scheme is obtained as a special case of the method proposed in J. Zhang [Zha04, Section 6]. The rate $|\pi|^{1/2}$ in the $\mathbb{S}_2 \times \mathbb{H}_2$-norm was obtained for both of these schemes under Lipschitz assumptions on the data.

From the perspective of implementation, a problem related to these time discretization algorithms is the computation of conditional expectations such as (4.6) using

Monte Carlo methods. These computations are considerably less complex in the context of random walk approximation, where the Brownian increments ΔW_{t_i} are replaced by random variables with finitely many values.

In [BDM01], P. Briand, B. Delyon, and J. Mémin studied the convergence of a random walk scheme based on Rademacher-distributed increments. Their main result concerns a setting where the approximating random walks $(W^n)_{n=1,2,\dots}$ given by (1.6) and the Brownian motion W are defined in the same probability space. Provided that $\sup_{t\in[0,T]} |W_t^n - W_t|$ converges to zero in probability as $n \to \infty$, it was shown that

$$\sup_{t\in[0,T]} |Y_t - Y_t^n| + \int_0^T |Z_t - Z_t^n|^2 dt \xrightarrow{n\to\infty} 0 \quad \text{in probability}$$

for the associated backward scheme (Y^n, Z^n).

4.3. **The setting and the main results.** The motivation behind [C]–[D] was to obtain a strong rate of convergence when the FBSDE (4.2) is approximated using a random walk. To the author's best knowledge, results about convergence rates for random walk schemes for BSDEs have not been previously obtained – unlike for time-discretization schemes as mentioned above. The case where the forward process X is simply the Brownian motion is studied in [C], whereas [D] concerns forward processes obtained as solutions to more general SDEs.

The approximation is based on the framework presented in [BDM01, Section 5]. For the moment assume that the data (b, σ, g, f) is sufficient in the sense that everything is well-defined; the precise assumptions will be specified below. Recall the notation for the equidistant partition π, the step-size in time $\Delta t_i = h = \frac{T}{n}$, and the discrete forward process X^n introduced in Subsection 1.2. Consider the discretized BSDE, where $Y_{t_n}^n := g(X_T^n)$, and

$$Y_{t_k}^n = g(X_T^n) + h \sum_{m=k}^{n-1} f(t_{m+1}, X_{t_m}^n, Y_{t_m}^n, Z_{t_m}^n) - \sqrt{h} \sum_{m=k}^{n-1} Z_{t_m}^n \xi_{m+1}, \quad 0 \le k < n. \tag{4.7}$$

Equation (4.7) is solved by backward recursion in a similar manner as (4.5). After having found $Z_{t_{i-1}}^n := h^{-1/2}\mathbb{E}[\xi_i Y_{t_i}^n | \mathcal{G}_{i-1}^n]$, where $\mathcal{G}_k^n := \sigma(\xi_1, \xi_2, \dots, \xi_k)$, the quantity $Y_{t_{i-1}}^n$ is obtained by solving the implicit equation

$$Y_{t_{i-1}}^n = Y_{t_i}^n + hf(t_i, X_{t_{i-1}}^n, Y_{t_{i-1}}^n, Z_{t_{i-1}}^n) - \sqrt{h}\xi_i Z_{t_{i-1}}^n. \tag{4.8}$$

For n large enough – assuming that f is Lipschitz continuous – one can prove by a fixed-point argument that equation (4.8) has a unique $\mathcal{G}_k^n$-measurable solution $Y_{t_k}^n$. The solution (X^n, Y^n, Z^n) is then extended into continuous time by letting $X_t^n = X_{t_k}^n$, $t \in [t_k, t_{k+1}), 0 \le k \le n-1$ (similarly for Y^n and Z^n).

The key idea behind the approach adopted in [C]–[D] is that the binary random walk associated to (4.7) is taken to be the *Skorokhod version* of W^n, i.e. from now on $(W_{t_k}^n)_{k\ge0}$ will be identified with $(W_{\tau_k})_{k\ge0}$ defined in Subsection 1.2. In particular, $\sqrt{h}\xi_k := W_{\tau_k} - W_{\tau_{k-1}}$ and $\mathcal{G}_k^n := \mathcal{F}_{\tau_k}$.

Assumption 4.1. *The terminal function $g : \mathbb{R}\to\mathbb{R}$ and the generator $f : [0,T]\times\mathbb{R}^3\to\mathbb{R}$, $(t,x,y,z) \mapsto f(t,x,y,z)$ satisfy the following conditions:*

(A_g) *g is locally α-Hölder continuous in the sense for some constants $p_0 \geq 0$, $C_g > 0$,*

$$|g(x) - g(y)| \leq C_g(1 + |x|^{p_0} + |y|^{p_0})|x - y|^{\alpha}, \quad x, y \in \mathbb{R}. \tag{4.9}$$

(A_f) *There exists a constant $L_f > 0$ such that*

$$|f(t,x,y,z) - f(t',x',y',z')| \leq L_f(\sqrt{|t-t'|} + |x - x'| + |y - y'| + |z - z'|).$$

Remark 4.2. *Note that (A_f) implies $K_f := \sup_{t\in[0,T]} |f(t,0,0,0)| < \infty$.*

For $X := W$ the standard Brownian motion and $X^n := W^n$ the Skorokhod version of the simple symmetric random walk (1.6), we have the following result.

Theorem 4.3 ([C], Theorem 3.1)**.** *Under (A_g) and (A_f), for $n \in \mathbb{N}$ large enough, it holds for all $s \in [t_k, t_{k+1})$ that*

$$\mathbb{E}|Y_s - Y_s^n|^2 \leq \frac{C_0}{n^{\alpha/2}}, \quad \mathbb{E}|Z_s - Z_s^n|^2 \leq \frac{C_0}{n^{\alpha/2}(T-t_k)} + \frac{C_1}{n^{\alpha/2}(T-s)^{\frac{1-\alpha}{2}}}\mathbb{1}_{s\neq t_k}, \tag{4.10}$$

where $C_0, C_1 > 0$ are constants depending on T, p_0, L_f, K_f, C_g, and α. In particular, for each $\beta \in (0,\alpha)$ it holds

$$\sup_{s\in[0,T]} \left(\mathbb{E}|Y_s - Y_s^n|^2\right)^{\frac{1}{2}} \leq \frac{C_0^{1/2}}{n^{\alpha/4}}, \qquad \|Z - Z^n\|_{\mathbb{H}_2} \leq \frac{C(C_0, C_1, \beta)}{n^{\beta/4}}.$$

The result concerning general processes X and X^n given by (1.1) and (1.7) is proved under the following assumptions.

Assumption 4.4. *Let $g : \mathbb{R} \to \mathbb{R}$, $f : [0,T] \times \mathbb{R}^3 \to \mathbb{R}$, $(t,x,y,z) \mapsto f(t,x,y,z)$, and $b, \sigma : [0,T] \times \mathbb{R} \to \mathbb{R}$, $(t,x) \mapsto b(t,x), \sigma(t,x)$ be functions satisfying the following criteria.*

$(\widehat{A_g})$ *$g \in C^2(\mathbb{R})$ and g, g', and g'' satisfy (4.9).*

$(\widehat{A_f})$ *f is continuous and each of the partial derivatives $\partial_x^i\partial_y^j\partial_z^k f$ with $0 \leq i+j+k \leq 2$ exists as a bounded and continuous function. Moreover, the second-order partial derivatives of f are uniformly Lipschitz continuous w.r.t. the spatial variables. In addition, the function $t \mapsto f(t,x,y,z)$ is $\frac{1}{2}$-Hölder continuous, uniformly in (x,y,z).*

$(\widehat{A_{b,\sigma}})$ *(i) b and σ are bounded and continuous functions having bounded and continuous first and second order partial derivatives w.r.t. the variable x.*

(ii) b and σ are $\frac{1}{2}$-Hölder continuous in t, uniformly in $x \in \mathbb{R}$.

(iii) There is a constant $\delta > 0$ such that $\sigma(t,x) \geq \delta > 0$ holds for each point $(t,x) \in [0,T] \times \mathbb{R}$.

(iv) b_{xx} and σ_{xx} are Lipschitz continuous w.r.t the variable x, uniformly in $t \in [0,T]$. Additionally, b_{xx} and σ_{xx} are γ-Hölder continuous on compact subsets of $[0,T] \times \mathbb{R}$ (for some $\gamma \in (0,1]$) w.r.t. the metric $d((t,x),(s,y)) := \sqrt{|t-s| + |x-y|^2}$.

Theorem 4.5 ([D], Theorem 3.2)**.** *Let $(\widehat{A_g})$, $(\widehat{A_f})$, and $(\widehat{A_{b,\sigma}})$ hold. Then, for $n \in \mathbb{N}$ large enough and for all $s \in [0,T)$,*

$$\|Y_s - Y_s^n\|_{L_2(\mathbb{P}_{0,x})} + \|Z_s - Z_s^n\|_{L_2(\mathbb{P}_{0,x})} \leq \frac{C\hat{\Psi}(x)}{n^{\frac{1}{4}\wedge\frac{\alpha}{2}}}, \tag{4.11}$$

where $\hat{\Psi}(x) := 1 + |x|^{6p_0+8}$, and where the constant C depends on T and p_0 together with the bounds of the functions b, σ, f, g and their derivatives.

4.4. **On the proof and some remarks.** Theorems 4.3 and 4.5 are proved using the following decomposition, where U is either Y or Z:

$$\|U_s - U_s^n\| \leq \|U_s - U_{t_k}\| + \|U_{t_k} - U_{t_k}^n\|, \quad s \in [t_k, t_{k+1}); \tag{4.12}$$

recall $U_s^n = U_{t_k}^n$ on $[t_k, t_{k+1})$. Here $\|\cdot\| = \|\cdot\|_{L_2(\mathbb{P}_{0,0})}$ if $X = W$ and otherwise $\|\cdot\| = \|\cdot\|_{L_2(\mathbb{P}_{0,x})}$.

For the evaluation of the middle term in (4.12), regularity estimates of [GGG12] are applied and specialized into our setting. Partially due to the fact that an improved estimate for $\|Z_s - Z_{t_k}\|$ can be obtained under the stronger assumptions of Theorem 4.5, the estimate (4.11) does not involve singularities as opposed to the estimate (4.10) of Theorem 4.3.

Another important tool applied in [C]–[D] is the following representation of the process Z: It holds that $Z_t = \sigma(t, X_t)u_x(t, X_t)$, where

$$u_x(t, X_t) := \mathbb{E}\Big[g(X_T)N_T^t + \int_s^T f(r, X_r, Y_r, Z_r)N_r^t dr \Big| \mathcal{F}_t\Big], \quad t \in [0,T]. \tag{4.13}$$

This representation was obtained for Lipschitz continuous f and g by J. Ma and J. Zhang [MZ02], and was later generalized for polynomially bounded g by J. Zhang [Zha05]. In (4.13), the *Malliavin weight* N_s^t is given by

$$N_s^t := \frac{1}{s-t}\int_t^s \frac{\nabla X_r}{\nabla X_t}\frac{dW_r}{\sigma(r, X_r)}, \quad t < s \leq T,$$

in terms of the *variational process* $\nabla X = (\nabla X_t)_{t\geq 0}$ associated to the forward SDE

$$\nabla X_t = 1 + \int_0^t b_x(r, X_r)\nabla X_r dr + \int_0^t \sigma_x(r, X_r)\nabla X_r dW_r, \quad t \geq 0.$$

The process ∇X can be seen as the derivative of $X = X^{0,x}$ w.r.t. the initial value $x \in \mathbb{R}$ in the $\mathbb{S}_2$-norm.

Provided that $(X, X^n) = (W, W^n)$, one simply has $\nabla X = 1$, and thus the weight is given by $N_s^t = (W_s - W_t)/(s-t)$. Moreover, the process Z^n solving (4.7) can be written as

$$Z_{t_k}^n = \mathbb{E}\Big[g(W_T^n)\frac{W_T^n - W_{t_k}^n}{T - t_k} + h\sum_{m=k+1}^{n-1} f(t_{m+1}, W_{t_m}^n, Y_{t_m}^n, Z_{t_m}^n)\frac{W_{t_m}^n - W_{t_k}^n}{t_m - t_k}\Big|\mathcal{G}_k^n\Big]. \tag{4.14}$$

The above representations for Z and Z^n play a key role in the proof of Theorem 4.3.

Consider then the case of Theorem 4.5 concerning a general forward process. A representation similar to (4.14) can be derived by letting

$$\hat{Z}^n_{t_k} := \mathbb{E}\left[g(X^n_T)N^{n,t_k}_T + h\sum_{m=k}^{n-1} f(t_{m+1}, X^n_{t_m}, Y^n_{t_m}, Z^n_{t_m})N^{n,t_k}_{t_m}\middle|\mathcal{G}^n_k\right]\sigma(t_{k+1}, X^n_{t_k}),$$

where $N^{n,t_k}_{t_m}$ is the natural extension of the discretized weight $(W^n_{t_m} - W^n_{t_k})/(t_m - t_k)$ given in terms of a discretized variational equation, see [D, Subsection 2.3]. As opposed to the Brownian case, however, the processes Z^n and $\hat{Z}^n$ are not equal in general.

In [BDM01, Proposition 5.1], a representation for the true solution Z^n was proved in terms of a finite-difference equation (FDE) involving Malliavin difference operators. See [BP18] for an in-depth discussion on discrete Malliavin calculus. Roughly speaking, the main difference between the true solution Z^n and the approximative solution $\hat{Z}^n$ is that the former involves *difference quotients* in space of b and σ, while the latter involves *partial derivatives* b_x and σ_x. This gives rise to additional technicalities especially in the case of a non-zero generator. When $f \neq 0$, the convergence $\|Z^n_{t_k} - \hat{Z}^n_{t_k}\| \to 0$ is obtained by exploiting regularity properties of the solution to the FDE. However, one needs to impose $(\widehat{A_g})$ on the terminal function g; for the zero generator case ([D, Proposition 3.1]) it is sufficient that $g \in C^1(\mathbb{R})$ and that g' is locally Hölder continuous.

One final important aspect related to the proof is the evaluation of conditional expectations using the properties of the Skorokhod representation. To illustrate this approach, consider for a Lipschitz continuous function φ the difference

$$\mathbb{E}[\varphi(W_T)|\mathcal{F}_{t_k}] - \mathbb{E}[\varphi(W^n_T)|\mathcal{G}^n_k] = \mathbb{E}[\varphi(W_T)|\mathcal{F}_{t_k}] - \mathbb{E}[\varphi(W_{\tau_n})|\mathcal{F}_{\tau_k}]$$

and let $\tilde{W}$ denote a Brownian motion independent of W defined on the same probability space. Denote by $\tilde{\tau}_1, \tilde{\tau}_2, \ldots$ the stopping times given by (1.8) for which W is replaced by $\tilde{W}$. Then, the strong Markov property and Jensen's inequality yield

$$\begin{aligned}\|\mathbb{E}[\varphi(W_T)|\mathcal{F}_{t_k}] - \mathbb{E}[\varphi(W^n_T)|\mathcal{G}^n_k]\| &\leq \|\tilde{\mathbb{E}}[\varphi(\tilde{W}_{t_{n-k}}+W_{t_k})] - \tilde{\mathbb{E}}[\varphi(\tilde{W}_{\tilde{\tau}_{n-k}}+W_{\tau_k})]\| \\ &\leq L_\varphi\left(\|W_{t_{n-k}}-W_{\tau_{n-k}}\| + \|W_{t_k}-W_{\tau_k}\|\right) \\ &\leq C(L_\varphi, T)h^{\frac{1}{4}},\end{aligned}$$

where the last inequality holds by Proposition 1.2 (*iii*). Similar bounds are derived in [D] in the case of a general forward process X and its discretization X^n.

The following example, closing this section, provides some insight into the sharpness of the rate in (4.10) in the case of smooth data.

Example 4.6. *Suppose that $f = 0$ and $\xi = g(W_T)$, where $g(x) = ax + b$ for some $a \neq 0$, $a, b \in \mathbb{R}$. Even in this very simple case the rate in (4.3) is not better than $n^{-\frac{1}{4}}$: First notice that $(Y_t, Z_t) = (aW_t + b, a)$ and $(Y^n_{t_k}, Z^n_{t_k}) = (aW_{\tau_k} + b, a)$ are the unique solutions to (4.2) and (4.7), respectively. It obviously holds that $\mathbb{E}|Z^n_{t_k} - Z^n_{t_k}|^2 = 0$, but more interestingly,*

$$\mathbb{E}|Y^n_{t_k} - Y_{t_k}|^2 = a^2\mathbb{E}|W_{\tau_k} - W_{t_k}|^2 = a^2\mathbb{E}|\tau_k - t_k| \tag{4.15}$$

(recall Proposition 1.2). Let $S_n := \sum_{i=1}^n(\rho_i - 1)$, where $(\rho_i)_{i=1,2,\ldots}$ are i.i.d. random variables distributed as $\tau_1/h \sim \inf\{t \geq 0 : |W_t| = 1\}$. Since $\mathbb{E}[\rho_i] = 1$ and $Var(\rho_i) =$

2/3, the sequence $n^{-1/2}S_n$ *converges in distribution to* $W_{2/3}$ *by the central limit theorem. In addition, using truncation one can show that*

$$\frac{\mathbb{E}|S_n|}{\sqrt{n}} \xrightarrow{n\to\infty} \mathbb{E}|W_{2/3}| = \frac{2}{\sqrt{3\pi}}.$$

Suppose then that $k = k_n$ *depends on* n *in such a way that* $\frac{kT}{n} \xrightarrow{n\to\infty} t_0 \in (0,T]$. *Then, for* n *large enough, there exists a constant* $C(T) > 0$ *such that*

$$\mathbb{E}|\tau_{k_n} - t_{k_n}| = \frac{\sqrt{t_{k_n}T}}{\sqrt{n}} \frac{\mathbb{E}|S_{k_n}|}{\sqrt{k_n}} \geq \frac{C(T)}{\sqrt{n}} \frac{\mathbb{E}|S_{k_n}|}{\sqrt{k_n}}. \tag{4.16}$$

Consequently, by (4.15), (4.16), *and* (1.9),

$$0 < \liminf_{n\to\infty} n^{1/2}\mathbb{E}|Y^n_{t_{k_n}} - Y_{t_{k_n}}|^2 \leq \limsup_{n\to\infty} n^{1/2}\mathbb{E}|Y^n_{t_{k_n}} - Y_{t_{k_n}}|^2 < \infty.$$

References

[Abu02] M. Abundo. Some conditional crossing results of Brownian motion over a piecewise-linear boundary. *Statist. Probab. Lett.*, 58(2):131—145, 2002.

[AKU16] S. Ankirchner, T. Kruse, and M. Urusov. Numerical approximation of irregular SDEs via Skorokhod embeddings. *J. Math. Anal. Appl.*, 440(2):692–715, 2016.

[Avi09] R. Avikainen. On generalized bounded variation and approximation of SDEs. Preprint, 2009. `http://www.math.jyu.fi/research/pspdf/383.pdf`

[Bal97] V. Bally. Approximation scheme for solutions of BSDE. In Backward stochastic differential equations, Pitman Res. Notes Math. Ser., 364:177–191, 1997.

[BT96] V. Bally and D. Talay. The law of the Euler scheme for stochastic differential equations. I. Convergence rate of the distribution function. *Probab. Theory Related Fields*, 104(1):43–60, 1996.

[BO99] L. Beghin and E. Orsingher. On the maximum of the generalized Brownian bridge. *Lith. Math. J.*, 39(2):157–167, 1999.

[BD07] C. Bender and R. Denk. A forward scheme for backward SDEs. *Stochastic Process. Appl.*, 117(12):1793–1812, 2007.

[BP18] C. Bender and P. Parczewski. Discretizing Malliavin calculus. *Stochastic Process. Appl.*, 128(8):2489–2537, 2018.

[BPY01] Ph. Biane, J. Pitman, and M. Yor. Probability laws related to the Jacobi theta and Riemann zeta functions. *Bull. Amer. Math. Soc. (N.S.)*, 38(4):435–465, 2001.

[Bil79] P. Billingsley. Probability and Measure. Wiley Series in Probability and Mathematical Statistics, John Wiley & Sons, 1979.

[BS15] A. N. Borodin and P. Salminen. Handbook of Brownian Motion - Facts and Formulae. Second edition, second printing. Probability and its Applications, Birkhäuser Verlag, 2015.

[BDM01] P. Briand, B. Delyon, and J. Memin. Donsker-Type theorem for BSDEs. *Electron. Comm. Probab.* 6:1–14, 2001.

[BT04] 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.

[BJ08] Z. A. Burq and O. D. Jones. Simulation of Brownian motion at first-passage times. *Math. Comput. Simulation*, 77(1):64–71, 2008.

[CU11] L. Chaumont and G. Uribe Bravo. Markovian bridges: weak continuity and pathwise constructions. *Ann. Probab.*, 39(2):609–647, 2011.

[Che97a] D. Chevance. Résolution numériques des équations différentielles stochastiques rétrogrades. Ph.D. thesis, 1997.

[Che97b] D. Chevance. Numerical methods for backward stochastic differential equations. In Numerical methods in Finance, Cambridge Univ. Press, 232–244, 1997.

[DK05] H. Dong and N. V. Krylov. Rate of convergence of finite-difference approximations for degenerate linear parabolic equations with C^1 and C^2 coefficients. *Electron. J. Differ. Equ.*, 2005(102):1–25, 2005.

[Doo49] J. L. Doob. Heuristic approach to the Kolmogorov-Smirnov theorems. *Ann. Math. Stat.*, 20(3):393–403, 1949.

[FPY92] P. Fitzsimmons, J. Pitman, and M. Yor. Markovian bridges: construction, Palm interpretation, and splicing. In Seminar on Stochastic Processes, 1992, *Progr. Probab*, 33:101–134, Birkhäuser, 1993.

[GGG12] 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.

[HZ00] S. Heston and G. Zhou. On the rate of convergence of discrete-time contingent claims. *Math. Finance*, 10(1):53–75, 2000.

[IM65] K. Itô and H. P. McKean, Jr. Diffusion processes and their sample paths. Second, corrected printing. Die Grundlehren der Mathematischen Wissenschaften 125, Berlin-New York, 1974.

[JY53] M. L. Juncosa and D. M. Young. On the order of convergence of solutions of a difference equation to a solution of the diffusion equation, *SIAM J. Appl. Math.*, 1(2):111–135, 1953.

[KP99] P. E. Kloeden and E. Platen. Numerical solution of stochastic differential equations. Corrected third printing. Probability and its Applications, Birkhäuser Verlag, 1999.

[Kni69] F. B. Knight. Brownian local times and taboo processes. *Trans. Amer. Math. Soc.* 143:173–185, 1969.

[Kol33] A. N. Kolmogorov. Sulla determinazione empirica delle leggi di probabilita. *Giorn. 1st. Ital. Attuari*, 4:1–11, 1933.

[LR00] D. Lamberton and L. C. G. Rogers. Optimal stopping and embedding. *J. Appl. Probab.*, 37(4):1143–1148, 2000.

[Lin07] L. Lindhagen. Finite difference equations and convergence rates in the central limit theorem. *Probab. Math. Statist.*, 27(2):153–166, 2007.

[Luo17] A. Luoto. Time-dependent weak rate of convergence for functions of generalized bounded variation. Preprint, 2017. arXiv:1609.05768v3.

[MPSMT02] J. Ma, P. Protter, J. San Mart´in, and S. Torres. Numerical method for backward stochastic differential equations. *Ann. Appl. Probab.*, 12(1):302–316, 2002.

[MZ02] J. Ma and J. Zhang. Representation theorems for backward stochastic differential equations. *Ann. Appl. Probab.*, 12(4):1390–1418, 2002.

[PP90] E. Pardoux and S. Peng. Adapted solution of a backward stochastic differential equation. *Systems Control Lett.*, 14(1):55–61, 1990.

[PP92] E. Pardoux and S. Peng. Backward stochastic differential equations and quasilinear parabolic partial differential equations. In Stochastic partial differential equations and their applications (Charlotte, NC, 1991). *Lect. Notes Control Inf. Sci.*, 176:200-217, Springer, 1992.

[PX11] S. Peng and M. Xu. Numerical algorithms for backward stochastic differential equation with 1-D Brownian motion: Convergence and simulations. *ESAIM Math. Model. Numer. Anal.*, 45(2):335–460, 2011.

[PW01] K. Pötzelberger and L. Wang. Boundary crossing probability for Brownian motion. *J. Appl. Prob.*, 38(1):152–164, 2001.

[Rey72] A. C. Reynolds. Convergent finite difference schemes for nonlinear parabolic equations. *SIAM J. Numer. Anal.*, 9(4):523–533, 1972.

[RS97] L. C. G. Rogers and E. J. Stapleton. Fast accurate binomial pricing. *Finance Stochast.*, 2(1):3–17, 1997.

[Rud74] W. Rudin. Real and complex analysis. Second edition. McGraw-Hill Series in Higher Mathematics, McGraw-Hill Book Co., 1974.

[Sal97] P. Salminen. On the last exit decomposition of linear diffusions. *Studia Sci. Math. Hungar.*, 33(1):251–262, 1997.

[SY11] P. Salminen and M. Yor. On hitting times of affine boundaries by reflecting Brownian motion and Bessel processes. *Period. Math. Hungar.*, 62(1):75–101, 2011.

[SP12] R. Schilling and L. Partzsch. Brownian Motion - An Introduction to Stochastic Processes. Second edition, De Gryuter, 2012.

[Sko65] A. V. Skorokhod. Studies in the theory of random processes. Translated from the Russian by Scripta Technica, Inc. Addison-Wesley Publishing Co., 1965.

[SS04] B. Székely and T. Szabados. Strong approximation of continuous local martingales by simple random walks. *Studia Sci. Math. Hungar*, 41(1):101–126, 2004.

[Wal03] J. B. Walsh. The rate of convergence of the binomial tree scheme. *Finance Stochast.*, 7(3):337–36, 2003.

[Zha04] J. Zhang. A numerical scheme for BSDEs. *Ann. Appl. Probab.*, 14(1):459–488, 2004.

[Zha05] J. Zhang. Representation of solutions to BSDEs associated with a degenerate FSDE. *Ann. Appl. Probab.*, 15(3):1798–1831, 2005.

[**A**]

TIME-DEPENDENT WEAK RATE OF CONVERGENCE FOR FUNCTIONS OF GENERALIZED BOUNDED VARIATION

A. LUOTO

Time-dependent weak rate of convergence for functions of generalized bounded variation

Antti Luoto

Abstract

Let W denote the Brownian motion. For any exponentially bounded Borel function g the function u defined by $u(t,x) = \mathbb{E}[g(x+\sigma W_{T-t})]$ is the stochastic solution of the backward heat equation with terminal condition g. Let $u^n(t,x)$ denote the according approximation produced by a simple symmetric random walk with steps $\pm\sigma\sqrt{T/n}$ where $\sigma > 0$. This paper is concerned with the rate of convergence of $u^n(t,x)$ to $u(t,x)$, and the behavior of the error $u^n(t,x) - u(t,x)$ as t tends to T. The terminal condition g is assumed to have bounded variation on compact intervals, or to be locally Hölder continuous.

Keywords: approximation using simple random walk, weak rate of convergence, finite difference approximation of the heat equation.
Mathematics Subject Classification (2010): Primary: 41A25, 65M15; Secondary: 35K05, 60G50.

Contents

1 Introduction

The objective of this paper is to study the rate of convergence of a finite-difference approximation scheme for the backward heat equation with an irregular terminal condition. Convergence rates of finite-difference schemes for parabolic boundary value problems have been studied during the past decades (see e.g. [3], [6], [8], [11] and [14]) with varying assumptions on the regularity of the initial/terminal condition, the domain of the solution, properties of the possible boundary data etc. Naturally, several techniques have been proposed in order to study the convergence. Our approach is probabilistic: The solution of the PDE is represented in terms of Brownian motion, and the approximation scheme is realized using an appropriately scaled sequence of simple symmetric random walks in the same probability space, in the spirit of Donsker's theorem. This method produces error bounds which are not uniform over the time-nets under consideration, and hence the time-dependence of the error is of particular interest here.

To explain our setting in more detail, fix a finite time horizon $T > 0$, a constant $\sigma > 0$, and consider the backward heat equation

$$\frac{\partial}{\partial t}u + \frac{\sigma^2}{2}\frac{\partial^2}{\partial x^2}u = 0, \quad (t,x) \in [0,T) \times \mathbb{R}, \qquad u(T,x) = g(x), \quad x \in \mathbb{R}. \tag{1.1}$$

University of Jyvaskyla, Department of Mathematics and Statistics, P. O. Box 35, FI-40014 University of Jyvaskyla
antti.k.luoto@student.jyu.com

The terminal condition $g : \mathbb{R} \to \mathbb{R}$ is assumed to belong to the class GBV_{exp} consisting of exponentially bounded functions that have bounded variation on compact intervals (see Definition 2.3 for the precise description of GBV_{exp}). The stochastic solution of the problem (1.1) is given by

$$u(t,x) := \mathbb{E}[g(\sigma W_T)|\sigma W_t = x] = \mathbb{E}[g(x+\sigma W_{T-t})], \quad (t,x) \in [0,T] \times \mathbb{R}, \tag{1.2}$$

where $(W_t)_{t\geq 0}$ denotes the standard Brownian motion. To approximate the solution (1.2), we proceed as follows. Given an even integer $n \in 2\mathbb{N}$, a level $z_0 \in \mathbb{R}$, and time and space step sizes δ and h, define

$$\mathcal{T}^n := \left\{ t_k^n := 2k\delta \,\middle|\, 0 \leq k \leq \tfrac{n}{2}, k \in \mathbb{Z} \right\}, \quad \mathcal{S}_{z_0}^n := \left\{ z_0 + 2mh \,\middle|\, m \in \mathbb{Z} \right\}.$$

The finite-difference scheme we will consider is given by the following system of equations defined on grids $\mathcal{G}_{z_0}^n := \mathcal{T}^n \times \mathcal{S}_{z_0}^n \subset [0,T] \times \mathbb{R}$,

$$\begin{cases} \dfrac{u^n(t_k^n,x) - u^n(t_{k-1}^n,x)}{t_k^n - t_{k-1}^n} + \dfrac{\sigma^2}{2} \dfrac{u^n(t_k^n,x+2h) - 2u^n(t_k^n,x) + u^n(t_k^n,x-2h)}{(2h)^2} = 0, \\ u^n(T,\,\cdot\,) = g. \end{cases} \tag{1.3}$$

Letting $\delta := \frac{T}{n}$ and $h := \sigma\sqrt{\frac{T}{n}}$, system (1.3) can be rewritten in an equivalent form as

$$\begin{cases} u^n(t_{k-1}^n,x) = \frac{1}{4}\left[u^n(t_k^n,x+2h) + 2u^n(t_k^n,x) + u^n(t_k^n,x-2h)\right], \\ u^n(T,\,\cdot\,) = g. \end{cases} \tag{1.4}$$

This scheme is explicit: Given the set of terminal values $\left\{g(x) \,\middle|\, x \in \mathcal{S}_{z_0}^n\right\}$, the solution u^n of (1.4) is uniquely determined by a backward recursion. We extend the function u^n in continuous time by letting

$$u^n(t,x) := u^n(t_k^n,x) \quad \text{for} \quad t \in [t_k^n, t_{k+1}^n),\ 0 \leq k < \tfrac{n}{2}, \tag{1.5}$$

and consider the error $\varepsilon_n(t,x)$ on $(t,x) \in [0,T) \times \mathcal{S}_{z_0}^n$, which is given by

$$\varepsilon_n(t,x) := u^n(t,x) - u(t,x). \tag{1.6}$$

The main result of this paper, Theorem 2.4 (A) states that for some constant $C > 0$ depending only on g,

$$|\varepsilon_n(t,x)| \leq \frac{C\psi(x)}{\sqrt{n(T-t)}} \mathbb{1}_{\{t \notin \mathcal{T}^n\}} + \frac{C\psi(x)}{\sqrt{n(T-t_k^n)}} \mathbb{1}_{[t_k^n, t_{k+1}^n)}(t), \quad (t,x) \in [0,T) \times \mathcal{S}_{z_0}^n, \tag{1.7}$$

where $\psi(x) = \psi(|x|, g, \sigma, T) > 0$ depends on the properties of g and will be given explicitly later.

Inequality (1.7) suggests that the convergence is not uniform in (t,x). However, if we consider uniform convergence on any compact subset of $[0,T) \times \mathbb{R}$, the rate is at least $n^{-1/2}$, and it will be shown in Subsection 4.4 that this rate is also sharp.

Already in 1953, Juncosa & Young [6] considered a finite difference approximation of the forward heat equation on a semi-infinite strip $[0,\infty) \times [0,1]$, where the initial condition was assumed to have bounded variation. Using Fourier methods, they proved in [6, Theorem 7.1] that the error is $O(n^{-1/2})$ uniformly on $[t,\infty) \times [0,1]$ for any fixed $t > 0$, but did not study the blow-up of the error as $t \downarrow 0$. Notice that the right-hand side of (1.7) undergoes a blow-up as $t \uparrow T$. The order of the singularity is even worse for time instants t not belonging to the lattice $\mathcal{T}^n$ due to the possible discontinuities of g. On the other hand, one observes that the order remains unchanged if the terminal condition g is Hölder continuous. Indeed, suppose that g belongs to the class $C_{\text{exp}}^{0,\alpha}$ (see Definition 2.1), which consists of exponentially bounded, locally α-Hölder continuous functions. By Theorem 2.4 (B), there exists a constant $C > 0$ depending only on g such that

$$|\varepsilon_n(t,x)| \leq \frac{C\psi(x)}{n^{\frac{\alpha}{2}}(T-t_k^n)^{\frac{\alpha}{2}}} \mathbb{1}_{[t_k^n, t_{k+1}^n)}(t), \quad (t,x) \in [0,T) \times \mathcal{S}_{z_0}^n, \tag{1.8}$$

where the function $\psi(x) = \psi(|x|, g, \sigma, T) > 0$ plays a similar role as in (1.7).

Recently, Dong & Krylov (2005) [3] considered the convergence of a finite-difference scheme for a very general parabolic PDE. By specializing their result [3, Theorem 2.12] to the setting of this paper, the error is seen to converge *uniformly* in (t,x) with rate $n^{-1/4}$ for a bounded and Lipschitz continuous terminal condition, in contrast to the time-dependent rate $n^{-1/2}$ implied by (1.8). In fact, an analogous uniform rate $n^{-\alpha/4}$ can be shown for the class $C^{0,\alpha}_{\exp}$ in our setting; the proof is sketched in Remark 2.6.

In this paper, the main results are derived using the following probabilistic approach. Let $(\xi_i)_{i=1,2,\dots}$ be a sequence of i.i.d. Rademacher random variables, and define

$$u^n(t,x) := \mathbb{E}\big[g\big(x+\sigma W^n_{T-t}\big)\big], \quad (t,x) \in [0,T] \times \mathbb{R}, \tag{1.9}$$

where $(W^n_t)_{t\in[0,T]}$ is the random walk given by

$$W^n_t := \sqrt{\tfrac{T}{n}} \sum_{i=1}^{2\lceil \frac{t}{2T/n} \rceil} \xi_i, \quad t \in [0,T] \tag{1.10}$$

($\lceil \cdot \rceil$ denotes the ceiling function). The key observation is that the function u^n in (1.9), when restricted to $\mathcal{G}^n_{z_0}$, is the unique solution of (1.4) for every $z_0 \in \mathbb{R}$; (1.5) also holds for this u^n by definition. Moreover, since the random walk $(W^n_t)_{t\in[0,T]}$ influences the value of u^n only through its distribution, we may consider a special setting where the random variables $\xi_1, \xi_2, \dots$ are chosen in a suitable way. Defining these variables as the values of the Brownian motion $(W_t)_{t\geq 0}$ sampled at certain stopping times (see Subsection 2.1) enables us to apply techniques from stochastic analysis for the estimation of the error (1.6).

The above procedure was used in J. B. Walsh [13] (2003) (cf. Rogers & Stapleton (1997) [12]) in relation to a problem arising in mathematical finance. More precisely, the weak rate of convergence of European option prices given by the binomial tree scheme (Cox-Ross-Rubinstein model) to prices implied by the Black-Scholes model is analyzed (cf. Heston & Zhou (2000) [5]). A detailed error expansion is presented in [13, Theorem 4.3] for terminal conditions belonging to a certain class of piecewise C^2 functions. Using similar ideas, we complement this result by considering more irregular functions and taking into account the time-dependence. It is argued in [13, Sections 7 and 12] that the rate remains unaffected if the geometric Brownian motion is replaced with a Brownian motion, and the binomial tree is replaced with a random walk. It seems plausible that also our time-dependent results in the Brownian setting can be transferred into the geometric setting with essentially the same upper bounds.

It should be mentioned here that the proof of (1.7) uses the general representation (2.7) for functions of generalized bounded variation, which allows us to estimate the error (1.6) in an explicit manner. This type of function classes of generalized bounded variation were studied first in Avikainen (2009) [1].

The paper is organized as follows. In Section 2 we introduce the notation, recall the construction of a simple random walk using first hitting times of the Brownian motion, and formulate the main result Theorem 2.4. Using the sequence of stopping times, we split the error (1.6) into three parts, which we refer to as the *adjustment error*, the *local error*, and the *global error*. The *adjustment error* is a consequence of the fact that the approximation $u^n(t,x)$ is constant in t on intervals of length $\frac{2T}{n}$, while $t \mapsto u(t,x)$ is continuous. The remaining two parts of the error appear because the construction of the simple random walk uses the Brownian motion sampled at a stopping time which can be larger of smaller than $T-t$, and for $u(t,x)$ we use W_{T-t}. The *local error* is influenced by the smoothness properties of the terminal condition g, while for the *global error* only integrability properties of g are needed. In Section 3, estimates for the adjustment error are computed. Section 4 treats the local error and follows in many places the ideas and the machinery of J. B. Walsh [13]. We also apply some results of [4] related to the first exit times of Brownian bridges to study the sharpness of the rate and to derive explicit upper estimates in Section 6. In Section 5, the global error is treated for exponentially bounded Borel functions, and our approach is similar to that of [13]. Section 6 contains a collection of moment estimates and tail behaviors of random times appearing in the description of the local and the global error. Again, it was possible to adjust methods from [13] to our setting.

2 The setting and the main result

2.1 Notation related to the random walk

Consider a standard Brownian motion $(W_t)_{t\geq 0}$ on a stochastic basis $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t\geq 0})$, where $(\mathcal{F}_t)_{t\geq 0}$ stands for the natural filtration of $(W_t)_{t\geq 0}$. We also let $(X_t)_{t\geq 0} := (\sigma W_t)_{t\geq 0}$, where $\sigma > 0$ is a given constant. By $\tau_{(-h,h)}$ we denote the first exit time of the process $(X_t)_{t\geq 0}$ from the open interval $(-h, h)$,

$$\tau_{(-h,h)} := \inf\{t \geq 0 : |X_t| = h\} = \inf\{t \geq 0 : |W_t| = h/\sigma\}, \quad h > 0.$$

The random variable $\tau_{(-h,h)}$ is a $(\mathcal{F}_t)_{t\geq 0}$-stopping time and its moment-generating function is given by

$$\mathbb{E}\big[e^{\lambda\tau_{(-h,h)}}\big] = \begin{cases} \cosh(h\sqrt{2|\lambda|}/\sigma)^{-1}, & \lambda \leq 0, \\ \cos(h\sqrt{2\lambda}/\sigma)^{-1}, & \lambda \in (0, \frac{\pi^2}{8}\frac{\sigma^2}{h^2}). \end{cases} \tag{2.1}$$

It follows that the exit time $\tau_{(-h,h)}$ has finite moments of all orders, and for every $K \in \mathbb{N}$ there exists a constant $C_K > 0$ such that

$$\mathbb{E}\big[\tau_{(-h,h)}^K\big] = C_K (h/\sigma)^{2K}. \tag{2.2}$$

In particular, $C_1 = 1$ and $C_2 = 5/3$. For relations (2.1) and (2.2), see [13, Proposition 11.1].

In order to represent the error (1.6), we construct a random walk on the space $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t\geq 0})$. Following [13], we define

$$\tau_0 := 0 \quad \text{and} \quad \tau_k = \tau_k(h) := \inf\big\{t \geq \tau_{k-1} : \big|X_t - X_{\tau_{k-1}}\big| = h\big\} \tag{2.3}$$

recursively for $k = 1, 2, \ldots$. Then τ_k is a $\mathbb{P}$-a.s. finite $(\mathcal{F}_t)_{t\geq 0}$-stopping time for all $k \geq 0$, and the process $(X_{\tau_k})_{k=0,1,\ldots}$ is a symmetric simple random walk on $\mathbb{Z}^h := \{mh : m \in \mathbb{Z}\}$. For every integer $k \geq 1$, we also let

$$\Delta\tau_k := \tau_k - \tau_{k-1} \quad \text{and} \quad \Delta X_{\tau_k} := X_{\tau_k} - X_{\tau_{k-1}}.$$

The strong Markov property of $(X_t)_{t\geq 0}$ implies that $(\Delta\tau_k, \Delta X_{\tau_k})_{k=1,2,\ldots}$ is an i.i.d. process such that, for each $k \geq 1$, we have $\mathbb{P}(\Delta X_{\tau_k} = \pm\, h) = 1/2$,

$$(\Delta\tau_k, \Delta X_{\tau_k}) \stackrel{d}{=} (\tau_{(-h,h)}, X_{\tau_{(-h,h)}}), \quad \text{and} \quad (\Delta\tau_k, \Delta X_{\tau_k}) \text{ is independent of } \mathcal{F}_{\tau_{k-1}+}.$$

Moreover, as shown in [12, Proposition 1], the increments ΔX_{τ_1} and $\Delta\tau_1$ are independent. Consequently, the *processes* $(\Delta\tau_k)_{k=1,2,\ldots}$ and $(\Delta X_{\tau_k})_{k=1,2,\ldots}$ are independent (see also [13, Proposition 11.1] and [7, Proposition 2.4]).

We deduce, in particular, that for all $N \geq 1$ the random variable X_{τ_N} is distributed as $h\sum_{k=1}^N \xi_k$, where $(\xi_k)_{k=1,2,\ldots}$ is an i.i.d. sequence of Rademacher random variables. Therefore, for W^n_{T-t} defined in (1.10), we have the equality in law

$$X_{\tau_N} \stackrel{d}{=} \sigma W^n_{T-t} \quad \text{provided that } (h, N) = \big(\sigma\sqrt{\tfrac{T}{n}}, 2\lceil\tfrac{T-t}{2T/n}\rceil\big).$$

Note that in this case the sequence of stopping times $(\tau_k)_{k=0,1,\ldots}$ (2.3) depends on n via $h = h(n)$.

The error (1.6) will be split into three parts, where each of these parts will take into account different properties of the given function g. For this purpose, let us introduce some more notation. For given $n \in 2\mathbb{N}$ and $t \in [0, T)$, we let

$$\theta_n := \frac{n_\theta T}{n}, \quad \text{where} \quad n_\theta := 2\left\lceil\frac{T-t}{2T/n}\right\rceil \in \{2, 4, \ldots, n\}. \tag{2.4}$$

By definition, θ_n is the smallest multiple of $\frac{2T}{n}$ greater than or equal to $T-t$. It is clear that

$$0 \leq \theta_n - (T-t) \leq \frac{2T}{n} \quad \text{and} \quad \theta_n \downarrow T-t \quad \text{as } n \to \infty.$$

The connection between lattice points $t_k^n = \frac{2kT}{n} \in \mathcal{T}^n$ and the time instant $\theta_n \in (0, T]$ is explained by

$$t \in [t_k^n, t_{k+1}^n) \quad \text{if and only if} \quad \theta_n = T - t_k^n, \quad 0 \leq k \leq \tfrac{n}{2}-1. \tag{2.5}$$

2.2 The function classes under consideration

The error (1.6) will be estimated for functions g belonging to the function class $C^{0,\alpha}_{\text{exp}}$ or GBV_{exp} defined below. More information regarding these classes is provided in Subsections 4.5 and A.1, respectively.

Definition 2.1 (The class $C^{0,\alpha}_{\text{exp}}$)**.** Denote by $C^{0,\alpha}_{\text{exp}}$ the class of all functions $g : \mathbb{R} \to \mathbb{R}$ for which there exist constants $A, \beta \geq 0$ such that for all $R > 0$,

$$\sup_{x,y \in [-R,R],\, x \neq y} \frac{|g(x) - g(y)|}{|x - y|^{\alpha}} \leq A e^{\beta R}. \tag{2.6}$$

The function class GBV_{exp} generalizes functions of bounded variation (which are bounded) by allowing exponential growth. For more information, see [1]. Before introducing the class GBV_{exp}, we recall

Definition 2.2 ([1, Definition 3.2])**.** Denote by $\mathcal{M}$ the class of all set functions

$$\mu : \{G \in \mathcal{B}(\mathbb{R}) : G \text{ is bounded}\} \to \mathbb{R}$$

that can be written as a difference of two measures $\mu^1, \mu^2 : \mathcal{B}(\mathbb{R}) \to [0, \infty]$ such that $\mu^1(K), \mu^2(K) < \infty$ for all compact sets $K \in \mathcal{B}(\mathbb{R})$.

Definition 2.3 (The class GBV_{exp})**.** Denote by GBV_{exp} the class of functions $g : \mathbb{R} \to \mathbb{R}$ which can be represented as

$$g(x) = c + \mu([0, x)) - \mu([x, 0)) + \sum_{i=1}^{\infty} \alpha_i \mathbb{1}_{\{x_i\}}(x), \quad x \in \mathbb{R}, \tag{2.7}$$

where $c \in \mathbb{R}$ is a constant, $\mu \in \mathcal{M}$, and $\mathcal{J} = (\alpha_i, x_i)_{i=1,2,\ldots} \subset \mathbb{R}^2$ is a countable set such that $x_i \neq x_j$ whenever $i \neq j$. In addition, we require that for some constant $\beta \geq 0$,

$$\int_{\mathbb{R}} e^{-\beta|x|} d|\mu|(x) + \sum_{i=1}^{\infty} |\alpha_i| e^{-\beta|x_i|} < \infty. \tag{2.8}$$

2.3 The main result

The following theorem is the main result of this paper.

Theorem 2.4. *Let $n \in 2\mathbb{N}$, and let u and u^n be the functions introduced in (1.2) and (1.9).*

(A) *Suppose that $g \in GBV_{\text{exp}}$ is a function given by (2.7) and that $\beta \geq 0$ is as in (2.8). Then, for all $(t, x) \in [0, T) \times \mathbb{R}$,*

$$\begin{aligned}
(i) \quad & |u^n(t, x) - u(t, x)| \leq \frac{C_{\beta,\sigma,T}}{\sqrt{n(T-t)}} e^{\beta|x|}, \quad t \neq t_k^n, \quad 0 \leq k < \tfrac{n}{2}, \\
(ii) \quad & |u^n(t_k^n, x) - u(t_k^n, x)| \leq \frac{C_{\beta,\sigma,T}}{\sqrt{n(T - t_k^n)}} e^{\beta|x|}, \quad 0 \leq k < \tfrac{n}{2},
\end{aligned}$$

where $C_{\beta,\sigma,T} := C(T \vee \sqrt{T}) e^{3\beta^2\sigma^2 T}$ and $C > 0$ is a constant depending only on g.

(B) *Suppose that the function* $g \in C^{0,\alpha}_{exp}$ *and that* $\beta \geq 0$ *is as in* (2.6)*. Then, for all* $(t,x) \in [0,T)\times\mathbb{R}$*,*

$$(iii) \qquad |u^n(t,x) - u(t,x)| \leq \frac{C_{\beta,\sigma,T}}{n^{\frac{\alpha}{2}}(T-t_k^n)^{\frac{\alpha}{2}}} e^{(\beta+1)|x|}, \qquad t \in [t_k^n, t_{k+1}^n),\ 0 \leq k < \tfrac{n}{2},$$

where $C_{\beta,\sigma,T} := (1+T)(2+\sigma)Ce^{4(\beta+1)^2\sigma^2 T}$ *and* $C > 0$ *is a constant depending only on* g*.*

Remark 2.5. Properties of the error bounds in (A) and (B) were already discussed in Section 1. Here we only point out that in general these error bounds grow exponentially as functions of x. A uniform bound w.r.t. x can be shown under additional assumptions: For $g \in GBV_{\text{exp}}$, it is sufficient that g satisfies the condition (2.8) with $\beta = 0$. For $g \in C^{0,\alpha}_{\text{exp}}$, it suffices to assume that g is bounded and satisfies (2.6) with $\beta = 0$.

Proof of Theorem 2.4. Following [13], we define an auxiliary random variable J_n on $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t\geq 0})$ by

$$J_n(\omega) := \inf\{2m \in 2\mathbb{N} : \tau_{2m}(\omega) > \theta_n\}, \tag{2.9}$$

where we assume that the step size related to $(\tau_k)_{k=0,1,\ldots}$ is $h = \sigma\sqrt{\frac{T}{n}}$. By definition, J_n is the index of the first even stopping time $\tau_0, \tau_2, \ldots$ exceeding the value θ_n. It holds that J_n is a stopping time w.r.t. $(\mathcal{F}_{\tau_k})_{k=0,1,\ldots}$. Moreover, τ_{J_n} is a stopping time w.r.t. $(\mathcal{F}_t)_{t\geq 0}$, and both J_n and τ_{J_n} are $\mathbb{P}$-a.s. finite. The error $\varepsilon_n(t,x)$ given by (1.6) is then decomposed as follows:

$$\varepsilon_n(t,x) = \varepsilon_n^{\text{glob}}(t,x) + \varepsilon_n^{\text{loc}}(t,x) + \varepsilon_n^{\text{adj}}(t,x), \tag{2.10}$$

where

$$\varepsilon_n^{\text{glob}}(t,x) := \mathbb{E}[g(x+X_{\tau_{n_\theta}}) - g(x+X_{\tau_{J_n}})], \qquad (\text{"the global error"}) \tag{2.11}$$

$$\varepsilon_n^{\text{loc}}(t,x) := \mathbb{E}[g(x+X_{\tau_{J_n}}) - g(x+X_{\theta_n})], \qquad (\text{"the local error"}) \tag{2.12}$$

$$\varepsilon_n^{\text{adj}}(t,x) := \mathbb{E}\left[g(x+X_{\theta_n}) - g(x+X_{T-t})\right]. \qquad (\text{"the adjustment error"}) \tag{2.13}$$

Assume that $0 \leq k < \frac{n}{2}$ is the integer for which $t \in [t_k^n, t_{k+1}^n)$ holds.

(A): By Remark A.1 (i), there exists a constant $A = A(\beta) \geq 0$ such that $|g(x)| \leq Ae^{\beta|x|}$ for all $x \in \mathbb{R}$. Hence, by Propositions 3.3 and 5.3 and Corollary 4.13, there exists a constant $C > 0$ such that

$$|\varepsilon_n(t,x)| \leq Ce^{\beta|x|+3\beta^2\sigma^2 T}\Big(\frac{T}{n(T-t)}\mathbb{1}_{\{t\neq t_k^n\}} + \frac{\sqrt{T}}{\sqrt{n(T-t_k^n)}} + \frac{T}{n(T-t_k^n)}\Big).$$

We then get the claim in both of the cases $t \in (t_k^n, t_{k+1}^n)$ and $t = t_k^n$: It holds that

$$\frac{T}{n(T-t_k^n)} \leq \frac{T}{n(T-t)} \wedge \frac{\sqrt{T}}{\sqrt{n(T-t_k^n)}} \quad \text{and} \quad \frac{\sqrt{T}}{\sqrt{n(T-t_k^n)}} \leq \frac{T}{\sqrt{n}(T-t)},$$

since $\sqrt{n(T-t_k^n)} \geq \sqrt{2T}$ for all integers $0 \leq k < \frac{n}{2}$.

(B): Given a constant $\delta > 0$, by assumption, we can derive the exponential bound

$$|g(x)| \leq A|x|^\alpha e^{\beta|x|} + |g(0)| \leq Ce^{(\beta+\delta)|x|}, \quad x \in \mathbb{R},$$

for some constant $C > 0$. For simplicity, let us choose $\delta = 1$. Consequently, by Propositions 3.3 and 5.3 (put $b = \beta + \delta$), and Corollary 4.17, we find another constant $\tilde{C} > 0$ such that

$$|\varepsilon_n(t,x)| \leq \tilde{C}e^{(\beta+1)|x|+4(\beta+1)^2\sigma^2 T}\left(\frac{\sigma^\alpha T^{\alpha/2}}{n^{\alpha/2}} + \frac{T}{n(T-t_k^n)}\right).$$

The claim follows, since $\left(\frac{T}{n(T-t_k^n)}\right)^\gamma \leq \left(\frac{T}{n(2T/n)}\right)^\gamma \leq 1$ for all $\gamma \in [0,1]$, and thus

$$\frac{\sigma^\alpha T^{\alpha/2}}{n^{\alpha/2}} + \frac{T}{n(T-t_k^n)} \leq \frac{\sigma^\alpha T^{\alpha/2}}{n^{\alpha/2}} + \left(\frac{T}{n(T-t_k^n)}\right)^{\alpha/2} \leq \frac{(T^{\alpha/2}+\sigma^\alpha T^\alpha)}{n^{\alpha/2}(T-t_k^n)^{\alpha/2}} \leq \frac{(1+T)(2+\sigma)}{n^{\alpha/2}(T-t_k^n)^{\alpha/2}}.$$

□

Remark 2.6. For $g \in C^{0,\alpha}_{\text{exp}}$, there exists a constant $C = C(A,\sigma,T) > 0$ such that for all $x \in \mathbb{R}$,

$$\sup_{t\in[0,T)} |u^n(t,x) - u(t,x)| \leq \frac{C}{n^{\frac{\alpha}{4}}} e^{4\beta|x|+8\beta^2\sigma^2 T}, \tag{2.14}$$

where $A, \beta \geq 0$ are as in (2.6). Hence, we get the uniform rate $n^{-\alpha/4}$ instead of the time-dependent rate $n^{-\alpha/2}$ implied by Theorem 2.4 (B). Note that for $g \in C^{0,\alpha}_{\text{exp}}$, the time-dependence of the error bound in Theorem 2.4 (B) is caused solely by the global error, and it remains unclear whether the associated upper bound (5.9) can be improved using the additional information about the regularity of g.

For the proof of (2.14), notice first that by the Hölder continuity and by Hölder's inequality,

$$|u^n(t,x) - u(t,x)| \leq A\sigma^\alpha \left(\mathbb{E}e^{q\beta|x+\sigma W_{T-t}|+q\beta|x+\sigma W_{\tau_{n_\theta}}|}\right)^{1/q} \left(\mathbb{E}\left|W_{T-t} - W_{\tau_{n_\theta}}\right|^{p\alpha}\right)^{1/p},$$

where $p := \frac{2}{\alpha}$ and $q := \frac{p}{p-1}$. To proceed, apply Lemma 5.1 (i) and the fact that for some $C(T) > 0$,

$$\mathbb{E}\left|W_{T-t} - W_{\tau_{n_\theta}}\right|^2 = \mathbb{E}\left|(T-t) - \tau_{n_\theta}\right| \leq C(T)n^{-1/2},$$

which follows from Itô's isometry and a slight generalization of [13, Proposition 11.1 (iv)].

3 The adjustment error

In this section we derive an upper bound for the adjustment error (2.13) for exponentially bounded Borel functions and for functions belonging to the class $C^{0,\alpha}_{\text{exp}}$, $\alpha \in (0,1]$.

Definition 3.1 (The class $\mathcal{B}_{\text{exp}}$)**.** A function $g : \mathbb{R} \to \mathbb{R}$ is said to be exponentially bounded, if there exist constants $A, b \geq 0$ such that

$$|g(x)| \leq Ae^{b|x|} \quad \text{for all } x \in \mathbb{R}. \tag{3.1}$$

The class of all Borel functions with the above property will be denoted by $\mathcal{B}_{\text{exp}}$.

Remark 3.2. By definition, $GBV_{\text{exp}} \subset \mathcal{B}_{\text{exp}}$ (see Remark A.1) and $C^{0,\alpha}_{\text{exp}} \subset \mathcal{B}_{\text{exp}}$ (see Subsection 4.5).

Proposition 3.3. *Let* $n \in 2\mathbb{N}$.

(i) *Let* $g \in \mathcal{B}_{exp}$ *and let* $A, b \geq 0$ *be as in (3.1). Then, for all* $(t_0, x_0) \in [0,T)\times\mathbb{R}$,

$$\left|\varepsilon_n^{adj}(t_0,x_0)\right| \leq \frac{8AT}{n(T-t_0)} e^{b|x_0|+b^2\sigma^2 T} \mathbb{1}_{\{t_0 \neq t_k^n \,\forall\, 0<k<\frac{n}{2}\}}.$$

(ii) *Let* $g \in C^{0,\alpha}_{exp}$ *and let* $A, \beta \geq 0$ *be as in (2.6). Then, for all* $(t_0, x_0) \in [0,T)\times\mathbb{R}$,

$$\left|\varepsilon_n^{adj}(t_0,x_0)\right| \leq \frac{2A\sigma^\alpha T^{\alpha/2}}{n^{\alpha/2}} e^{\beta|x_0|+4\beta^2\sigma^2 T} \mathbb{1}_{\{t_0 \neq t_k^n \,\forall\, 0<k<\frac{n}{2}\}}.$$

Proof. (i): Denote by p_t the density of $X_t = \sigma W_t$ for $t > 0$, and consider the function

$$u(t, x_0) = \mathbb{E}[g(x_0{+}X_{T-t})] = \int_{\mathbb{R}} g(x_0{+}y)p_{T-t}(y)dy, \quad 0 \le t < T.$$

Since $g \in \mathcal{B}_{\text{exp}}$, we can use differentiation under the integral sign to show that

$$\frac{\partial}{\partial t}u(t, x_0) = \int_{\mathbb{R}} g(x_0{+}y)\frac{\partial}{\partial t}p_{T-t}(y)dy = \int_{\mathbb{R}} \frac{g(x_0{+}y)p_{T-t}(y)}{2(T{-}t)}\left(1 - \frac{y^2}{\sigma^2(T{-}t)}\right)dy. \tag{3.2}$$

Fix $n \in 2\mathbb{N}$ and suppose that $t_k^n = \frac{2kT}{n}$ is the lattice point such that $t_0 \in [t_k^n, t_{k+1}^n)$. If $t_0 = t_k^n$, (2.5) implies that $\theta_n = T{-}t_0$, and thus $\varepsilon_n^{\text{adj}}(t_0, x_0) = 0$ by (2.13). For $t_0 \in (t_k^n, t_{k+1}^n)$, by the mean value theorem and (3.2), there exists some $\eta \in (t_k^n, t_0)$ such that

$$|u(t_k^n, x_0) - u(t_0, x_0)| \le \frac{t_0 - t_k^n}{2(T - t_0)}C_{T-\eta} \le \frac{T}{n(T - t_0)} \sup_{r \in (T-t_0, T-t_k^n)} C_r, \tag{3.3}$$

where $C_r := \big|\int_{\mathbb{R}} g(x_0{+}y)p_r(y)\big(1{-}\frac{y^2}{\sigma^2 r}\big)dy\big|$, $r > 0$. Let Z be a standard normal random variable. Since $g \in \mathcal{B}_{\text{exp}}$, it holds for all $r \in (0, T]$ that

$$\begin{aligned} C_r &\le Ae^{b|x_0|}\mathbb{E}\left[e^{b|X_r|}\right] + Ae^{b|x_0|}\mathbb{E}\left[e^{b|X_r|}\left(\tfrac{X_r}{\sigma\sqrt{r}}\right)^2\right] \\ &\le 2Ae^{b|x_0|}\left(\mathbb{E}\left[e^{b\sigma\sqrt{r}Z}\right] + \mathbb{E}\left[Z^2 e^{b\sigma\sqrt{r}Z}\right]\right) \\ &= 2Ae^{b|x_0|}\left(e^{\frac{1}{2}b^2\sigma^2 r} + e^{\frac{1}{2}b^2\sigma^2 r}\mathbb{E}\left[\left(Z + b\sigma\sqrt{r}\right)^2\right]\right) \\ &\le 2A(2 + b^2\sigma^2 T)e^{b|x_0| + \frac{1}{2}b^2\sigma^2 T} \\ &\le 8Ae^{b|x_0| + b^2\sigma^2 T}. \end{aligned} \tag{3.4}$$

Since $\left|\varepsilon_n^{\text{adj}}(t_0, x_0)\right| = |u(t_k^n, x_0) - u(t_0, x_0)|$, (3.3) and (3.4) imply the claim.

(ii): Let $0 \le k < \frac{n}{2}$ be such that $t_0 \in (t_k^n, t_{k+1}^n)$ holds; the case $t_0 = t_k^n$ follows from (2.5) and (2.13). Hölder's inequality implies that

$$\begin{aligned} \left|\varepsilon_n^{\text{adj}}(t_0, x_0)\right| &\le \mathbb{E}\left|g(x_0{+}X_{T-t_k^n}) - g(x_0{+}X_{T-t_0})\right| \\ &\le A\mathbb{E}\left[e^{\beta(|x_0|+|X_{T-t_k^n}|+|X_{T-t_0}|)}\left|X_{T-t_k^n} - X_{T-t_0}\right|^\alpha\right] \\ &\le A\left(\mathbb{E}\left[e^{q\beta(|x_0|+|X_{T-t_k^n}|+|X_{T-t_0}|)}\right]\right)^{1/q}\left(\mathbb{E}\left|X_{T-t_k^n} - X_{T-t_0}\right|^{p\alpha}\right)^{1/p}, \end{aligned} \tag{3.5}$$

for some $p, q \in (1, \infty)$ with $\frac{1}{p} + \frac{1}{q} = 1$. The choice $p = \frac{2}{\alpha}$, $q = \frac{2}{2-\alpha}$ and the fact $|t_0 - t_k^n| \le \frac{2T}{n}$ yield

$$\left(\mathbb{E}\left|X_{T-t_k^n}{-}X_{T-t_0}\right|^{p\alpha}\right)^{1/p} \le \left(\sigma^2\mathbb{E}\left|W_{T-t_k^n}{-}W_{T-t_0}\right|^2\right)^{\alpha/2} \le \frac{\sigma^\alpha 2^{\alpha/2}T^{\alpha/2}}{n^{\alpha/2}}. \tag{3.6}$$

Moreover, for a standard normal random variable Z, Hölder's inequality implies that

$$\begin{aligned} \mathbb{E}\left[e^{q\beta(|x_0|+|X_{T-t_k^n}|+|X_{T-t_0}|)}\right] &\le e^{q\beta|x_0|}\left(\mathbb{E}\left[e^{2q\beta\sigma\sqrt{T-t_k^n}|Z|}\right]\right)^{1/2}\left(\mathbb{E}\left[e^{2q\beta\sigma\sqrt{T-t_0}|Z|}\right]\right)^{1/2} \\ &\le 2e^{q\beta|x_0|+2q^2\beta^2\sigma^2 T}. \end{aligned} \tag{3.7}$$

The claim then follows by (3.5), (3.6), and (3.7). □

4 The local error

4.1 Notation and definitions

Suppose that $(h,\theta) \in (0,\infty)\times(0,T]$. The aim of this section is to derive an upper bound for the absolute value of the error

$$\varepsilon^{\text{loc}}_{h,\theta}(g) := \mathbb{E}[g(X_{\tau_J}) - g(X_\theta)] \tag{4.1}$$

as a function of (h,θ), where the function g belongs to GBV_{exp} or $C^{0,\alpha}_{\text{exp}}$. The random variable J is given by

$$J = J(h,\theta) = \inf\{2m : \tau_{2m} > \theta\}.$$

Afterwards, upper bounds for the error (4.1) are derived in the dynamical setting, where the step size h and the level θ will depend on n. Observe that J agrees with J_n defined in (2.9) for $(h,\theta) = (\sigma\sqrt{\frac{T}{n}}, \frac{2T}{n}\lceil\frac{T-t}{2T/n}\rceil)$.

Let us start by introducing the following notation:

$$\mathbb{Z}^h_o := \{(2k+1)h : k \in \mathbb{Z}\}, \quad \mathbb{Z}^h_e := \{2kh : k \in \mathbb{Z}\}$$

(o refers to 'odd' and e refers to 'even'); then $\mathbb{Z}^h = \mathbb{Z}^h_o \cup \mathbb{Z}^h_e$. In addition, we will abbreviate

$$d_o(x) := \text{dist}(x,\mathbb{Z}^h_o), \quad d_e(x) := \text{dist}(x,\mathbb{Z}^h_e) = h - d_o(x), \quad x \in \mathbb{R}. \tag{4.2}$$

As in [13], we project functions onto piecewise linear functions in order to compute the conditional expectation $\mathbb{E}[g(X_{\tau_J})|\mathcal{F}_\theta]$.

Definition 4.1. Define operators Π_o and Π_e acting on functions $u : \mathbb{R} \to \mathbb{R}$ by

$$\Pi_e u(x) := u(x) \text{ if } x \in \mathbb{Z}^h_e \quad\text{and}\quad x \mapsto \Pi_e u(x) \quad\text{linear in}\quad [2kh,(2k+2)h] \quad \forall k \in \mathbb{Z},$$

$$\Pi_o u(x) := u(x) \text{ if } x \in \mathbb{Z}^h_o \quad\text{and}\quad x \mapsto \Pi_o u(x) \quad\text{linear in}\quad [(2k-1)h,(2k+1)h] \quad \forall k \in \mathbb{Z}.$$

The key ingredient in the estimation of the error $\varepsilon^{\text{loc}}_{h,\theta}(g)$ is the following result, which was proposed in [13, Section 9]. For the convenience of the reader, a sketch of the proof is given below. Recall Definition 3.1 for the class $\mathcal{B}_{\text{exp}}$, and denote by $\mathbb{N}_0 := \{0,1,2,\ldots\}$ the set of non-negative integers.

Proposition 4.2. *Let $(h,\theta) \in (0,\infty)\times(0,T]$ and define a random variable*

$$L = L(h,\theta) := \sup\{m \in \mathbb{N}_0 : \tau_m < \theta\}$$

(τ_L is equal to the largest of the stopping times $\tau_0,\tau_1,\ldots$ less than θ). Then, given a function $g \in \mathcal{B}_{exp}$,

$$\varepsilon^{loc}_{h,\theta}(g) = \mathbb{E}\big[\Pi_e g(X_\theta) - g(X_\theta)\big] + \mathbb{E}\big[\big(\Pi_o\Pi_e g(X_\theta) - \Pi_e g(X_\theta)\big)\mathbb{P}(L \text{ even}|X_\theta)\big]. \tag{4.3}$$

Proof. If $g \in \mathcal{B}_{\text{exp}}$, then also $\Pi_e g \in \mathcal{B}_{\text{exp}}$ and $\Pi_o\Pi_e g \in \mathcal{B}_{\text{exp}}$. The expectations on the right-hand side of (4.3) thus exist and are finite. Using the Markov property of the process $(X_t)_{t\geq 0}$, it can be shown that

$$\begin{aligned} \mathbb{E}\big[g(X_{\tau_J})\big|\mathcal{F}_\theta\big] &= \Pi_e g(X_\theta) \qquad \mathbb{P}\text{-a.s. on } \{L \text{ odd}\}, \\ \mathbb{E}\big[g(X_{\tau_J})\big|\mathcal{F}_\theta\big] &= \Pi_o\Pi_e g(X_\theta) \quad \mathbb{P}\text{-a.s. on } \{L \text{ even}\}, \end{aligned}$$

see [13, Section 9]. Consequently, since $\mathbb{1}_{\{L \text{ odd}\}} + \mathbb{1}_{\{L \text{ even}\}} = 1$ $\mathbb{P}$-a.s.,

$$\begin{aligned} \mathbb{E}[g(X_{\tau_J})] &= \mathbb{E}\left[\mathbb{E}\left[g(X_{\tau_J})\big|\mathcal{F}_\theta\right]\mathbb{1}_{\{L \text{ odd}\}}\right] + \mathbb{E}\left[\mathbb{E}\left[g(X_{\tau_J})\big|\mathcal{F}_\theta\right]\mathbb{1}_{\{L \text{ even}\}}\right] \\ &= \mathbb{E}\left[\Pi_e g(X_\theta)\mathbb{P}\left(L \text{ odd}\big|X_\theta\right)\right] + \mathbb{E}\left[\Pi_o\Pi_e g(X_\theta)\mathbb{P}\left(L \text{ even}\big|X_\theta\right)\right] \\ &= \mathbb{E}\left[\Pi_e g(X_\theta)\right] + \mathbb{E}\left[\left(\Pi_o\Pi_e g(X_\theta) - \Pi_e g(X_\theta)\right)\mathbb{P}(L \text{ even}|X_\theta)\right]. \end{aligned}$$

□

4.2 Evaluation of the conditional probability $\mathbb{P}(L \text{ even}|X_\theta)$

In this subsection we derive a representation for the function

$$y \mapsto \mathbb{P}(L \text{ even}|X_\theta = y) \tag{4.4}$$

based on first exit time probabilities of a Brownian bridge. This representation (4.11) together with the associated bounds presented in this subsection are applied in the proof of Propositions 4.14 and 6.1 below.

Definition 4.3 (Brownian bridge)**.** Let $x, y \in \mathbb{R}$ and $l > 0$. A Gaussian process $(B_t^{x,l,y})_{t\in[0,l]}$ with mean and covariance functions given by

$$\mathbb{E}[B_t^{x,l,y}] = x + \tfrac{t}{l}(y-x), \quad 0 \le t \le l,$$
$$\mathrm{Cov}(B_s^{x,l,y}, B_t^{x,l,y}) = s\left(1 - \tfrac{t}{l}\right), \quad 0 \le s \le t \le l,$$

is called a (generalized) Brownian bridge from x to y of length l.

Remark 4.4. By comparing mean and covariance functions, it is easy to verify that a Brownian bridge $(B_t^{x,l,y})_{t\in[0,l]}$ is equal in law with the transformed processes below:

$$(B_{l-t}^{y,l,x})_{t\in[0,l]} \qquad \text{('time reversal')} \tag{4.5}$$
$$(x + B_t^{0,l,y-x})_{t\in[0,l]} \qquad \text{('translation')} \tag{4.6}$$
$$(-B^{-x,l,-y})_{t\in[0,l]} \qquad \text{('reflection around the } x\text{-axis')}. \tag{4.7}$$

A continuous version of a Brownian bridge $(B_t^{x,\theta,y})_{t\in[0,\theta]}$ can be thought as a random function on the canonical space $(C[0,\theta], \mathcal{B}(C[0,\theta]), \mathbb{P}_{x,\theta,y})$, where $\mathbb{P}_{x,\theta,y}$ denotes the associated probability measure. In the following proposition we give different characterizations for the function (4.4) in terms of hitting times. For all $c \in \mathbb{R}$, $a < b$, and $\omega \in C[0,\theta]$, we let

$$H_c(\omega) := \inf\{t \in [0,\theta] : \omega_t = c\}, \qquad H_{(a,b)}(\omega) := \inf\{t \in [0,\theta] : \omega_t \notin (a,b)\},$$
$$\hat{H}_c(\omega) := \sup\{t \in [0,\theta] : \omega_t = c\}, \qquad \hat{H}_{(a,b)}(\omega) := \sup\{t \in [0,\theta] : \omega_t \notin (a,b)\}.$$

Proposition 4.5. *Let $(h,\theta) \in (0,\infty)\times(0,T]$. Suppose that $(B_t^{y/\sigma,\theta,0})_{t\in[0,\theta]}$ is a Brownian bridge on a probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$, and define*

$$q(y) = q(y,h,\theta) := \tilde{\mathbb{P}}((B_t^{y/\sigma,\theta,0})_{t\in[0,\theta]} \text{ hits } \mathbb{Z}_e^{h/\sigma} \text{ before hitting } \mathbb{Z}_o^{h/\sigma}), \quad y \in \mathbb{R}. \tag{4.8}$$

Then, for all $k \in \mathbb{Z}$,

$$(i) \quad q(y) = \mathbb{P}(L \text{ even}|X_\theta = y), \quad y \notin \mathbb{Z}^h, \tag{4.9}$$

$$(ii) \quad q(y) = \begin{cases} \mathbb{P}_{y/\sigma,\theta,0}(H_{2kh/\sigma} < H_{(2k+1)h/\sigma}), & y \in (2kh,(2k+1)h), \\ \mathbb{P}_{y/\sigma,\theta,0}(H_{2kh/\sigma} < H_{(2k-1)h/\sigma}), & y \in ((2k-1)h, 2kh), \end{cases} \tag{4.10}$$

$$(iii) \quad q(y) = \begin{cases} \dfrac{d_o(y)}{h} + \dfrac{\sigma}{h}\mathbb{E}_{\tilde{\mathbb{P}}}\left[B^{0,\theta,y/\sigma}_{\tilde{H}_{(-((2k+1)h-y)/\sigma,(y-2kh)/\sigma)}}\right], & y \in (2kh,(2k+1)h), \\ \dfrac{d_o(y)}{h} - \dfrac{\sigma}{h}\mathbb{E}_{\tilde{\mathbb{P}}}\left[B^{0,\theta,y/\sigma}_{\tilde{H}_{(-(2kh-y)/\sigma,(y-(2k-1)h)/\sigma)}}\right], & y \in ((2k-1)h, 2kh). \end{cases} \tag{4.11}$$

Here $\tilde{H}_{(a,b)} = \inf\{t \in [0,\theta] : B_t^{0,\theta,y/\sigma} \notin (a,b)\}$, and $\mathbb{P}$ refers to the probability measure on the space $(\Omega, \mathcal{F}, \mathbb{P})$ considered in Section 2.

Remark 4.6. It is clear by (4.7) that the function q is symmetric.

Proof of Proposition 4.5. Item (ii) is clear. To show (i), observe that if $X_\theta(\omega) \in (2kh,(2k+1)h)$ and $L(\omega)$ is even, the path $t \mapsto X_t(\omega)$ does hit $2kh$ at $\tau_L(\omega)$ and afterwards, i.e. on $[\tau_L(\omega), \theta)$, it does not hit any other mh ($m \neq 2k$) and hence stays inside $((2k-1)h,(2k+1)h)$. Therefore, the last entry of this path into $(2kh,(2k+1)h)$ occurs via $2kh$, and thus

$$\begin{aligned}\mathbb{P}(L \text{ even}, X_\theta \in (2kh,(2k+1)h)) &= \mathbb{P}_0(\sigma\omega_{\hat{H}_{(2kh,(2k+1)h)}(\omega)} = 2kh,\ \sigma\omega_\theta \in (2kh,(2k+1)h)) \\ &= \mathbb{P}_0\big(\omega_{\hat{H}_{(2kh/\sigma,(2k+1)h/\sigma)}(\omega)} = \tfrac{2kh}{\sigma},\ \omega_\theta \in (\tfrac{2kh}{\sigma}, \tfrac{(2k+1)h}{\sigma})\big) \\ &= \mathbb{P}_0(\hat{H}_{2kh/\sigma} > \hat{H}_{(2k+1)h/\sigma}),\end{aligned}$$

where $\mathbb{P}_0$ denotes the Wiener measure on $(C[0,\theta],\mathcal{B}(C[0,\theta]))$. Thus, for $y \in (2kh,(2k+1)h)$,

$$\mathbb{P}(L \text{ even}|X_\theta = y) = \mathbb{P}_{0,\theta,y/\sigma}(\hat{H}_{2kh/\sigma} > \hat{H}_{(2k+1)h/\sigma}) = \mathbb{P}_{y/\sigma,\theta,0}(H_{2kh/\sigma} < H_{(2k+1)h/\sigma}) = q(y),$$

where we used relations (4.5), (4.10), and the fact that $\mathbb{P}(\,\cdot\,|X_\theta = y) = \mathbb{P}_{0,\theta,y/\sigma}$ on $(C[0,\theta],\mathcal{B}(C[0,\theta]))$ (see e.g. [10, Chapter 1]). The case $y \in ((2k-1)h, 2kh)$ is similar.

For (iii), assume $y \in ((2k-1)h, 2kh)$; the case $y \in (2kh,(2k+1)h)$ is similar. It is clear that whenever $z \notin (a,b)$, $a < 0 < b$, and $\tilde{H}_{(a,b)} = \inf\{t \in [0,\theta] : B_t^{0,\theta,z} \notin (a,b)\}$,

$$\mathbb{P}_{0,\theta,z}(H_a < H_b) = \frac{b}{b-a} - \frac{1}{b-a}\mathbb{E}_{\tilde{\mathbb{P}}}\Big[B^{0,\theta,z}_{\tilde{H}_{(a,b)}}\Big]. \tag{4.12}$$

In addition, from (4.10) we deduce that

$$\begin{aligned}q(y) &= \mathbb{P}_{y/\sigma,\theta,0}(H_{2kh/\sigma} < H_{(2k-1)h/\sigma}) \\ &= \mathbb{P}_{0,\theta,-y/\sigma}(H_{(2kh-y)/\sigma} < H_{((2k-1)h-y)/\sigma}) \\ &= \mathbb{P}_{0,\theta,y/\sigma}(H_{(y-2kh)/\sigma} < H_{(y-(2k-1)h)/\sigma})\end{aligned} \tag{4.13}$$

by (4.6) and (4.7). Substitute $z = \frac{y}{\sigma}$, $a = \frac{y-2kh}{\sigma}$, and $b = \frac{y-(2k-1)h}{\sigma}$. Then $z \notin (a,b)$, $a < 0 < b$, $b-a = \frac{h}{\sigma}$, and hence by (4.12), (4.13), and $d_o(y) = y-(2k-1)h$,

$$q(y) = \frac{d_o(h)}{h} - \frac{\sigma}{h}\mathbb{E}_{\tilde{\mathbb{P}}}\left[B^{0,\theta,y/\sigma}_{\tilde{H}_{((y-2kh)/\sigma,(y-(2k-1)h)/\sigma)}}\right].$$

□

The probability for the Brownian motion $(W_t + y/\sigma)_{t\geq 0}$ to hit the set $\mathbb{Z}_e^{h/\sigma}$ before hitting the set $\mathbb{Z}_o^{h/\sigma}$ is equal to $d_o(y)/h$ (cf. (4.8)). As pointed out in [13, Section 9], the piecewise linear function $y \mapsto d_o(y)/h$ can be used to approximate the function $y \mapsto q(y)$ for small $h > 0$. Estimates related to this approximation, which are also applied in the proof of Proposition 6.1, are presented in the proposition below. We denote by $p = p(\,\cdot\,,\theta)$ the density of the random variable X_θ.

Proposition 4.7. *Suppose that* $(h,\theta) \in (0,\infty)\times(0,T]$ *and define*

$$\varrho : \mathbb{R} \to \mathbb{R}, \quad \varrho(y) = \varrho(y,h,\theta) := q(y) - d_o(y)/h, \tag{4.14}$$

where $q = q(\,\cdot\,,h,\theta)$ *was introduced in (4.8). Then* ϱ *is symmetric, and it holds that*

$$(i) \quad \int_0^h |\varrho(y)|\,p(y)dy \leq \frac{26}{10}\frac{h}{\sigma\sqrt{\theta}} + \frac{h^2}{\sigma^2\theta}, \qquad (ii) \quad \int_h^\infty |\varrho(y)|\,p(y)dy \leq \frac{29}{10}\frac{h}{\sigma\sqrt{\theta}} + \frac{h^2}{\sigma^2\theta}.$$

Proposition 4.7 can be seen as a generalization of [4, Corollary 3.3] to the time-dependent setting. A detailed proof is presented can be found in the arXiv version [9, Subsection 4.2]. The proof uses certain estimates of [4] related to the expected first hitting times of Brownian bridges [4, Lemma 3.1 and Lemma 3.2 (i)]. For the further use of these estimates in the proof of Proposition 4.14, we present them in the lemma below using the notation of this subsection.

Lemma 4.8. *Let* $(h,\theta)\in(0,\infty)\times(0,\infty)$ *and suppose that* $a<0<b$ *and* $y\notin(a,b)$*. Then*

$$\mathbb{E}_{\tilde{\mathbb{P}}}\big[B^{0,\theta,y}_{\tilde{H}_{(a,b)}}\big]\le\frac{\mathbb{E}_{0,\theta,y}[H_{(a,b)}]}{\theta}\Big(|y|+2(|a|\vee b)+3\sqrt{2\theta}\Big),\tag{4.15}$$

$$\mathbb{E}_{0,\theta,y}[H_{(a,b)}]\le\begin{cases}b(2|a|+y)\wedge\theta, & y\ge b,\\ |a|(2b+|y|)\wedge\theta, & y\le a.\end{cases}\tag{4.16}$$

4.3 The local error for $g\in GBV_{\text{exp}}$

The estimation of the local error for the class GBV_{exp} relies on the following observation: If $g\in GBV_{\text{exp}}$ is given by (2.7) and if $g^{x_0}:=g(x_0+\cdot\,)$ for some $x_0\in\mathbb{R}$, then

$$g^{x_0}(x)=c+\int\limits_{[0,\infty)}\mathbb{1}_{(y-x_0,\infty)}(x)d\mu(y)-\int\limits_{(-\infty,0)}\mathbb{1}_{(-\infty,y-x_0]}(x)d\mu(y)+\sum_{i=1}^{\infty}\alpha_i\mathbb{1}_{\{x_i-x_0\}}(x).\tag{4.17}$$

Using the representation (4.17) and linearity, the estimation of the error $\varepsilon^{\text{loc}}_{h,\theta}(g^{x_0})$ essentially reduces to the estimation of integrals, where the integrands consist of indicator functions or their linear approximations given by the operators Π_e and Π_o (introduced in Definition 4.1). The following proposition enables us to interchange the order of integration or summation with the application of these operators.

Recall that $p=p(\,\cdot\,,\theta)$ denotes the density of X_θ and that $q=q(\,\cdot\,,h,\theta)$ is the function defined in (4.8).

Proposition 4.9. *Suppose that* $(h,\theta)\in(0,\infty)\times(0,T]$ *and that* $g\in GBV_{exp}$ *admits the representation (2.7). Then, for all* $x_0\in\mathbb{R}$*,*

$$(i)\quad \Pi_e g^{x_0}(x)=c+\int\limits_{[0,\infty)}\Pi_e\mathbb{1}_{(y-x_0,\infty)}(x)d\mu(y)-\int\limits_{(-\infty,0)}\Pi_e\mathbb{1}_{(-\infty,y-x_0]}(x)d\mu(y)+\sum_{i\in\mathbb{N}:x_i-x_0\in\mathbb{Z}^h_e}\alpha_i\Pi_e\mathbb{1}_{\{x_i-x_0\}}(x),\quad x\in\mathbb{R},$$

$$(ii)\quad \Pi_o\Pi_e g^{x_0}(x)=c+\int\limits_{[0,\infty)}\Pi_o\Pi_e\mathbb{1}_{(y-x_0,\infty)}(x)d\mu(y)-\int\limits_{(-\infty,0)}\Pi_o\Pi_e\mathbb{1}_{(-\infty,y-x_0]}(x)d\mu(y)+\sum_{i\in\mathbb{N}:x_i-x_0\in\mathbb{Z}^h_e}\alpha_i\Pi_o\Pi_e\mathbb{1}_{\{x_i-x_0\}}(x),\quad x\in\mathbb{R}.$$

Idea of the proof. Items (i)–(ii) follow by using the representation (4.17), linearity of the operations $f\mapsto\Pi_e f$, $f\mapsto\Pi_o f$, and $f\mapsto\int f d|\mu|$, and relation (A.13). □

Proposition 4.10. *Let* $(h,\theta)\in(0,\infty)\times(0,T]$*. Suppose that* $g\in GBV_{exp}$ *admits the representation (2.7) and that* $\beta\ge0$ *is as in (2.8). Then, for all* $x_0\in\mathbb{R}$*,*

$$\big|\mathbb{E}[g(x_0+X_{\tau_J})-g(x_0+X_\theta)]\big|\le\frac{7}{\sqrt{2\pi}}\frac{h}{\sigma\sqrt{\theta}}e^{3\beta h+\beta|x_0|+\beta^2\sigma^2T/2}\times\Bigg(\int_{\mathbb{R}}e^{-\beta|y|}d|\mu|(y)+\sum_{i\in\mathbb{N}:x_i-x_0\in\mathbb{Z}^h_e}|\alpha_i|\,e^{-\beta|x_i|}\Bigg).\tag{4.18}$$

Proof. For given $x_0\in\mathbb{R}$, we apply (4.3) for the function $g(x_0+\cdot\,)$. By Proposition 4.9 and by the relation $\mathbb{P}(L\text{ even}|X_\theta=x)=q(x)$ (Leb-a.e.), we may decompose the expectation on the left-hand side of (4.18) in the following way:

$$
\begin{aligned}
&\mathbb{E}[g(x_0+X_{\tau_J}) - g(x_0+X_\theta)] \\
&= \int_{\mathbb{R}} \int_{[0,\infty)} \left[\Pi_e \mathbb{1}_{(y-x_0,\infty)}(x) - \mathbb{1}_{(y-x_0,\infty)}(x)\right] d\mu(y) p(x) dx \\
&\quad + \int_{\mathbb{R}} \int_{(-\infty,0)} \left[\Pi_e \mathbb{1}_{(-\infty,y-x_0]}(x) - \mathbb{1}_{(-\infty,y-x_0]}(x)\right] d\mu(y) p(x) dx \\
&\quad + \int_{\mathbb{R}} \int_{[0,\infty)} \left[\Pi_o \Pi_e \mathbb{1}_{(y-x_0,\infty)}(x) - \Pi_e \mathbb{1}_{(y-x_0,\infty)}(x)\right] d\mu(y) q(x) p(x) dx \\
&\quad + \int_{\mathbb{R}} \int_{(-\infty,0)} \left[\Pi_o \Pi_e \mathbb{1}_{(-\infty,y-x_0]}(x) - \Pi_e \mathbb{1}_{(-\infty,y-x_0]}(x)\right] d\mu(y) q(x) p(x) dx \\
&\quad + \int_{\mathbb{R}} \sum_{i\in\mathbb{N}: x_i - x_0 \in \mathbb{Z}_e^h} \alpha_i \Pi_e \mathbb{1}_{\{x_i - x_0\}}(x) p(x) dx \\
&\quad + \int_{\mathbb{R}} \sum_{i\in\mathbb{N}: x_i - x_0 \in \mathbb{Z}_e^h} \alpha_i \left[\Pi_o \Pi_e \mathbb{1}_{\{x_i - x_0\}}(x) - \Pi_e \mathbb{1}_{\{x_i - x_0\}}(x)\right] q(x) p(x) dx \\
&=: E^{(1)} + E^{(2)} + E^{(3)} + E^{(4)} + E^{(5)} + E^{(6)}.
\end{aligned}
$$

We will derive upper estimates for the quantities $|E^{(i)}|$, $1 \le i \le 6$, in the following steps.
Step 1: $E^{(1)}$ **and** $E^{(2)}$**.** Suppose that $y - x_0 \in [2kh, (2k+2)h)$ for some $k \in \mathbb{Z}$. Then

$$
\left|\Pi_e \mathbb{1}_{(y-x_0,\infty)}(x) - \mathbb{1}_{(y-x_0,\infty)}(x)\right| \le \mathbb{1}_{[2kh,(2k+2)h)}(x),
$$

and since for each $x \in [2kh, (2k+2)h)$ it holds that $|y| \le 2h + |x_0| + |x|$, we have

$$
\begin{aligned}
&e^{\beta|y|} \int_{\mathbb{R}} \left|\Pi_e \mathbb{1}_{(y-x_0,\infty)}(x) - \mathbb{1}_{(y-x_0,\infty)}(x)\right| p(x) dx \\
&\le e^{2\beta h + \beta|x_0|} \int_{\mathbb{R}} e^{\beta|x|} \left|\Pi_e \mathbb{1}_{(y-x_0,\infty)}(x) - \mathbb{1}_{(y-x_0,\infty)}(x)\right| p(x) dx \\
&\le e^{2\beta h + \beta|x_0|} \int_{2kh}^{(2k+2)h} e^{\beta|x|} p(x) dx \\
&\le \frac{2}{\sqrt{2\pi}} e^{2\beta h + \beta|x_0| + \beta^2\sigma^2 T/2} \frac{h}{\sigma\sqrt{\theta}}.
\end{aligned}
$$

Consequently, by Fubini's theorem,

$$
\begin{aligned}
\left|E^{(1)}\right| &\le \int_{[0,\infty)} e^{-\beta|y|} \Big(e^{\beta|y|} \int_{\mathbb{R}} \left|\Pi_e \mathbb{1}_{(y-x_0,\infty)}(x) - \mathbb{1}_{(y-x_0,\infty)}(x)\right| p(x) dx \Big) d|\mu|(y) \\
&\le \frac{h}{\sigma\sqrt{\theta}} \frac{2}{\sqrt{2\pi}} e^{2\beta h + \beta|x_0| + \beta^2\sigma^2 T/2} \int_{[0,\infty)} e^{-\beta|y|} d|\mu|(y). \qquad (4.19)
\end{aligned}
$$

In fact, it also holds that

$$
\left|E^{(2)}\right| \le \frac{h}{\sigma\sqrt{\theta}} \frac{2}{\sqrt{2\pi}} e^{2\beta h + \beta|x_0| + \beta^2\sigma^2 T/2} \int_{(-\infty,0)} e^{-\beta|y|} d|\mu|(y) \qquad (4.20)
$$

since $\left|\Pi_e \mathbb{1}_{(-\infty,y-x_0]}(x) - \mathbb{1}_{(-\infty,y-x_0]}(x)\right| = \left|\Pi_e \mathbb{1}_{(y-x_0,\infty)}(x) - \mathbb{1}_{(y-x_0,\infty)}(x)\right|$ for all $x \in \mathbb{R}$, which is a direct consequence of the relation

$$
\Pi_e \mathbb{1}_{(-\infty,r]} = 1 - \Pi_e \mathbb{1}_{(r,\infty)}, \quad r \in \mathbb{R}. \qquad (4.21)
$$

Step 2: $E^{(3)}$ **and** $E^{(4)}$**.** Suppose $y-x_0 \in [2kh, (2k+2)h)$ for some $k \in \mathbb{Z}$. Then $|y| \leq 3h + |x_0| + |x|$ holds for all $x \in [(2k-1)h, (2k+3)h)$, and by (A.14) we may estimate

$$\begin{aligned}
&e^{\beta|y|} \int_{\mathbb{R}} \left| \Pi_o \Pi_e \mathbb{1}_{(y-x_0,\infty)}(x) - \Pi_e \mathbb{1}_{(y-x_0,\infty)}(x) \right| q(x)p(x)dx \\
&\leq e^{3\beta h + \beta|x_0|} \int_{\mathbb{R}} e^{\beta|x|} \left| \Pi_o \Pi_e \mathbb{1}_{(y-x_0,\infty)}(x) - \Pi_e \mathbb{1}_{(y-x_0,\infty)}(x) \right| q(x)p(x)dx \\
&\leq e^{3\beta h + \beta|x_0|} \int_{(2k-1)h}^{(2k+3)h} e^{\beta|x|} \frac{d_o(x)}{4h} q(x)p(x)dx \\
&\leq \frac{1}{4} e^{3\beta h + \beta|x_0|} \int_{(2k-1)h}^{(2k+3)h} e^{\beta|x|} p(x)dx \\
&\leq \frac{1}{\sqrt{2\pi}} e^{3\beta h + \beta|x_0| + \beta^2\sigma^2 T/2} \frac{h}{\sigma\sqrt{\theta}}.
\end{aligned}$$

Hence, by Fubini's theorem,

$$\begin{aligned}
\left|E^{(3)}\right| &\leq \int_{[0,\infty)} e^{-\beta|y|} \Big(e^{\beta|y|} \int_{\mathbb{R}} \left| \Pi_o \Pi_e \mathbb{1}_{(y-x_0,\infty)}(x) - \Pi_e \mathbb{1}_{(y-x_0,\infty)}(x) \right| q(x)p(x)dx \Big) d|\mu|(y) \\
&\leq \frac{h}{\sigma\sqrt{\theta}} \frac{1}{\sqrt{2\pi}} e^{3\beta h + \beta|x_0| + \beta^2\sigma^2 T/2} \int_{[0,\infty)} e^{-\beta|y|} d|\mu|(y). \qquad (4.22)
\end{aligned}$$

Moreover, by (4.21) and by the linearity of Π_o, we obtain

$$\left|E^{(4)}\right| \leq \frac{h}{\sigma\sqrt{\theta}} \frac{1}{\sqrt{2\pi}} e^{3\beta h + \beta|x_0| + \beta^2\sigma^2 T/2} \int_{(-\infty,0)} e^{-\beta|y|} d|\mu|(y), \qquad (4.23)$$

since $\left|\Pi_o \Pi_e \mathbb{1}_{(-\infty,y-x_0]}(x) - \Pi_e \mathbb{1}_{(-\infty,y-x_0]}(x)\right| = \left|\Pi_o \Pi_e \mathbb{1}_{(y-x_0,\infty)}(x) - \Pi_e \mathbb{1}_{(y-x_0,\infty)}(x)\right|$, $x \in \mathbb{R}$.
Step 3: $E^{(5)}$**.** By (A.13), $\Pi_e \mathbb{1}_{\{\xi\}} \equiv 0$ if $\xi \notin \mathbb{Z}_e^h$, and by (A.15), $\Pi_e \mathbb{1}_{\{\xi\}} \leq \mathbb{1}_{[\xi-2h,\xi+2h]}$ if $\xi \in \mathbb{Z}_e^h$. In addition, since $|x_i| \leq 2h + |x_0| + |x|$ whenever $|x - (x_i - x_0)| \leq 2h$,

$$\begin{aligned}
\left|E^{(5)}\right| &\leq \sum_{i\in\mathbb{N}: x_i - x_0 \in \mathbb{Z}_e^h} \left| \alpha_i \int_{\mathbb{R}} \Pi_e \mathbb{1}_{\{x_i - x_0\}}(x) p(x)dx \right| \\
&\leq \sum_{i\in\mathbb{N}: x_i - x_0 \in \mathbb{Z}_e^h} |\alpha_i| \, e^{-\beta|x_i|} \int_{\mathbb{R}} e^{\beta|x_i|} p(x) \mathbb{1}_{[(x_i-x_0)-2h,(x_i-x_0)+2h]}(x)dx \\
&\leq \sum_{i\in\mathbb{N}: x_i - x_0 \in \mathbb{Z}_e^h} |\alpha_i| \, e^{-\beta|x_i|} \int_{(x_i-x_0)-2h}^{(x_i-x_0)+2h} e^{\beta|x_i|} p(x)dx \\
&\leq \frac{h}{\sigma\sqrt{\theta}} \frac{4}{\sqrt{2\pi}} e^{2\beta h + \beta|x_0| + \beta^2\sigma^2 T/2} \sum_{i\in\mathbb{N}: x_i - x_0 \in \mathbb{Z}_e^h} |\alpha_i| \, e^{-\beta|x_i|}. \qquad (4.24)
\end{aligned}$$

Step 4: $E^{(6)}$**.** If $\xi \in \mathbb{Z}_e^h$, relations (A.13), (A.18), and the linearity of Π_o imply that

$$\begin{aligned}
&\Pi_o \Pi_e \mathbb{1}_{\{\xi\}}(x) - \Pi_e \mathbb{1}_{\{\xi\}}(x) \\
&= \frac{1}{4h} \Big(\Pi_o \left| \cdot - (\xi - 2h) \right|(x) - |x - (\xi - 2h)| \Big) - \frac{1}{2h} \Big(\Pi_o \left| \cdot - \xi \right|(x) - |x - \xi| \Big) \\
&\quad + \frac{1}{4h} \Big(\Pi_o \left| \cdot - (\xi + 2h) \right|(x) - |x - (\xi + 2h)| \Big)
\end{aligned}$$

$$= \frac{d_o(x)}{4h}\Big(\mathbb{1}_{[\xi-3h,\xi-h)}(x) - 2\mathbb{1}_{[\xi-h,\xi+h)}(x) + \mathbb{1}_{[\xi+h,\xi+3h)}(x)\Big), \quad x \in \mathbb{R}.$$

In addition, we have $\Pi_o\Pi_e\mathbb{1}_{\{\xi\}} - \Pi_e\mathbb{1}_{\{\xi\}} \equiv 0$ for $\xi \notin \mathbb{Z}_e^h$ by (A.13). Therefore, since $|x_i| \le 3h + |x| + |x_0|$ whenever $|x - (x_i - x_0)| \le 3h$, we get

$$\begin{aligned}
\left|E^{(6)}\right| &\le \sum_{i\in\mathbb{N}:x_i-x_0\in\mathbb{Z}_e^h} |\alpha_i| \int_{\mathbb{R}} \left|\Pi_o\Pi_e\mathbb{1}_{\{x_i-x_0\}}(x) - \Pi_e\mathbb{1}_{\{x_i-x_0\}}(x)\right| q(x)p(x)dx \\
&\le \sum_{i\in\mathbb{N}:x_i-x_0\in\mathbb{Z}_e^h} |\alpha_i|\, e^{-\beta|x_i|} \int_{(x_i-x_0)-3h}^{(x_i-x_0)+3h} e^{\beta|x_i|}\frac{d_o(x)}{2h} q(x)p(x)dx \\
&\le \sum_{i\in\mathbb{N}:x_i-x_0\in\mathbb{Z}_e^h} \frac{|\alpha_i|}{2} e^{-\beta|x_i|+3\beta h+\beta|x_0|} \int_{(x_i-x_0)-3h}^{(x_i-x_0)+3h} e^{\beta|x|}p(x)dx \\
&\le \frac{h}{\sigma\sqrt{\theta}}\frac{3}{\sqrt{2\pi}} e^{3\beta h+\beta|x_0|+\beta^2\sigma^2T/2} \sum_{i\in\mathbb{N}:x_i-x_0\in\mathbb{Z}_e^h} |\alpha_i|\, e^{-\beta|x_i|}. \qquad (4.25)
\end{aligned}$$

It remains to observe that the sum the right-hand sides of (4.19), (4.20), (4.22), (4.23), (4.24), and (4.25) are bounded from above by the right-hand side of (4.18). □

In order to distinguish between the general setting (h, θ) and the specific n-dependent setting (h_n, θ_n), we will refer to the assumption below.

Assumption 4.11. *For given $t_0 \in [0, T)$ and $n \in 2\mathbb{N}$, we substitute $(h, \theta) = (h_n, \theta_n)$, where*

$$h_n = \sigma\sqrt{\frac{T}{n}}, \quad \theta_n = \frac{n_\theta T}{n} \quad \textit{and} \quad n_\theta = 2\left\lceil \frac{T-t_0}{2T/n} \right\rceil$$

as in (2.4). For notational convenience, we will drop the subscript n from h_n.

Remark 4.12. The special choice $(h, \theta) = (h_n, \theta_n)$ in Assumption 4.11 affects the objects below used throughout this text:

$$\begin{aligned}
\tau_k &= \inf\left\{t > \tau_{k-1} : |X_t - X_{\tau_{k-1}}| = h\right\}, \quad (X_{\tau_k})_{k=0,1,\dots}, \quad (\mathcal{F}_{\tau_k})_{k=0,1,\dots}, \\
J_n &= J = \inf\{2m \in 2\mathbb{N} : \tau_{2m} > \theta_n\}, \quad L_n = L = \sup\{m \in \mathbb{N}_0 : \tau_m < \theta_n\}, \\
\mathbb{Z}_e^h &= \{2kh : k \in \mathbb{Z}\}, \quad \mathbb{Z}_o^h = \{(2k+1)h : k \in \mathbb{Z}\}, \quad \mathbb{Z}^h = \mathbb{Z}_o^h \cup \mathbb{Z}_e^h, \\
d_o(x) &= \operatorname{dist}(x, \mathbb{Z}_o^h), \quad d_e(x) = \operatorname{dist}(x, \mathbb{Z}_e^h), \quad p(x) = \mathbb{P}(X_{\theta_n} \in dx)/dx.
\end{aligned}$$

This choice also affects the functions $q = q(\,\cdot\,, h, \theta)$ and $\varrho = \varrho(\,\cdot\,, h, \theta)$ defined in (4.5) and (4.14), respectively. In particular, Proposition 4.5 implies that

$$q(x) = \mathbb{P}(L_n \text{ even} | X_{\theta_n} = x), \quad x \notin \mathbb{Z}^h.$$

For the main result of this subsection, recall that $\varepsilon_n^{\text{loc}}(t_0, x_0) = \mathbb{E}[g(x_0+X_{\tau_{J_n}}) - g(x_0+X_{\theta_n})]$.

Corollary 4.13. *Let $n \in 2\mathbb{N}$. Suppose that the function $g \in GBV_{exp}$ admits the representation (2.7) and that $\beta \ge 0$ is as in (2.8). Then, under Assumption 4.11, there exists a constant $C > 0$ such that for all $(t_0, x_0) \in [0, T)\times\mathbb{R}$,*

$$\left|\varepsilon_n^{loc}(t_0, x_0)\right| \le \frac{C\sqrt{T}}{\sqrt{n(T - t_k^n)}} e^{\beta|x_0|+3\beta^2\sigma^2T}, \quad t_0 \in [t_k^n, t_{k+1}^n), \quad 0 \le k < \tfrac{n}{2}.$$

Proof. Proposition 4.10 and the relation $h(\sigma^2\theta_n)^{-1/2} = n_\theta^{-1/2}$ imply that

$$\left|\varepsilon_n^{\mathrm{loc}}(t_0, x_0)\right| \le \frac{C_{\beta,\sigma,T} e^{\beta|x_0|}}{\sqrt{n_\theta}} \left(\int_{\mathbb{R}} e^{-\beta|y|} d|\mu|(y) + \sum_{i\in\mathbb{N}: x_i - x_0 \in \mathbb{Z}_e^h} |\alpha_i|\, e^{-\beta|x_i|} \right),$$

where the coefficient $C_{\beta,\sigma,T} > 0$ implied by (4.18) can be estimated as follows:

$$C_{\beta,\sigma,T} = \frac{7}{\sqrt{2\pi}} e^{3\beta h + \beta^2\sigma^2 T/2} \le \frac{7}{\sqrt{2\pi}} e^{\frac{5}{2}\beta\sigma\sqrt{T} + \beta^2\sigma^2 T/2} \le C e^{3\beta^2\sigma^2 T}$$

for a constant $C > 0$. Since $n_\theta T = n(T - t_k^n)$ for $t_0 \in [t_k^n, t_{k+1}^n)$ by (2.5), we obtain the desired result. □

4.4 On the sharpness of the rate for the class GBV_{exp}

The following lemma indicates that the rate $n^{-1/2}$ for the class GBV_{exp} is sharp.

Proposition 4.14. *Under Assumption 4.11, there exists a function $g \in GBV_{exp}$ such that*

$$0 < \liminf_{n\to\infty} n^{1/2} \varepsilon_n(0,0) \le \limsup_{n\to\infty} n^{1/2}\varepsilon_n(0,0) < \infty. \tag{4.26}$$

Proof. For simplicity, let $T = \sigma = 1$ and $g := \mathbb{1}_{[0,\infty)}$. Then $h = n^{-1/2}$, $g \in GBV_{\mathrm{exp}}$, and the location of the jump of g belongs to the set $\mathbb{Z}_e^h$ for all $n \in \mathbb{N}$. Observe that then $\varepsilon_n^{\mathrm{adj}}(0,0) = 0$ by Proposition 3.3 and $|\varepsilon_n^{\mathrm{glob}}(0,0)| \le Cn^{-1}$ by Proposition 5.3 below, where $C > 0$ is some constant. Consequently, it suffices to show that (4.26) is valid for the local error $\varepsilon_n^{\mathrm{loc}}(0,0)$; recall (2.10).

The expression $n^{1/2}\varepsilon_n^{\mathrm{loc}}(0,0)$ is bounded from above by Corollary 4.13. For the lower bound, we note that by Definition 4.1,

$$\Pi_e \mathbb{1}_{[0,\infty)}(x) = \left(1 \wedge \tfrac{x+2h}{2h}\right) \mathbb{1}_{[-2h,\infty)}(x), \quad \Pi_o\Pi_e\mathbb{1}_{[0,\infty)}(x) = \left(1 \wedge \tfrac{x+3h}{4h}\right)\mathbb{1}_{[-3h,\infty)}(x), \quad x \in \mathbb{R}.$$

Consequently, for $(h,\theta) = (n^{-1/2}, 1)$, Proposition 4.2 and relation (4.9) yield

$$\begin{aligned}
\varepsilon_n^{\mathrm{loc}}(0,0) &= \mathbb{E}\left[\Pi_e\mathbb{1}_{[0,\infty)}(W_1) - \mathbb{1}_{[0,\infty)}(W_1)\right] + \mathbb{E}\left[\left(\Pi_o\Pi_e\mathbb{1}_{[0,\infty)}(W_1) - \Pi_e\mathbb{1}_{[0,\infty)}(W_1)\right) q(W_1)\right] \\
&= \int_{-2h}^{0} \frac{x+2h}{2h} p(x)dx + \int_{-3h}^{h} \left[\frac{x+3h}{4h} - \left(\frac{x+2h}{2h}\mathbb{1}_{[-2h,0)}(x) + \mathbb{1}_{[0,\infty)}(x)\right)\right] q(x)p(x)dx \\
&= \int_{-2h}^{0} \frac{x+2h}{2h}(1-q(x))p(x)dx + \int_{-3h}^{0} \frac{x+3h}{4h} q(x)p(x)dx + \int_0^h \frac{x-h}{4h} q(x)p(x)dx \\
&\ge p(h)\int_{-h}^{0} \frac{x+2h}{2h}(1-q(x))dx
\end{aligned}$$

by the symmetry of the functions $p > 0$ and $q \in [0,1]$. Moreover, in terms of the function ϱ defined in (4.14), we deduce for $x \in (-h, 0)$ that

$$1 - q(x) = 1 - \frac{d_o(x)}{h} - \varrho(x) \ge \frac{d_e(x)}{h} - |\varrho(x)| \ge \frac{|x|}{h} - 3h(h + \sqrt{2}), \tag{4.27}$$

where the last inequality on the right-hand side of (4.27) follows by applying relations (4.11) , (4.15), and (4.16) for $k = 0$ and $\sigma = \theta = 1$;

$$\begin{aligned}
|\varrho(x)| = \frac{1}{h}\mathbb{E}_{\tilde{\mathbb{P}}}\big[B^{0,1,x}_{\tilde{H}_{(x,h+x)}}\big] &\le \frac{1}{h}\Big(|x| + 2(|x| \vee (x+h)) + 3\sqrt{2}\Big)\mathbb{E}_{0,1,x}[H_{(x,h+x)}] \\
&\le (3h + 3\sqrt{2})\frac{|x|}{h}\Big(2h - |x|\Big) \\
&\le 3h(h + \sqrt{2}).
\end{aligned}$$

Hence, there exist constants $C_1, C_2 > 0$ not depending on h such that

$$\varepsilon_n^{\mathrm{loc}}(0,0) \geq p(h) \int_{-h}^{0} \frac{x+2h}{2h} \frac{|x|}{h} dx - 3h(h+\sqrt{2})p(h) \int_{-h}^{0} \frac{x+2h}{2h} dx \geq \left[C_1 h - C_2 h^2 (h+\sqrt{2})\right] p(h).$$

The relation $h = n^{-1/2}$ then implies that $\liminf_{n\to\infty} n^{1/2} \varepsilon_n^{\mathrm{loc}}(0,0) \geq C_1 p(0) > 0$. □

Remark 4.15. In [13, Proposition 9.8] it is stated that the rate for the local error is h (i.e. $n^{-1/2}$) instead of h^2 (i.e. n^{-1}) whenever the terminal condition g has a discontinuity at a non-lattice point $x \notin \mathbb{Z}^h$. By contrast, Proposition 4.10 implies that only the jumps that occur at even lattice points contribute to the error. This discrepancy is a result of the choice of different step functions: In [13], only step functions of the the type $\tilde{\mathbb{1}}_{[a,\infty)} := \mathbb{1}_{(a,\infty)} + \frac{1}{2}\mathbb{1}_{\{a\}}$ are considered.

4.5 The local error for $g \in \mathbf{C}^{0,\alpha}_{\exp}$

A function $g : \mathbb{R} \to \mathbb{R}$ is called locally α-Hölder continuous (write $g \in C^{0,\alpha}_{\mathrm{loc}}$), if for each compact $K \subset \mathbb{R}$

$$\sup_{x,y\in K,\, x\neq y} \frac{|g(x)-g(y)|}{|x-y|^\alpha} < \infty.$$

The class $C^{0,\alpha}_{\exp}$ (see Definition 2.1) consists of all locally α-Hölder continuous functions with exponentially bounded Hölder constants in the sense of (2.6). In fact, $C^{0,\alpha}_{\exp} \subset C^{0,\alpha}_{\mathrm{loc}} \cap \mathcal{B}_{\exp}$, $\alpha \in (0,1]$, and this inclusion is strict at least for $\alpha = 1$: The function $f(x) = \sin(e^{x^2} - 1)$ belongs to $C^{0,1}_{\mathrm{loc}} \cap \mathcal{B}_{\exp}$, whereas $f \notin C^{0,1}_{\exp}$, since $f' \notin \mathcal{B}_{\exp}$.

Recall that $p = p(\,\cdot\,,\theta)$ denotes the density of X_θ and that $\varepsilon^{\mathrm{loc}}_{h,\theta}(g) = \mathbb{E}[g(X_{\tau_J}) - g(X_\theta)]$.

Proposition 4.16. *Let $(h,\theta) \in (0,\infty)\times(0,T]$. Suppose that $g \in C^{0,\alpha}_{exp}$ and that $A, \beta \geq 0$ are as in (2.6). Then, for every $x_0 \in \mathbb{R}$ it holds that*

$$\left|\mathbb{E}[g(x_0+X_{\tau_J}) - g(x_0+X_\theta)]\right| \leq 2^{3+\alpha} h^\alpha A e^{2\beta h + \beta|x_0| + \beta^2\sigma^2\theta/2}.$$

Proof. The property $g \in C^{0,\alpha}_{\exp}$ implies that both g and $g^{x_0} = g(x_0 + \cdot\,)$ belong to $\mathcal{B}_{\exp}$, and

$$\left|\varepsilon^{\mathrm{loc}}_{h,\theta}(g^{x_0})\right| \leq \mathbb{E}|\mathit{\Pi}_e g^{x_0}(X_\theta) - g^{x_0}(X_\theta)| + \mathbb{E}|\mathit{\Pi}_o \mathit{\Pi}_e g^{x_0}(X_\theta) - \mathit{\Pi}_e g^{x_0}(X_\theta)| \tag{4.28}$$

holds by Proposition 4.2. Moreover, whenever $x \in [2kh, (2k+2)h]$ for some $k \in \mathbb{Z}$,

$$\begin{aligned} |\mathit{\Pi}_e g^{x_0}(x) - g^{x_0}(x)| &\leq \frac{(2k+2)h - x}{2h} |g^{x_0}(2kh) - g^{x_0}(x)| + \frac{x-2kh}{2h} |g^{x_0}((2k+2)h) - g^{x_0}(x)| \\ &\leq 2^\alpha h^\alpha A e^{\beta|x_0| + 2\beta h + \beta|x|} \end{aligned}$$

since $g \in C^{0,\alpha}_{\exp}$ and $|2kh| \vee |(2k+2)h| \leq 2h + |x|$. Hence, by $\mathbb{E}\left[e^{\beta|X_\theta|}\right] \leq 2e^{\beta^2\sigma^2\theta/2}$,

$$\begin{aligned} \mathbb{E}\,|\mathit{\Pi}_e g^{x_0}(X_\theta) - f(X_\theta)| &\leq 2^\alpha h^\alpha A e^{\beta|x_0| + 2\beta h} \sum_{k=-\infty}^{\infty} \int_{2kh}^{(2k+2)h} e^{\beta|x|} p(x) dx \\ &\leq 2^{1+\alpha} h^\alpha A e^{\beta|x_0| + 2\beta h + \beta^2\sigma^2\theta/2}. \end{aligned} \tag{4.29}$$

For the remaining expectation on the right-hand side of (4.28), observe that if $y, z \in [2mh, (2m+2)h]$ for given $m \in \mathbb{Z}$, then

$$\begin{aligned} |\mathit{\Pi}_e g^{x_0}(y) - \mathit{\Pi}_e g^{x_0}(z)| &= \left|\frac{z-y}{2h} g^{x_0}(2mh) - \frac{z-y}{2h} g^{x_0}((2m+2)h)\right| \\ &\leq 2^\alpha h^\alpha A e^{\beta(|x_0| + |2mh| \vee |(2m+2)h|)}. \end{aligned} \tag{4.30}$$

Therefore, for $x \in [2kh, (2k+2)h]$ with $k \in \mathbb{Z}$, by (4.30) it holds that

$$
\begin{aligned}
|\Pi_o \Pi_e g^{x_0}(x) - \Pi_e g^{x_0}(x)| &= \frac{(2k+1)h - x}{2h} |\Pi_e g^{x_0}((2k-1)h) - \Pi_e g^{x_0}(x)| \\
&\quad + \frac{x - (2k-1)h}{2h} |\Pi_e g^{x_0}((2k+1)h) - \Pi_e g^{x_0}(x)| \\
&\leq \frac{(2k+1)h - x}{2h} A e^{\beta|x_0|} \left[e^{\beta(|(2k-2)h| \vee |2kh|)} + e^{\beta(|2kh| \vee |(2k+2)h|)} \right] 2^\alpha h^\alpha \\
&\quad + \frac{x - (2k-1)h}{2h} A e^{\beta(|x_0| + |2kh| \vee |(2k+2)h|)} 2^\alpha h^\alpha \\
&\leq 2^{\alpha+1} h^\alpha A e^{\beta(|x_0| + |(2k-2)h| \vee |(2k+2)h|)}.
\end{aligned}
$$

Using the symmetry (in k) of this upper bound, we obtain

$$
\begin{aligned}
&\mathbb{E}|\Pi_o \Pi_e g^{x_0}(X_\theta) - \Pi_e g^{x_0}(X_\theta)| \\
&\leq 2^{\alpha+2} h^\alpha A \sum_{k=0}^{\infty} \int_{2kh}^{(2k+2)h} e^{\beta(|x_0| + |(2k-2)h| \vee |(2k+2)h|)} p(x) dx \\
&\leq 2^{\alpha+2} h^\alpha A e^{\beta|x_0| + 2\beta h} \sum_{k=0}^{\infty} \int_{2kh}^{(2k+2)h} e^{\beta x} p(x) dx \\
&\leq 2^{\alpha+2} h^\alpha A e^{\beta|x_0| + 2\beta h + \beta^2 \sigma^2 \theta / 2}.
\end{aligned} \tag{4.31}
$$

The claim follows by applying the estimates (4.29) and (4.31) to (4.28). □

Corollary 4.17. *Let $n \in 2\mathbb{N}$. Suppose that $g \in C^{0,\alpha}_{exp}$ and that $\beta \geq 0$ is as in (2.6). Then, under Assumption 4.11, there exist a constant $C > 0$ such that for all $(t_0, x_0) \in [0, T) \times \mathbb{R}$,*

$$
\left|\varepsilon_n^{loc}(t_0, x_0)\right| \leq \frac{C \sigma^\alpha T^{\alpha/2}}{n^{\alpha/2}} e^{\beta|x_0| + 2\beta^2 \sigma^2 T}.
$$

Proof. Since $h = \sigma\sqrt{\frac{T}{n}} \leq \sigma\sqrt{\frac{T}{2}}$ and $\varepsilon_n^{\text{loc}}(t_0, x_0) = \varepsilon_{h,\theta_n}^{\text{loc}}(g^{x_0})$ by Assumption 4.11 and (2.12), Proposition 4.16 implies the result. □

5 The global error

Our aim is to derive an upper bound for the global error

$$
\varepsilon_n^{\text{glob}}(t_0, x_0) = \mathbb{E}[g(x_0 + X_{\tau_{n_\theta}}) - g(x_0 + X_{\tau_{J_n}})]
$$

defined in (2.11), where g is an exponentially bounded Borel function and $(X_{\tau_k})_{k=0,1,\ldots}$ is the random walk considered in Subsection 2.1. For this purpose, we need a collection of estimates related to the behavior of the random walk (X_{τ_k}) and the stopping time J_n. A part of these are given in this section, while the more involved ones are presented later in Section 6.

Note: *The Assumption 4.11 is taken as a standing assumption throughout Section 5.*

Recall the definitions of n_θ and θ_n given in (2.4). Recall also that $J_n(\omega) = \inf\{2m \in 2\mathbb{N} : \tau_{2m}(\omega) > \theta_n\}$ as was defined in (2.9). A result similar to the lemma below was proved in [13, Corollary 11.4].

Lemma 5.1. *For any $b \geq 0$, it holds that*

$$
(i) \qquad \mathbb{E}\left[e^{b|X_{\tau_{n_\theta}}|}\right] \leq 2e^{b^2 \sigma^2 T/2}, \tag{5.1}
$$

$$(ii) \qquad \mathbb{E}\big[e^{b|X_{\tau_{J_n}}|}\big] \le 2e^{b\sigma\sqrt{2T}+b^2\sigma^2T/2}. \tag{5.2}$$

Proof. (i): Since $X_{\tau_{n_\theta}} = \sum_{k=1}^{n_\theta} \Delta X_{\tau_k}$, where $(\Delta X_{\tau_k})_{k=1,2,\ldots}$ is a sequence of i.i.d. random variables with $\mathbb{P}(\Delta X_{\tau_k} = \pm h) = 1/2$ for $h = \sigma\sqrt{\frac{T}{n}}$ (see Subsection 2.1),

$$\mathbb{E}\big[e^{b|X_{\tau_{n_\theta}}|}\big] \le 2\mathbb{E}\big[e^{bX_{\tau_{n_\theta}}}\big] = 2\left(\mathbb{E}\big[e^{b\Delta X_{\tau_1}}\big]\right)^{n_\theta} = 2\left(\cosh(bh)\right)^{n_\theta} \le 2e^{b^2h^2n_\theta/2} \le 2e^{b^2\sigma^2T/2}$$

by the inequality $\cosh(x) \le e^{x^2/2}$, $x \in \mathbb{R}$.

(ii): Firstly, observe that by the definition of J_n we have $\big|X_{\tau_{J_n}} - X_{\theta_n}\big| \le 2h$. Secondly, since for a standard normal Z random variable it holds that $\mathbb{E}\big[e^{u|Z|}\big] \le 2e^{u^2/2}$ $(u \in \mathbb{R})$,

$$\mathbb{E}\big[e^{b|X_{\tau_{J_n}}|}\big] \le \mathbb{E}\big[e^{b|X_{\tau_{J_n}}-X_{\theta_n}|+b|X_{\theta_n}|}\big] \le e^{2bh}\mathbb{E}\big[e^{b\sigma\sqrt{\theta_n}|Z|}\big] \le 2e^{b\sigma\sqrt{2T}+b^2\sigma^2T/2}.$$

□

In Proposition 5.2, we present some more upper bounds which are used to estimate the global error.

Proposition 5.2.

(i) *Suppose that $p \ge 0$, $g \in \mathcal{B}_{exp}$, and that $b \ge 0$ is as in (3.1). Then there exists a constant $C_p > 0$ such that for all $x_0 \in \mathbb{R}$,*

$$\sup_{(n,t_0)\in 2\mathbb{N}\times[0,T)} \left|\mathbb{E}\left[\left(\frac{|X_{\tau_{n_\theta}}|}{\sqrt{\sigma^2\theta_n}}\right)^p g(x_0+X_{\tau_{n_\theta}})\right]\right| \le C_p e^{b|x_0|+b^2\sigma^2T}. \tag{5.3}$$

Moreover, for every $p > 0$ there exists a constant $C_p > 0$ such that

$$(ii) \qquad \sup_{(n,t_0)\in 2\mathbb{N}\times[0,T)} n_\theta^p \mathbb{P}\big(\big|X_{\tau_{n_\theta}}/h\big| > n_\theta^{3/5}\big) \le C_p, \tag{5.4}$$

$$(iii) \qquad \sup_{(n,t_0)\in 2\mathbb{N}\times[0,T)} n_\theta^p \mathbb{P}\big(\,|J_n - n_\theta| > n_\theta^{3/5}\big) \le C_p. \tag{5.5}$$

Proof. (i): Observe that

$$S_{n_\theta} := \frac{X_{\tau_{n_\theta}}}{\sqrt{\sigma^2\theta_n}} = \frac{1}{\sqrt{\sigma^2\theta_n}}\sum_{k=1}^{n_\theta}\Delta X_{\tau_k} \overset{d}{=} \frac{1}{\sqrt{n_\theta}}\sum_{k=1}^{n_\theta}\xi_i,$$

where $(\xi_i)_{i=1,2,\ldots}$ is an i.i.d. Rademacher sequence (see Subsection 2.1). Hence,

$$\mathbb{E}\left[e^{tS_{n_\theta}}\right] = \big(\cosh(\tfrac{t}{\sqrt{n_\theta}})\big)^{n_\theta} \le \big(e^{t^2/(2n_\theta)}\big)^{n_\theta} = e^{t^2/2}, \quad t \in \mathbb{R}.$$

Consequently, by the symmetricity of S_{n_θ} and Markov's inequality,

$$\mathbb{P}(|S_{n_\theta}| > t) = 2\mathbb{P}(e^{tS_{n_\theta}} > e^{t^2}) \le 2e^{-t^2}\mathbb{E}\left[e^{tS_{n_\theta}}\right] \le 2e^{-t^2/2}, \quad t > 0,$$

and thus, uniformly in (n, t_0), for $p > 0$,

$$\mathbb{E}|S_{n_\theta}|^p = p\int_0^\infty t^{p-1}\mathbb{P}(|S_{n_\theta}| > t)dt \le 2p\int_0^\infty t^{p-1}e^{-t^2/2}dt := \tilde{C}_p < \infty. \tag{5.6}$$

Hölder's inequality, (5.6), and (5.1) then imply that

$$\left|\mathbb{E}\left[|S_{n_\theta}|^p\, g(x_0+X_{\tau_{n_\theta}})\right]\right| \le Ae^{b|x_0|}\left(\mathbb{E}|S_{n_\theta}|^{2p}\right)^{1/2}\left(\mathbb{E}\big[e^{2b|X_{\tau_{n_\theta}}|}\big]\right)^{1/2} \le 2A\tilde{C}_{2p}^{1/2}e^{b|x_0|+b^2\sigma^2T}.$$

This proves (5.3) for $p>0$, and the case $p=0$ can be seen from the last line as well.

(ii): Since $h\sqrt{n_\theta}=\sqrt{\sigma^2\theta_n}$, by Markov's inequality and (5.6) we obtain

$$\mathbb{P}\big(|X_{\tau_{n_\theta}}/h|>n_\theta^{3/5}\big)=\mathbb{P}\big(\,|S_{n_\theta}|>n_\theta^{1/10}\big)\le\mathbb{E}\,|S_{n_\theta}|^q\,n_\theta^{-q/10}\le C_q n_\theta^{-q/10} \tag{5.7}$$

for all $q>0$. Choose $q\ge 10p$ and multiply both sides of (5.7) by n_θ^p to obtain (5.4).

(iii): For every $K>0$, Markov's inequality and Proposition 6.3 below imply that

$$\mathbb{P}\big(\,|J_n-n_\theta|>n_\theta^{3/5}\big)\le\mathbb{E}\,|J_n-n_\theta|^K\,n_\theta^{-3K/5}\le C_K n_\theta^{-K/10} \tag{5.8}$$

for some constant $C_K>0$. For given $p>0$, it remains to choose $K\ge 10p$ and multiply both sides of (5.8) by n_θ^p. □

The proof of the main result of this section follows closely the proof of [13, Theorem 8.1].

Proposition 5.3. *Let $n\in 2\mathbb{N}$. Suppose that $g\in\mathcal{B}_{exp}$ and that $b\ge 0$ is as in (3.1). Then there exists a constant $C>0$ such that for all $(t_0,x_0)\in[0,T)\times\mathbb{R}$,*

$$\big|\varepsilon_n^{glob}(t_0,x_0)\big|\le\frac{CT}{n(T-t_k^n)}e^{b|x_0|+3b^2\sigma^2T},\quad t_0\in[t_k^n,t_{k+1}^n),\quad 0\le k<\tfrac{n}{2}. \tag{5.9}$$

Proof. Define a set

$$\Gamma_{n_\theta}:=\big\{\big|X_{\tau_{n_\theta}}/h\big|\vee|J_{n_\theta}-n_\theta|\le n_\theta^{3/5}\big\} \tag{5.10}$$

and decompose the error $\varepsilon_n^{\rm glob}(t_0,x_0)$ into the sum of expectations $E^{(1)}$ and $E^{(2)}$, where

$$E^{(1)}:=\mathbb{E}[g(x_0+X_{\tau_{n_\theta}})-g(x_0+X_{\tau_{J_n}});\Gamma_{n_\theta}],\quad E^{(2)}:=\mathbb{E}[g(x_0+X_{\tau_{n_\theta}})-g(x_0+X_{\tau_{J_n}});\Gamma_{n_\theta}^{\complement}]. \tag{5.11}$$

Using the estimates of Lemma 5.1 and Proposition 5.2, it can be shown that

$$\big|E^{(2)}\big|\le\tilde{C}_0 n_\theta^{-3/2}e^{b|x_0|+b^2\sigma^2T+b\sigma\sqrt{2T}} \tag{5.12}$$

for some constant $\tilde{C}_0>0$; this is done in Lemma A.3 (i). Estimation of $\big|E^{(1)}\big|$ requires more subtlety. Denote the probability mass functions of $X_{\tau_{n_\theta+k}}/h$ and J_n-n_θ by

$$P_{n_\theta+k}(x):=\mathbb{P}(X_{\tau_{n_\theta+k}}=hx)\quad\text{and}\quad P^J_{n_\theta}(x):=\mathbb{P}(J_n-n_\theta=x),\qquad x\in\mathbb{Z}. \tag{5.13}$$

By Lemma A.3 (ii), there exists a constant $\tilde{C}_1>0$ such that

$$\begin{aligned}\big|E^{(1)}\big|\le&\bigg|\sum_{k=2-n_\theta}^{\infty}\sum_{x=-n_\theta}^{n_\theta}g(x_0+xh)P^J_{n_\theta}(k)P_{n_\theta}(x)\Big(\frac{k}{2n_\theta}-\frac{3k^2+4kx^2}{8n_\theta^2}+\frac{3k^2x^2}{4n_\theta^3}-\frac{k^2x^4}{8n_\theta^4}\Big)\bigg|\\&+\tilde{C}_1n_\theta^{-3/2}e^{b|x_0|+b^2\sigma^2T}.\end{aligned} \tag{5.14}$$

Next, we use relation (A.5) in order to rewrite the double sum on the right-hand side of (5.14) as

$$\begin{aligned}E^{(3)}:=&\sum_{k=2-n_\theta}^{\infty}\sum_{x=-n_\theta}^{n_\theta}g(x_0+xh)P^J_{n_\theta}(k)P_{n_\theta}(x)\Big(\frac{k}{2n_\theta}-\frac{3k^2+4kx^2}{8n_\theta^2}+\frac{3k^2x^2}{4n_\theta^3}-\frac{k^2x^4}{8n_\theta^4}\Big)\\=&\frac{1}{n_\theta}\bigg\{\frac12\mathbb{E}\left[g(x_0+X_{\tau_{n_\theta}})\right]\mathbb{E}[J_n-n_\theta]-\frac38\mathbb{E}\left[g(x_0+X_{\tau_{n_\theta}})\right]\mathbb{E}\left(\tfrac{J_n-n_\theta}{\sqrt{n_\theta}}\right)^2\\&-\frac12\mathbb{E}\Big[\Big(\tfrac{X_{\tau_{n_\theta}}}{\sqrt{\sigma^2\theta_n}}\Big)^2g(x_0+X_{\tau_{n_\theta}})\Big]\mathbb{E}[J_n-n_\theta]+\frac34\mathbb{E}\Big[\Big(\tfrac{X_{\tau_{n_\theta}}}{\sqrt{\sigma^2\theta_n}}\Big)^2g(x_0+X_{\tau_{n_\theta}})\Big]\mathbb{E}\left(\tfrac{J_n-n_\theta}{\sqrt{n_\theta}}\right)^2\end{aligned}$$

$$-\frac{1}{8}\mathbb{E}\left[\left(\frac{X_{\tau_{n_\theta}}}{\sqrt{\sigma^2\theta_n}}\right)^4 g(x_0+X_{\tau_{n_\theta}})\right]\mathbb{E}\left(\frac{J_n-n_\theta}{\sqrt{n_\theta}}\right)^2\Bigg\}$$

$$=\frac{1}{n_\theta}\Bigg\{\mathbb{E}\left[g(x_0+X_{\tau_{n_\theta}})\right]\left(\frac{1}{2}\mathbb{E}[J_n-n_\theta]-\frac{3}{8}\mathbb{E}\left(\frac{J_n-n_\theta}{\sqrt{n_\theta}}\right)^2\right)$$
$$+\mathbb{E}\Big[\left(\frac{X_{\tau_{n_\theta}}}{\sqrt{\sigma^2\theta_n}}\right)^2 g(x_0+X_{\tau_{n_\theta}})\Big]\left(\frac{3}{4}\mathbb{E}\left(\frac{J_n-n_\theta}{\sqrt{n_\theta}}\right)^2-\frac{1}{2}\mathbb{E}[J_n-n_\theta]\right)$$
$$-\frac{1}{8}\mathbb{E}\left[\left(\frac{X_{\tau_{n_\theta}}}{\sqrt{\sigma^2\theta_n}}\right)^4 g(x_0+X_{\tau_{n_\theta}})\right]\mathbb{E}\left(\frac{J_n-n_\theta}{\sqrt{n_\theta}}\right)^2\Bigg\}. \tag{5.15}$$

By Proposition 6.1, there exist constants $c_1, c_2 > 0$ such that $\left|\mathbb{E}[J_n-n_\theta]-\frac{4}{3}\right| \le \frac{c_1}{\sqrt{n_\theta}}$ and $\mathbb{E}\big[\big(\frac{J_n-n_\theta}{\sqrt{n_\theta}}\big)^2-\frac{2}{3}\big] \le \frac{c_2}{\sqrt{n_\theta}}$, and thus

$$\left|\frac{1}{2}\mathbb{E}[J_n-n_\theta]-\frac{3}{8}\mathbb{E}\left(\frac{J_n-n_\theta}{\sqrt{n_\theta}}\right)^2-\frac{5}{12}\right| \le \frac{c_1+c_2}{\sqrt{n_\theta}}, \quad \left|\frac{1}{8}\mathbb{E}\left(\frac{J_n-n_\theta}{\sqrt{n_\theta}}\right)^2-\frac{1}{12}\right| \le \frac{c_2}{\sqrt{n_\theta}} \quad \text{and}$$
$$\left|\frac{3}{4}\mathbb{E}\left(\frac{J_n-n_\theta}{\sqrt{n_\theta}}\right)^2-\frac{1}{2}\mathbb{E}[J_n-n_\theta]+\frac{1}{6}\right| \le \frac{c_1+c_2}{\sqrt{n_\theta}}.$$

Consequently, by (5.15) and (5.3), there exist constants $\tilde{C}_2, \tilde{C}_3 > 0$ such that

$$\left|E^{(3)}\right| \le \frac{5}{12n_\theta}\left|\mathbb{E}\left[g(x_0+X_{\tau_{n_\theta}})\right]\right| + \frac{1}{6n_\theta}\left|\mathbb{E}\Big[\left(\frac{X_{\tau_{n_\theta}}}{\sqrt{\sigma^2\theta_n}}\right)^2 g(x_0+X_{\tau_{n_\theta}})\Big]\right|$$
$$+\frac{1}{12n_\theta}\left|\mathbb{E}\Big[\left(\frac{X_{\tau_{n_\theta}}}{\sqrt{\sigma^2\theta_n}}\right)^4 g(x_0+X_{\tau_{n_\theta}})\Big]\right| + \frac{\tilde{C}_2 e^{b|x_0|+b^2\sigma^2 T}}{n_\theta^{3/2}}$$
$$\le \frac{\tilde{C}_3}{n_\theta}e^{b|x_0|+b^2\sigma^2 T} + \frac{\tilde{C}_2 e^{b|x_0|+b^2\sigma^2 T}}{n_\theta^{3/2}}. \tag{5.16}$$

To complete the proof, it remains to observe that $\frac{1}{n_\theta^{3/2}} \le \frac{1}{\sqrt{2}}\frac{1}{n_\theta}$, to combine (5.11), (5.12), (5.14), and (5.16), and to recall that $n_\theta T = n(T-t_k^n)$ for $t_0 \in [t_k^n, t_{k+1}^n)$. □

6 Moment estimates for the stopping time J_n

In this section we present moment estimates for the random variable $J_n = \inf\{2m \in 2\mathbb{N} : \tau_{2m} > \theta_n\}$ introduced in (2.9), which are used for the estimation of the global error in Section 5.

Proposition 6.1. *Suppose that Assumption 4.11 holds. Then there exists a constant $C > 0$ such that for all $(n, t_0) \in 2\mathbb{N}\times[0, T)$,*

$$(i) \quad \left|\mathbb{E}[J_n]-n_\theta-\frac{4}{3}\right| \le \frac{C}{\sqrt{n_\theta}}, \qquad (ii) \quad \left|\mathbb{E}\left(\frac{J_n-n_\theta}{\sqrt{n_\theta}}\right)^2-\frac{2}{3}\right| \le \frac{C}{\sqrt{n_\theta}}.$$

The proof of Proposition 6.1 is given in the arXiv version of this paper; see [9, Subsection 6.1]. To derive an estimate for $\mathbb{E}\left|J_n-n_\theta\right|^K$ for any $K > 0$, we recall (see e.g. [2, Theorem 14.12]) a version of the Azuma–Hoeffding inequality.

Proposition 6.2 (Azuma–Hoeffding inequality)**.** *Suppose that $(M_j)_{j=0,1,\ldots}$ is a martingale with $M_0 = 0$. In addition, assume that for all $i \ge 1$ there exists a constant $\alpha_i > 0$ such that $|M_i - M_{i-1}| \le \alpha_i$ a.s. Then, for all $k \in \mathbb{N}$ and every $t > 0$,*

$$\mathbb{P}(M_k \ge t) \le \exp\left(-\frac{t^2}{2\sum_{j=1}^k \alpha_i^2}\right).$$

For $t_0 = 0$, the following statement can be found also in [13, Proposition 11.2 (iv)].

Proposition 6.3. *Suppose that Assumption 4.11 holds, and let $K > 0$. Then there exists a constant $C_K > 0$ depending at most on K such that*

$$\mathbb{E}\,|J_n - n_\theta|^K \leq C_K n_\theta^{K/2} \quad \textit{for all } (n, t_0) \in 2\mathbb{N} \times [0, T). \tag{6.1}$$

Proof. It suffices to prove the claim for $K \geq 2$, since the case $K \in (0, 2)$ then follows by Jensen's inequality. Since $|J_n - n_\theta|$ is a non-negative random variable,

$$\frac{1}{K}\mathbb{E}\,|J_n - n_\theta|^K = \int_0^\infty z^{K-1}\mathbb{P}(|J_n - n_\theta| > z)dz.$$

We show that there exist constants $C_K^{(1)}, C_K^{(2)}, C_K^{(3)} > 0$ corresponding to the sets $A_1 = (0, 2]$, $A_2 = (2, n_\theta]$ and $A_3 = (n_\theta, \infty)$ such that

$$I_k(n_\theta) := \int_{A_k} z^{K-1}\mathbb{P}(|J_n - n_\theta| > z)dz \leq C_K^{(k)} n_\theta^{K/2} \quad \text{for all } n_\theta. \tag{6.2}$$

Step 1: Since $K \geq 2$ and $n_\theta \geq 2$, we have that

$$I_1(n_\theta) = \int_0^2 z^{K-1}\mathbb{P}(|J_n - n_\theta| > z)dz \leq \int_0^2 z^{K-1}dz \leq 2^K/K \leq C_K^{(1)} n_\theta^{K/2}.$$

Step 2: Suppose that $n_\theta > 2$ and define $\delta_{n_\theta}(u) := \frac{2}{n_\theta}\lfloor \frac{n_\theta u}{2} \rfloor$. Then

$$\begin{aligned} I_2(n_\theta) &= \int_2^{n_\theta} z^{K-1}\mathbb{P}(|J_n - n_\theta| > z)dz = n_\theta^K \int_{2/n_\theta}^1 u^{K-1}\mathbb{P}(|J_n - n_\theta| > n_\theta u)du \\ &\leq n_\theta^K \int_{2/n_\theta}^1 u^{K-1}\mathbb{P}(|J_n - n_\theta| > \delta_{n_\theta}(u) n_\theta)du. \end{aligned} \tag{6.3}$$

Fix a constant $a \in (0, 1]$ small enough such that for every $m \in \mathbb{N}$,

$$\delta_m(u) < \tfrac{\pi^2}{12+\pi^2} \quad \text{and} \quad H\Big(\sqrt{\tfrac{3\delta_m(u)}{1+\delta_m(u)}}\Big) \wedge H\Big(\sqrt{\tfrac{3\delta_m(u)}{1-\delta_m(u)}}\Big) > 1/4 \quad \text{hold for all} \quad u \leq a, \tag{6.4}$$

where the function H is defined below in (A.12). Depending on the value of n_θ, we split the right-hand side of (6.3) into the sum of the integrals

$$\begin{aligned} I_{2,1}(n_\theta) &:= n_\theta^K \int_{2/n_\theta}^a u^{K-1}\mathbb{P}(|J_n - n_\theta| > \delta_{n_\theta}(u) n_\theta)du \quad (\text{for } a > 2/n_\theta, \text{ otherwise } 0), \\ I_{2,2}(n_\theta) &:= n_\theta^K \int_{a \vee (2/n_\theta)}^1 u^{K-1}\mathbb{P}(|J_n - n_\theta| > \delta_{n_\theta}(u) n_\theta)du. \end{aligned}$$

If $a \in (2/n_\theta, 1)$, by (6.4) and the fact that $n_\theta(1 + \delta_{n_\theta}(u))$ and $n_\theta(1 - \delta_{n_\theta}(u))$ are (even) integers, we may apply Lemma A.6 and estimate

$$\begin{aligned} I_{2,1}(n_\theta) &= n_\theta^K \int_{2/n_\theta}^a u^{K-1}\Big[\mathbb{P}\big(J_n > n_\theta(1 + \delta_{n_\theta}(u))\big) + \mathbb{P}\big(J_n < n_\theta(1 - \delta_{n_\theta}(u))\big)\Big]du \\ &\leq n_\theta^K \int_{2/n_\theta}^a u^{K-1}\left[\exp\left(-\tfrac{3}{8}\tfrac{n_\theta \delta_{n_\theta}^2(u)}{1+\delta_{n_\theta}(u)}\right) + \exp\left(-\tfrac{3}{8}\tfrac{n_\theta \delta_{n_\theta}^2(u)}{1-\delta_{n_\theta}(u)}\right)\right] du \\ &\leq 2n_\theta^K \int_{2/n_\theta}^a u^{K-1} \exp\left(-\tfrac{3}{8}\tfrac{n_\theta \delta_{n_\theta}^2(u)}{1+\delta_{n_\theta}(u)}\right) du. \end{aligned} \tag{6.5}$$

By the properties of the floor function, for $u \in (0,1]$ it holds that

$$\frac{\delta_{n_\theta}^2(u)}{1+\delta_{n_\theta}(u)} = \frac{\left(\frac{2}{n_\theta}\lfloor\frac{n_\theta u}{2}\rfloor\right)^2}{1+\frac{2}{n_\theta}\lfloor\frac{n_\theta u}{2}\rfloor} \geq \frac{\left(\frac{2}{n_\theta}\left(\frac{n_\theta u}{2}-1\right)\right)^2}{1+u} > \frac{\left(u-\frac{2}{n_\theta}\right)^2}{2}, \tag{6.6}$$

and thus the right-hand side of (6.5) can be bounded from above by

$$\begin{aligned} 2n_\theta^K \int_{2/n_\theta}^{a} u^{K-1} e^{-\frac{3n_\theta(u-2/n_\theta)^2}{16}} du &\leq 2n_\theta^K \int_0^{1-2/n_\theta} \left(u+\tfrac{2}{n_\theta}\right)^{K-1} e^{-\frac{3n_\theta u^2}{16}} du \\ &\leq 2^{K-1} n_\theta^K \int_0^{1-2/n_\theta} \left(u^{K-1}+(\tfrac{2}{n_\theta})^{K-1}\right) e^{-\frac{3n_\theta u^2}{16}} du. \end{aligned} \tag{6.7}$$

By substituting $x = u\sqrt{n_\theta}$ and identifying the right-hand side of (6.7) as an integral with respect to a Gaussian measure, it can be verified that this integral multiplied by $n_\theta^{K/2}$ is bounded by some constant $\tilde{C}_K^{(2,1)} > 0$. Hence, $I_{2,1}(n_\theta) \leq C_K^{(2,1)} n_\theta^{K/2}$ for all n_θ, where $C_K^{(2,1)} > 0$ depends only on K.

Let us then consider the integral $I_{2,2}(n_\theta)$. If $a/3 \leq 2/n_\theta$, then $n_\theta \leq 6/a$, $a \leq a \vee (2/n_\theta)$, and thus $I_{2,2}(n_\theta) \leq 6^K (Ka^K)^{-1}$. On the other hand, if $a/3 \in (2/n_\theta, 1)$, by Lemma A.6,

$$\begin{aligned} I_{2,2}(n_\theta) &= n_\theta^K \int_a^1 u^{K-1}\Big[\mathbb{P}\big(J_n > n_\theta(1+\delta_{n_\theta}(u))\big) + \mathbb{P}\big(J_n < n_\theta(1-\delta_{n_\theta}(u))\big)\Big] du \\ &\leq n_\theta^K \int_a^1 u^{K-1}\Big[\mathbb{P}\big(J_n > n_\theta(1+\delta_{n_\theta}(a/3))\big) + \mathbb{P}\big(J_n < n_\theta(1-\delta_{n_\theta}(a/3))\big)\Big] du \\ &\leq n_\theta^K \Big[\exp\left(-\tfrac{3}{8}\tfrac{n_\theta \delta_{n_\theta}(a/3)^2}{1+\delta_{n_\theta}(a/3)}\right) + \exp\left(-\tfrac{3}{8}\tfrac{n_\theta \delta_{n_\theta}(a/3)^2}{1-\delta_{n_\theta}(a/3)}\right)\Big] \int_a^1 u^{K-1} du \\ &\leq 2n_\theta^K \exp\left(-\tfrac{3}{8}\tfrac{n_\theta \delta_{n_\theta}(a/3)^2}{1+\delta_{n_\theta}(a/3)}\right) \int_a^1 u^{K-1} du \\ &\leq 2n_\theta^K \exp\left(-\tfrac{3n_\theta(a/3-2/n_\theta)^2}{16}\right). \end{aligned}$$

Here we used the fact that $u \mapsto \delta_{n_\theta}(u)$ is nondecreasing, that $n_\theta(1+\delta_{n_\theta}(a/3))$ and $n_\theta(1-\delta_{n_\theta}(a/3))$ are (even) integers, condition (6.4), and inequality (6.6). Notice that the right-hand side converges to zero as $n_\theta \to \infty$. Consequently, there exists a constant $C_K^{(2,2)} > 0$ such that $I_{2,2}(n_\theta) \leq C_K^{(2,2)}$ for all n_θ, and (6.2) for $k = 2$ follows.

Step 3: To estimate $I_3(n_\theta)$, we apply the Azuma–Hoeffding inequality to the tail distribution of the random variable $J_n = \inf\{2m \in 2\mathbb{N} : \tau_{2m} > \theta_n\}$. Recall that $\tau_i - \tau_{i-1}$, $i = 1, 2, \dots$, are i.i.d. (see Subsection 2.1) and that $\frac{n}{T}\theta_n = n_\theta$ according to (2.4). Let $\zeta_i := \frac{n}{T}(\tau_i - \tau_{i-1})$, $i \geq 1$. Then, for all $m \in \mathbb{N}$, we have

$$\begin{aligned} \mathbb{P}(J_n \geq 2m) = \mathbb{P}(\tau_{2m-2} \leq \theta_n) &= \mathbb{P}\Big(\sum_{i=1}^{2m-2} \zeta_i \leq \tfrac{n}{T}\theta_n\Big) \leq \mathbb{P}\Big(\sum_{i=1}^{2m-2} \zeta_i \wedge N \leq n_\theta\Big) \\ &= \mathbb{P}\Big(\sum_{i=1}^{2m-2} (c_N - \zeta_i \wedge N) \geq (2m-2)c_N - n_\theta\Big), \end{aligned} \tag{6.8}$$

where $N \in \mathbb{N}$ is chosen such that $3/4 < c_N := \mathbb{E}[\zeta_i \wedge N] < \mathbb{E}[\zeta_i] = 1$. Then $|\mathbb{E}[\zeta_i \wedge N] - \zeta_i \wedge N| \leq N$ for all $i \geq 1$, and by (6.8) and the Azuma–Hoeffding inequality (Proposition 6.2),

$$\mathbb{P}(J_n \geq 2m) \leq \exp\left(-\tfrac{((2m-2)c_N - n_\theta)^2}{2(2m-2)N^2}\right), \quad m \in \mathbb{N}.$$

Since $J_n > 0$, we have

$$
\begin{aligned}
I_3(n_\theta) &= \int_{n_\theta}^{\infty} z^{K-1}\mathbb{P}\left(J_n - n_\theta \geq z\right) dz \\
&= \int_{n_\theta}^{2n_\theta+2} z^{K-1}\mathbb{P}\left(J_n \geq z + n_\theta\right) dz + \sum_{m=3}^{\infty} \int_{(m-1)n_\theta+2}^{mn_\theta+2} z^{K-1}\mathbb{P}(J_n \geq z + n_\theta) dz \\
&\leq \int_{n_\theta}^{2n_\theta+2} z^{K-1}\mathbb{P}\left(J_n \geq 2n_\theta\right) dz + \sum_{m=3}^{\infty} \int_{(m-1)n_\theta+2}^{mn_\theta+2} z^{K-1}\mathbb{P}(J_n \geq mn_\theta + 2) dz \\
&\leq \int_{n_\theta}^{2n_\theta+2} z^{K-1} e^{-\frac{n_\theta}{2N^2}\frac{((2-2/n_\theta)c_N-1)^2}{(2-2/n_\theta)}} dz + \sum_{m=3}^{\infty} \int_{(m-1)n_\theta+2}^{mn_\theta+2} z^{K-1} e^{-\frac{n_\theta}{2N^2}\frac{(mc_n-1)^2}{m}} dz \\
&=: I_{3,1}(n_\theta) + I_{3,2}(n_\theta).
\end{aligned}
$$

Since $c_N \in (3/4, 1)$, there exist constants $c, c' > 0$ such that

$$
\begin{aligned}
I_{3,1}(n_\theta) &= \int_{n_\theta}^{2n_\theta+2} z^{K-1} e^{-\frac{n_\theta}{2N^2}\frac{((2-2/n_\theta)c_N-1)^2}{(2-2/n_\theta)}} dz \leq 2(2n_\theta+2)^{K-1} e^{-\frac{n_\theta c}{N^2}} \leq C_K^{(3,1)}, \\
I_{3,2}(n_\theta) &= \sum_{m=3}^{\infty} \int_{(m-1)n_\theta+2}^{mn_\theta+2} z^{K-1} e^{-\frac{n_\theta}{2N^2}\frac{(mc_N-1)^2}{m}} dz \leq \sum_{m=3}^{\infty} (mn_\theta+2)^K e^{-\frac{n_\theta m c'}{N^2}} \leq C_K^{(3,2)},
\end{aligned}
$$

where $C_K^{(3,1)}, C_K^{(3,2)} > 0$ depend at most on K. This proves (6.2) for $k = 3$. □

A Appendix

A.1 The class $GBV_{\textbf{exp}}$

For a function $g : \mathbb{R} \to \mathbb{R}$, let

$$
T_g(x) := \sup \left\{ \sum_{i=1}^{N} |g(x_i) - g(x_{i-1})|, \; N \in \mathbb{N}, \; -\infty < x_0 < x_1 < \cdots < x_N = x \right\}.
$$

If $\lim_{x\to\infty} T_g(x) < \infty$, the function g is said to be of bounded variation. The class of functions with this property will be denoted by BV.

A function $g \in BV$ is by definition a bounded function. An error estimation carried out merely for the class BV would rule out e.g. polynomials, which on the other hand have bounded variation on every compact interval. Therefore, instead of the class BV, we consider a class of functions of generalized bounded variation allowing exponential growth. In order to find an applicable representation for a large class of such functions, we will follow the presentation given in [1].

Recall the class $\mathcal{M}$ given by Definition 2.2, which consists of set functions μ (acting on bounded Borel sets on $\mathbb{R}$) that can be written as a difference of two measures $\mu^1, \mu^2 : \mathcal{B}(\mathbb{R}) \to [0, \infty]$ such that $\mu^1(K)$ and $\mu^2(K)$ are finite for all compact sets $K \in \mathcal{B}(\mathbb{R})$. In [1, Theorem 3.3] it is proved that such a decomposition can be chosen to be orthogonal and minimal: There exists a unique pair of measures μ^+, μ^- on $\mathcal{B}(\mathbb{R})$ such that μ^+ and μ^- are mutually singular, and $\mu^+ \leq \mu^1$ and $\mu^- \leq \mu^2$ hold for all the other decompositions $\mu = \mu^1 - \mu^2$. Even though $\mu \in \mathcal{M}$ is not itself a signed measure (it is undefined on unbounded sets), the aforementioned result, based on the Hahn decomposition theorem, allows us to define the total variation measure associated to μ by setting

$$
|\mu| : \mathcal{B}(\mathbb{R}) \to [0, \infty], \quad |\mu| := \mu^+ + \mu^-.
$$

Consequently, the integral appearing in (2.8) is defined.

Some properties of the class GBV_{exp} (recall Definition 2.3) are given below.

Remark A.1.

(i): It holds that $BV \subset GBV_{\text{exp}} \subset \mathcal{B}_{\text{exp}}$. See [1, Theorem 4.3] for the first inclusion. The second inclusion can be seen as follows: Suppose that $g \in GBV_{\text{exp}}$ is given by (2.7). Then (3.1) is satisfied if $b = \beta$ and A is equal to the sum of $|c|$ and the left-hand side of (2.8).

(ii): A function $g \in GBV_{\text{exp}}$ has bounded variation on each compact interval $I \subset \mathbb{R}$.

(iii): Every polynomial $f(x) = \sum_{k=0}^{N} a_k x^k$, $a_k \in \mathbb{R}, N \in \mathbb{N}$, belongs to the class GBV_{exp}. Indeed, the function f admits the representation (2.7) by letting

$$c = a_0, \quad d\mu = \sum_{k=1}^{N} k a_k x^{k-1} dx, \quad \text{and} \quad \mathcal{J} = \emptyset.$$

Notice that this μ satisfies the condition (2.8), since for every $\beta > 0$,

$$\int_{\mathbb{R}} e^{-\beta|x|} d|\mu|(x) = \int_{\mathbb{R}} e^{-\beta|x|} \left| \sum_{k=1}^{N} k a_k x^{k-1} \right| dx < \infty.$$

A.2 Auxiliary results for Section 5

Under Assumption 4.11, let us recall from (5.13) the notation $P_{n_\theta+k}(x) = \mathbb{P}(X_{\tau_{n_\theta+k}} = hx)$ and $P^J_{n_\theta}(x) = \mathbb{P}(J_n - n_\theta = x)$, $x \in \mathbb{Z}$. Notice also that for all $k \in 2\mathbb{N}$,

$$P_k(x) = \binom{k}{\frac{k+x}{2}} 2^{-k}, \qquad x \in 2\mathbb{Z}, \quad |x| \le k.$$

As in [13], we define the 'effective order' of a monomial $\frac{k^p x^q}{n^r}$ with $p, q, r \in \mathbb{N}_0$ to be

$$\breve{O}\left(\frac{k^p x^q}{n^r}\right) := \frac{p+q}{2} - r.$$

We will use the following result from [13] in the proof of Lemma A.3.

Proposition A.2 ([13, Proposition 11.5])**.** *Let*

$$R : \mathcal{D}(R) \to \mathbb{R}, \quad R(n,k,x) := \frac{P_{n+k}(x)}{P_n(x)}; \tag{A.1}$$

$$R^{(1)} : 2\mathbb{N}\times(2\mathbb{Z})^n \to \mathbb{R}, \quad R^{(1)}(n,k,x) := \frac{k}{2n} - \frac{3k^2 + 4kx^2}{8n^2} + \frac{3k^2x^2}{4n^3} - \frac{k^2x^4}{8n^4}, \tag{A.2}$$

where

$$\mathcal{D}(R) := \left\{(n,k,x) \in 2\mathbb{N}\times(2\mathbb{Z})^2 : |k| \vee |x| \le n^{3/5}\right\}.$$

Then there exists a constant $C_0 > 0$, an integer n_0, and a finite sum $R^{(2)}$ of monomials of effective order at most $-3/2$ such that for all $(n,k,x) \in \mathcal{D}(R)$ with $n > n_0$,

$$\left| R(n,k,x) - [1 - R^{(1)}(n,k,x) + R^{(2)}(n,k,x)] \right| \le C_0 n^{-3/2}. \tag{A.3}$$

Lemma A.3. *Suppose that $g \in \mathcal{B}_{exp}$ and that $b \ge 0$ is as in (3.1). Suppose also that $R^{(1)}$ is as in (A.2) and that Γ_{n_θ} is given by (5.10). Then there exists a constant $C > 0$ such that for all $x_0 \in \mathbb{R}$ and $n_\theta \in 2\mathbb{N}$,*

$$(i) \quad \left| \mathbb{E}\left[g(x_0 + X_{\tau_{n_\theta}}) - g(x_0 + X_{\tau_{J_n}}); \Gamma^{\complement}_{n_\theta}\right] \right| \le C n_\theta^{-3/2} e^{b|x_0| + b^2\sigma^2 T + b\sigma\sqrt{2T}},$$

$$(ii)\quad \Bigg|\mathbb{E}\big[g(x_0{+}X_{\tau_{n_\theta}}){-}g(x_0{+}X_{\tau_J});\Gamma_{n_\theta}\big] - \sum_{k=2-n_\theta}^{\infty}\sum_{x=-n_\theta}^{n_\theta} g(x_0{+}xh)P^J_{n_\theta}(k)P_{n_\theta}(x)R^{(1)}(n_\theta,k,x)\Bigg| \le Cn_\theta^{-3/2}e^{b|x_0|+b^2\sigma^2T}.$$

Proof. (i): Since $\Gamma^{\complement}_{n_\theta}\subset\big\{\big|X_{\tau_{n_\theta}}/h\big|>n_\theta^{3/5}\big\}\cup\big\{\,|J_n-n_\theta|>n_\theta^{3/5}\big\}$, we may use Hölder's inequality, (5.4), and (5.5) to show that there exists a constant $C'>0$ such that

$$\Big|\mathbb{E}\big[g(x_0{+}X_{\tau_{n_\theta}})-g(x_0{+}X_{\tau_{J_n}});\Gamma^{\complement}_{n_\theta}\big]\Big| \le C'n_\theta^{-3/2}\left(\mathbb{E}\Big|g(x_0{+}X_{\tau_{n_\theta}})-g(x_0{+}X_{\tau_{J_n}})\Big|^2\right)^{1/2}.$$

The claim follows, since by the triangle inequality, (5.1), (5.2), and the fact that $g\in\mathcal{B}_{\text{exp}}$, there exists another constant $\tilde{C}>0$ such that

$$\left(\mathbb{E}\Big|g(x_0{+}X_{\tau_{n_\theta}})-g(x_0{+}X_{\tau_{J_n}})\Big|^2\right)^{1/2} \le \tilde{C}e^{b|x_0|+b^2\sigma^2T+b\sigma\sqrt{2T}}.$$

(ii): The proof of item (ii) is done in several intermediate steps and only sketched here for the sake of brevity. More details can be found in [9, Lemma A.3].

Step 1: Let us first show that there exists a constant $C>0$ such that for all $x_0\in\mathbb{R}$ and n_θ,

$$\Bigg|\sum_{k=2-n_\theta}^{\infty}\sum_{x=-n_\theta}^{n_\theta} g(x_0{+}xh)P^J_{n_\theta}(k)P_{n_\theta}(x)R^{(2)}(n_\theta,k,x)\mathbb{1}_{\left\{|x|\vee|k|\le n_\theta^{3/5}\right\}}\Bigg| \le Cn_\theta^{-3/2}e^{b|x_0|+b^2\sigma^2T}, \tag{A.4}$$

where $R^{(2)}$ is as in Proposition A.2. Using the relations $h=\sigma\sqrt{\frac{T}{n}}$, $\theta_n=\frac{n_\theta T}{2}$ and (5.13), it can be shown that for given integers $p,q,r\in\mathbb{N}_0$ and subsets $\Lambda_1,\Lambda_2\subset\mathbb{Z}$,

$$\begin{aligned}&\sum_{k=2-n_\theta}^{\infty}\sum_{x=-n_\theta}^{n_\theta} g(x_0{+}xh)P^J_{n_\theta}(k)P_{n_\theta}(x)\frac{k^px^q}{n_\theta{}^r}\mathbb{1}_{\{x\in\Lambda_1,k\in\Lambda_2\}}\\ &= n_\theta^{(p+q)/2-r}\mathbb{E}\left[\left(\tfrac{X_{\tau_{n_\theta}}}{\sqrt{\sigma^2\theta_n}}\right)^q g(x_0{+}X_{\tau_{n_\theta}});X_{\tau_{n_\theta}}/h\in\Lambda_1\right]\mathbb{E}\left[\left(\tfrac{J_n-n_\theta}{\sqrt{n_\theta}}\right)^p;J_n{-}n_\theta\in\Lambda_2\right].\end{aligned} \tag{A.5}$$

By the definition of $R^{(2)}$, there exists an integer $N\in\mathbb{N}$, a vector $(a_i)_{i=1}^N\subset\mathbb{R}$, and vectors $(p_i)_{i=1}^N,(q_i)_{i=1}^N,(r_i)_{i=1}^N\in\mathbb{N}_0^N$ such that $\frac{p_i+q_i}{2}-r_i\le-3/2$ for all $1\le i\le N$, and

$$R^{(2)}(n_\theta,k,x)=\sum_{i=1}^N a_i\frac{k^{p_i}x^{q_i}}{n_\theta^{r_i}}\quad\text{for }\ (n_\theta,k,x)\in\mathcal{D}(R).$$

Therefore, by the relation (A.5), the left-hand side of (A.4) can be rewritten and estimated by

$$\begin{aligned}&\Bigg|\sum_{i=1}^N a_in_\theta^{\frac{p_i+q_i}{2}-r_i}\mathbb{E}\left[\left(\tfrac{X_{\tau_{n_\theta}}}{\sqrt{\sigma^2\theta_n}}\right)^{q_i}g(x_0{+}X_{\tau_{n_\theta}});\big|X_{\tau_{n_\theta}}/h\big|\le n_\theta^{3/5}\right]\mathbb{E}\left[\left(\tfrac{J-n_\theta}{\sqrt{n_\theta}}\right)^{p_i};|J_n{-}n_\theta|\le n_\theta^{3/5}\right]\Bigg|\\ &\le n_\theta^{-3/2}\sum_{i=1}^N|a_i|\,\mathbb{E}\left[\left(\tfrac{|X_{\tau_{n_\theta}}|}{\sqrt{\sigma^2\theta_n}}\right)^{q_i}\big|g(x_0{+}X_{\tau_{n_\theta}})\big|\right]\mathbb{E}\left(\tfrac{|J_n-n_\theta|}{\sqrt{n_\theta}}\right)^{p_i}\\ &\le \tilde{C}n_\theta^{-3/2}e^{b|x_0|+b^2\sigma^2T},\end{aligned}$$

where $\tilde{C}>0$ is some constant implied by (5.3) and (6.1). This proves (A.4).

Step 2: Let us show that for some constant $C>0$ and for all $x_0\in\mathbb{R}$ and $n_\theta\in 2\mathbb{N}$,

$$\Big|\sum_{k=2-n_\theta}^{\infty}\sum_{x=-n_\theta}^{n_\theta} g(x_0{+}xh)P^J_{n_\theta}(k)P_{n_\theta}(x)R^{(1)}(n_\theta,k,x)\mathbb{1}_{\{|x|\vee|k|>n_\theta^{3/5}\}}\Big|\le Cn_\theta^{-3/2}e^{b|x_0|+b^2\sigma^2T}. \tag{A.6}$$

By (A.2) and (A.5), it is sufficient to prove that for given $p,q,r\in\mathbb{N}_0$ there exists a constant $C_{p,q,r}>0$ such that for all $x_0\in\mathbb{R}$,

$$\Big|\sum_{k=2-n_\theta}^{\infty}\sum_{x=-n_\theta}^{n_\theta} g(x_0{+}xh)P^J_{n_\theta}(k)P_{n_\theta}(x)\frac{k^px^q}{n_\theta{}^r}\mathbb{1}_{\{|x|\vee|k|>n_\theta^{3/5}\}}\Big|\le C_{p,q,r}n_\theta^{-3/2}e^{b|x_0|+b^2\sigma^2T}. \tag{A.7}$$

Relation (A.7) can be verified by writing $\{|x|\vee|k|>n_\theta^{3/5}\}=\{|x|>n_\theta^{3/5}\}\cup\{|k|>n_\theta^{3/5},|x|\le n_\theta^{3/5}\}$ and considering the corresponding sums separately using relation (A.5) and similar calculations as in Step 1. Indeed, the case $\{|x|>n_\theta^{3/5}\}$ can be shown using Hölder's inequality and relations (5.1), (5.4), (5.6), and (6.1). The case $\{|k|>n_\theta^{3/5},|x|\le n_\theta^{3/5}\}$ follows from Hölder's inequality, (5.3), (5.5), and (6.1).

Step 3: Since the processes $(\Delta\tau_k)_{k=1,2,\dots}$ and $(\Delta X_{\tau_k})_{k=1,2,\dots}$ are independent (see Subsection 2.1), the random variable J_n and the process $(X_{\tau_k})_{k=0,1,\dots}$ are also independent. Taking also into account that $\mathrm{supp}P_{n_\theta+k}=\{m\in 2\mathbb{Z}:|m|\le n_\theta+k\}$ (for each $k\in 2\mathbb{N}$) and $\mathrm{supp}P^J_{n_\theta}=\{m{-}n_\theta:m\in 2\mathbb{N}\}$, it can be shown that

$$\begin{aligned}&\mathbb{E}\big[g(x_0{+}X_{\tau_{n_\theta}})-g(x_0{+}X_{\tau_{J_n}});\Gamma_{n_\theta}\big]\\&=\sum_{k=2-n_\theta}^{\infty}\sum_{x=-n_\theta}^{n_\theta} g(x_0{+}xh)P^J_{n_\theta}(k)P_{n_\theta}(x)\Big(1-\frac{P_{n_\theta+k}(x)}{P_{n_\theta}(x)}\Big)\mathbb{1}_{\{|x|\vee|k|\le n_\theta^{3/5}\}}.\end{aligned}$$

Thus, by (A.1)–(A.5), there exist constants $C_0,C_1>0$ and $n_0\in 2\mathbb{N}$ such that whenever $n_\theta>n_0$,

$$\begin{aligned}&\Big|\sum_{k=2-n_\theta}^{\infty}\sum_{x=-n_\theta}^{n_\theta} g(x_0{+}xh)P^J_{n_\theta}(k)P_{n_\theta}(x)R^{(1)}(n_\theta,k,x)\mathbb{1}_{\{|x|\vee|k|\le n_\theta^{3/5}\}}\\&\qquad-\mathbb{E}\big[g(x_0{+}X_{\tau_{n_\theta}})-g(x_0{+}X_{\tau_J});\Gamma_{n_\theta}\big]\Big|\\&=\Big|\sum_{k=2-n_\theta}^{\infty}\sum_{x=-n_\theta}^{n_\theta} g(x_0{+}xh)P^J_{n_\theta}(k)P_{n_\theta}(x)\big(R^{(1)}(n_\theta,k,x)-[1-R(n_\theta,k,x)]\big)\mathbb{1}_{\{|x|\vee|k|\le n_\theta^{3/5}\}}\Big|\\&\le C_0n_\theta^{-3/2}\sum_{k=2-n_\theta}^{\infty}\sum_{x=-n_\theta}^{n_\theta}|g(x_0{+}xh)|\,P^J_{n_\theta}(k)P_{n_\theta}(x)\mathbb{1}_{\{|x|\vee|k|\le n_\theta^{3/5}\}}\\&\qquad+\Big|\sum_{k=2-n_\theta}^{\infty}\sum_{x=-n_\theta}^{n_\theta} g(x_0{+}xh)P^J_{n_\theta}(k)P_{n_\theta}(x)R^{(2)}(n_\theta,k,x)\mathbb{1}_{\{|x|\vee|k|\le n_\theta^{3/5}\}}\Big|\\&\le C_0n_\theta^{-3/2}\mathbb{E}\Big[\big|g(x_0{+}X_{\tau_{n_\theta}})\big|;\big|X_{\tau_{n_\theta}}/h\big|\le n_\theta^{3/5}\Big]+C_1n_\theta^{-3/2}e^{b|x_0|+b^2\sigma^2T}\\&\le C_2n_\theta^{-3/2}e^{b|x_0|+b^2\sigma^2T}\end{aligned} \tag{A.8}$$

for some constant $C_2>0$ implied by (5.3). Consequently, we get the claim for all $n_\theta>n_0$ by the triangle inequality, (A.6), and (A.8). By letting

$$M:=\sup_{(n,k,x):n\le n_0}\Big|\big(R^{(1)}(n,k,x)-[1-R(n,k,x)]\big)\mathbb{1}_{\{|x|\vee|k|\le n_\theta^{3/5}\}}\Big|<\infty,$$

for $n_\theta \le n_0$ we find another constant $C_4 = C_4(n_0) > 0$ such that

$$\begin{aligned}
&\Big|\sum_{k=2-n_\theta}^{\infty}\sum_{x=-n_\theta}^{n_\theta} g(x_0+xh)P^J_{n_\theta}(k)P_{n_\theta}(x)\big(R^{(1)}(n_\theta,k,x)-[1-R(n_\theta,k,x)]\big)\mathbb{1}_{\{|x|\vee|k|\le n_\theta^{3/5}\}}\Big| \\
&\le M\sum_{k=2-n_\theta}^{\infty}\sum_{x=-n_\theta}^{n_\theta} |g(x_0+xh)|\,P^J_{n_\theta}(k)P_{n_\theta}(x)\mathbb{1}_{\{|x|\vee|k|\le n_\theta^{3/5}\}} \\
&\le M\mathbb{E}\left[\big|g(x_0+X_{\tau_{n_\theta}})\big|;\big|X_{\tau_{n_\theta}}/h\big|\le n_\theta^{3/5}\right]\mathbb{P}\left(|J_{n_\theta}-n_\theta|\le n_\theta^{3/5}\right) \\
&\le C_4 n_\theta^{-3/2}e^{b|x_0|+b^2\sigma^2T}
\end{aligned} \tag{A.9}$$

by (5.3). Combine (A.6), (A.8), and (A.9) to complete the proof. □

A.3 Other auxiliary results

Lemmata A.4 and A.6 below are essential for the proof of Proposition 6.3.

Lemma A.4. *Under Assumption 4.11, suppose that $n_\theta \in 2\mathbb{N}$ and a constant $\xi > 0$ are such that $n_\theta\xi \in \mathbb{N}$. Then for every $\rho \in (0, \frac{\pi^2}{12}\xi\theta_n\sqrt{n_\theta})$ it holds that*

$$(i)\quad \mathbb{P}\left(\sqrt{n_\theta}(\tau_{n_\theta\xi}-\xi\theta_n)>\rho\right)\le \exp\left(-\frac{3}{2}\frac{\rho^2}{\xi\theta_n^2}H\left(\sqrt{\frac{3\rho}{\xi\theta_n\sqrt{n_\theta}}}\right)\right), \tag{A.10}$$

$$(ii)\quad \mathbb{P}\left(\sqrt{n_\theta}(\tau_{n_\theta\xi}-\xi\theta_n)<-\rho\right)\le \exp\left(-\frac{3}{2}\frac{\rho^2}{\xi\theta_n^2}H\left(\sqrt{\frac{3\rho}{\xi\theta_n\sqrt{n_\theta}}}\right)\right), \tag{A.11}$$

where the function $H : (0,\pi/2) \to \mathbb{R}$ is given by

$$H(x) := 1 + \tfrac{6}{x^4}\left(\tfrac{x^2}{2} + \log\cos x\right). \tag{A.12}$$

Remark A.5. The above estimates are non-trivial only whenever H is positive. Since $H(0+) = 1/2$, it holds that $H(x) > 0$ for small enough x. Notice that the condition $\rho \in (0, \frac{\pi^2}{12}\xi\theta_n\sqrt{n_\theta})$ ensures that $\sqrt{\frac{3\rho}{\xi\theta_n\sqrt{n_\theta}}} \in (0, \pi/2)$, which is the domain of H.

The proof of Lemma A.4 is given in [9, Subsection 6.2]. The following result, which is applied in the proof of Proposition 6.3, resembles inequality (42) in [13]. However, the time-dependent setting causes some changes.

Lemma A.6. *Under Assumption 4.11, suppose that $n_\theta \in 2\mathbb{N}$, $\delta \in (0, \frac{\pi^2}{12+\pi^2})$, and let H be as in (A.12). Then*

$$(i)\quad \mathbb{P}(J_n > n_\theta(1+\delta)) \le \exp\Big(-\tfrac{3}{2}\tfrac{n_\theta\delta^2}{1+\delta}H\Big(\sqrt{\tfrac{3\delta}{1+\delta}}\Big)\Big) \quad \textit{if} \quad n_\theta(1+\delta)\in\mathbb{N},$$

$$(ii)\quad \mathbb{P}(J_n < n_\theta(1-\delta)) \le \exp\Big(-\tfrac{3}{2}\tfrac{n_\theta\delta^2}{1-\delta}H\Big(\sqrt{\tfrac{3\delta}{1-\delta}}\Big)\Big) \quad \textit{if} \quad n_\theta(1-\delta)\in\mathbb{N}.$$

Proof. Fix $n_\theta \in 2\mathbb{N}$, $\delta \in (0, \frac{\pi^2}{12+\pi^2})$, and let $\rho := \delta\theta_n\sqrt{n_\theta}$. For (i), let $\xi := 1+\delta$ and suppose that $n_\theta(1+\delta) = n_\theta\xi \in \mathbb{N}$. Then

$$\mathbb{P}(J_n > n_\theta\xi) = \mathbb{P}(\tau_{n_\theta\xi} < \theta_n) = \mathbb{P}(\sqrt{n_\theta}(\tau_{n_\theta\xi}-\xi\theta_n) < -\rho) \le \exp\left(-\frac{3}{2}\frac{\rho^2}{\xi\theta_n^2}H\left(\sqrt{\frac{3\rho}{\xi\theta_n\sqrt{n_\theta}}}\right)\right)$$

by (A.11), since the choice of δ ensures that the pair (ξ,ρ) satisfies the assumptions of Lemma A.4. To show (ii), let now $\xi := 1-\delta$ and suppose that $n_\theta(1-\delta) = n_\theta\xi \in \mathbb{N}$. Then by (A.10),

$$\mathbb{P}(J_n < n_\theta\xi) = \mathbb{P}(\tau_{n_\theta\xi} > \theta_n) = \mathbb{P}(\sqrt{n_\theta}(\tau_{n_\theta\xi}-\xi\theta_n) > \rho) \le \exp\left(-\frac{3}{2}\frac{\rho^2}{\xi\theta_n^2}H\left(\sqrt{\frac{3\rho}{\xi\theta_n\sqrt{n_\theta}}}\right)\right)$$

since the pair (ξ, ρ) satisfies the assumptions of Lemma A.4 due to the choice of δ. □

The following identities are applied in the proofs of Propositions 4.9 and 4.10.

Lemma A.7. *Let $h > 0$ and recall the operators Π_e and Π_o given by Definition 4.1.*

(i) *For all $\xi \in \mathbb{R}$, it holds that*

$$\Pi_e \mathbb{1}_{\{\xi\}}(x) = \frac{\mathbb{1}_{\{\xi \in \mathbb{Z}_e^h\}}}{4h}\big(|x-(\xi-2h)| + |x-(\xi+2h)| - 2\,|x-\xi|\big), \quad x \in \mathbb{R}. \tag{A.13}$$

(ii) *If $y \in [2kh, (2k+2)h)$ for $k \in \mathbb{Z}$, then in terms of d_o defined in (4.2),*

$$\big|\Pi_o \Pi_e \mathbb{1}_{(y,\infty)}(x) - \Pi_e \mathbb{1}_{(y,\infty)}(x)\big| = \frac{d_o(x)}{4h}\mathbb{1}_{[(2k-1)h,(2k+3)h)}(x), \quad x \in \mathbb{R}. \tag{A.14}$$

Proof. (i): It is obvious by the definition of Π_e that $\Pi_e \mathbb{1}_{\{\xi\}} \equiv 0$ for $\xi \notin \mathbb{Z}_e^h$. If $\xi \in \mathbb{Z}_e^h$, then

$$\Pi_e \mathbb{1}_{\{\xi\}}(x) = \begin{cases} \frac{x-(\xi-2h)}{2h}, & (\xi-2h) \le x < \xi, \\ \frac{(\xi+2h)-x}{2h}, & \xi \le x < (\xi+2h), \end{cases} \tag{A.15}$$

and zero elsewhere, so it suffices to verify that (A.15) agrees with the representation given in (A.13).

(ii): Suppose that $y \in [2kh, (2k+2)h)$ for some $k \in \mathbb{Z}$. One checks that

$$\Pi_e \mathbb{1}_{(y,\infty)}(x) = \frac{1}{2} + \frac{1}{4h}|x-2kh| - \frac{1}{4h}|x-(2k+2)h|\,, x \in \mathbb{R}. \tag{A.16}$$

Then, by the linearity of Π_o and by (A.16), we have for every $x \in \mathbb{R}$ that

$$\begin{aligned} &\Pi_o \Pi_e \mathbb{1}_{(y,\infty)}(x) - \Pi_e \mathbb{1}_{(y,\infty)}(x) \\ &= \frac{1}{4h}\big(\Pi_o\,|\cdot - 2kh|\,(x) - |x-2kh|\big) - \frac{1}{4h}\big(\Pi_o\,|\cdot-(2k+2)h|\,(x) - |x-(2k+2)h|\big) \\ &= \frac{d_o(x)}{4h}\left(\mathbb{1}_{[(2k-1)h,(2k+1)h)}(x) - \mathbb{1}_{[(2k+1)h,(2k+3)h)}(x)\right), \end{aligned} \tag{A.17}$$

since it holds for all $x \in \mathbb{R}$ and $m \in \mathbb{Z}$ that

$$\Pi_o\,|\cdot - 2mh|\,(x) - |x-2mh| = d_o(x)\mathbb{1}_{[(2m-1)h,(2m+1)h)}(x). \tag{A.18}$$

Taking the absolute values of both sides of (A.17) then completes the proof. □

Acknowledgement The author was financially supported by the Magnus Ehrnrooth foundation during the preparation of this manuscript.

References

[1] R. Avikainen, *On generalized bounded variation and approximation of SDEs*, preprint (2009). http://www.math.jyu.fi/research/pspdf/383.pdf.

[2] A. DasGupta, *Probability for Statistics and Machine Learning*, Springer 2001.

[3] H. Dong, N. V. Krylov, *Rate of convergence of finite-difference approximations for degenerate linear parabolic equations with C^1 and C^2 coefficients*, Electron. J. Differ. Equ., Vol. 2005 (2005), No. 102, pp. 1–25.

[4] C. Geiss, A. Luoto, P. Salminen, *On first exit times and their means for Brownian bridges*, preprint. https://arxiv.org/pdf/1711.06107.pdf.

[5] S. Heston, G. Zhou, *On the rate of convergence of discrete-time contingent claims*, Math. Finance, Vol. 10 (2000), No. 1, pp. 53–75.

[6] M. L. Juncosa, D. M. Young, *On the order of convergence of solutions of a difference equation to a solution of the diffusion equation*, SIAM J. Appl. Math., Vol. 1 (1953), No. 2, pp. 111–135.

[7] D. Lamberton, L. C. G. Rogers, *Optimal stopping and embedding*, J. Appl. Probab., Vol. 37 (2000), No. 4, pp. 1143–1148.

[8] L. Lindhagen, *Finite difference equations and convergence rates in the central limit theorem*, Probab. Math. Statist., Vol. 27 (2007), No. 2, pp. 153–166.

[9] A. Luoto, *Time-dependent weak rate of convergence for functions of generalized bounded variation*, preprint. https://arxiv.org/abs/1609.05768.

[10] D. Revuz, M. Yor, *Continuous Martingales and Brownian Motion*, 3nd Edition, Springer 1999.

[11] A. C. Reynolds, *Convergent finite difference schemes for nonlinear parabolic equations*, SIAM J. Numer. Anal., Vol. 9 (1972), No. 4, pp. 523–533.

[12] L. C. G. Rogers, E. J. Stapleton, *Fast accurate binomial pricing*, Finance Stochast., Vol. 2 (1997), No. 1, pp. 3–17.

[13] J. B. Walsh, *The rate of convergence of the binomial tree scheme*, Finance Stochast., Vol. 7 (2003), No. 3, pp. 337–361.

[14] S. Zhu, G. Yuan, W. Sun, *Convergence and stability of explicit/implicit schemes for parabolic equations with discontinuous coefficients*, Int. J. Numer. Anal. Model., Vol. 1 (2004), No. 2, pp. 131–145.

[B]

ON FIRST EXIT TIMES AND THEIR MEANS FOR BROWNIAN BRIDGES

C. GEISS, A. LUOTO, AND P. SALMINEN

On first exit times and their means for Brownian bridges

Christel Geiss[1] Antti Luoto[2] Paavo Salminen[3]

Abstract

For a Brownian bridge from 0 to y we prove that the mean of the first exit time from interval $(-h, h),\ h > 0$, behaves as $O(h^2)$ when $h \downarrow 0$. Similar behavior is seen to hold also for the 3-dimensional Bessel bridge. For Brownian bridge and 3-dimensional Bessel bridge this mean of the first exit time has a puzzling representation in terms of the Kolmogorov distribution. The result regarding the Brownian bridge is applied to prove in detail an estimate needed by Walsh to determine the convergence of the binomial tree scheme for European options.

Keywords: Brownian motion, Brownian bridge, Bessel process, Bessel bridge, first exit time, last exit time, Kolmogorov distribution function, binomial tree scheme
AMS 60J65, 60J60; 91G60

1 Introduction

In this paper, we consider diffusion bridges and present an integral representation of the mean of their first exit time from an interval. With the help of this representation we deduce the limiting behavior of the mean, when the length of the interval around the starting value of the bridge decreases to 0. For a continuous process $(X_t)_{t\geq 0}$ with $X_0 = x$ we denote the first exit time from (a, b) by

$$\mathcal{T}^X_{(a,b)} := \inf\{t > 0 : X_t \notin (a, b)\}, \quad a < x < b,$$

with the convention $\mathcal{T}^X_{(a,b)} := \infty$ if $\{t > 0 : X_t \notin (a, b)\} = \emptyset$. We simply write $\mathcal{T}_{(a,b)}$ if there is no ambiguity about the underlying process. The main part of the paper concerns

[1]Department of Mathematics and Statistics, University of Jyväskylä, Finland.
christel.geiss@jyu.fi
[2]Department of Mathematics and Statistics, University of Jyväskylä, Finland.
antti.k.luoto@student.jyu.fi
[3]Faculty of Science and Engineering, Åbo Akademi University, Finland.
paavo.salminen@abo.fi

Brownian bridge and Bessel bridge (with dimension parameter 3). For the Brownian bridge $(B_t^{x,T,y})_{0\le t<T}$ from x to y of length T (a similar result holds for Bessel bridges), it is shown in Theorem 3.4 that

$$\lim_{h\downarrow 0}\frac{\mathbb{E}_{0,T,y}[\mathcal{T}_{(-h,h)}]}{h^2}=1,\quad y\neq 0, \tag{1}$$

where $\mathbb{E}_{x,T,y}$ stands for the corresponding expectation of the Brownian bridge.

Let $(W_t^x)_{t\ge 0}$ be the Brownian motion started at $x\in\mathbb{R}$. Recall that (this can be deduced, e.g., from the Laplace transform of $\mathcal{T}_{(a,b)}$ given in [2, Part II, Section 1, 3.0.1]) it holds

$$\mathbb{E}_x\left[\mathcal{T}_{(a,b)}\right]=(b-x)(x-a),\quad a<x<b, \tag{2}$$

where again $\mathbb{E}_x$ denotes the corresponding expectation. Consequently, by (1),

$$\lim_{h\downarrow 0}\frac{\mathbb{E}_{0,T,y}[\mathcal{T}_{(-h,h)}]}{\mathbb{E}_0[\mathcal{T}_{(-h,h)}]}=1,\quad y\neq 0.$$

For some other diffusion bridges a similar asymptotic behavior can be found. But we have not been able to prove such a result in generality.

The paper is organized as follows. In Section 2 we describe our setting and the construction of bridge measures for regular diffusions by Doob's h-transform. We also recall some properties of the Brownian bridge. Section 3 contains our main results. We start by deriving integral representations for the mean of $\mathcal{T}^X_{(a,b)}$ when X is, firstly, a general regular diffusion and, secondly, a corresponding diffusion bridge. We also calculate for a regular diffusion $X,\ X_0=x$, the limiting behavior of the mean of $\mathcal{T}^X_{(x-h,x+h)}$ as $h\to 0$. In Subsection 3.3 we focus on Brownian bridge and 3-dimensional Bessel bridge starting from $x>0$ and find the limiting behavior of the mean of $\mathcal{T}_{(x-h,x+h)}$ as $h\to 0$. In Theorem 3.5 we state a weak convergence result for $\mathcal{T}^X_{(x-h,x+h)}/h^2$. For the means of the first exit times – considered in Subsection 3.3 – Subsection 3.4 provides puzzling representations in terms of the Kolmogorov distribution function. We discuss also other properties of the Kolmogorov distribution, in particular, its connection with the last exit time distribution of Brownian motion. As an application, in Section 4, we use our results concerning Brownian bridge to give a detailed proof of an estimate in Walsh [15] needed therein when deriving the convergence rate of an option price calculated from the binomial tree scheme to the Black-Scholes price. The estimate is used also in a forthcoming paper [11].

2 Preliminaries

We start with the description of our setting. Let $C[0,\infty)$ denote the space of continuous functions $\omega:[0,\infty)\mapsto\mathbb{R}$, and

$$\mathcal{C}_t:=\sigma\{\omega(s):s\le t\}$$

the smallest σ-algebra making the coordinate mappings up to time t measurable. Furthermore, let $\mathcal{C}$ be the smallest σ-algebra containing all $\mathcal{C}_t$, $t \geq 0$. For an interval $I \subset \mathbb{R}$ let $(\mathbb{P}_x^X)_{x \in I}$ be a family of probability measures defined in the filtered canonical space $(C[0,\infty), \mathcal{C}, (\mathcal{C}_t)_{t\geq 0})$ such that under $\mathbb{P}_x^X$ for a given $x \in I$ the coordinate process $X = (X_t)_{t\geq 0} := (\omega_t)_{t \geq 0}$ is a regular diffusion taking values in I and starting from x. Here, X is considered in the sense of Itô and McKean [7], see also [2]. A crucial property of X is that there exists a (speed) measure m^X such that the transition probability has a continuous strictly positive density $(t, x, y) \mapsto q_t(x, y)$, $t > 0, x, y \in I$ with respect to m^X i.e.,

$$\mathbb{P}_x^X(X_t \in dy) = q_t(x, y) m^X(dy), \tag{3}$$

see [7, page 149 and 157].

For $(X_t)_{t\geq 0}$, $X_0 = x$, as defined above and $T > 0$, one can construct a new non-homogeneous strong Markov process by conditioning X to be at a given point $y \in I$ at time T. Although the conditioning is, in general, with respect to a zero set $\{X_T = y\}$, it can be realized using the Bayes formula and the notion of regular conditional distributions. Another approach is to apply the theory of the Doob h-transforms. To explain this briefly, consider X in space-time i.e., the process $\bar{X} = ((X_t, t))_{t\geq 0}$. Introduce for $z \in I$ and $t < T$ the function

$$h(z, t) := h(z, t; y, T) := q_{T-t}(z, y).$$

By the Chapman-Kolmogorov equation it holds for $x \in I$ and $s < t$

$$\begin{aligned} \mathbb{E}_{(x,s)}^X[h(X_t, t)] &= \int_I q_{t-s}(x, z) h(z, t)\, m(dz) \\ &= \int_I q_{t-s}(x, z) q_{T-t}(z, y)\, m(dz) \\ &= q_{T-s}(x, y) \\ &= h(x, s), \end{aligned}$$

where $\mathbb{E}_{(x,s)}^X$ refers to the expectation associated with the space-time process $\bar{X}$ initiated from x at time s. Consequently, we may define for $f \in \mathcal{B}_b(I)$(= bounded measurable functions on I) and $s < t < T$ a non-homogeneous Markov semigroup $P_{s,t}^{X,h}$, $0 < s < t < T$, via

$$P_{s,t}^{X,h} f(x) := \mathbb{E}_x^X\Big[f(X_{t-s}) \frac{h(X_{t-s}, t)}{h(x, s)}\Big]. \tag{4}$$

The process governed by the probability measure induced by this semigroup is called the X-bridge to y of length T. The notations $X^{x,T,y}$, $\mathbb{P}_{x,T,y}^X$, and $\mathbb{E}_{x,T,y}^X$ are used for this process, its probability measure and the expectation, respectively, when the initial state is $x \in I$. From (4) one may deduce the following absolute continuity relation for $A_t \in \mathcal{C}_t$, $t < T$,

$$\mathbb{P}_{x,T,y}^X(A_t) = \mathbb{E}_x^X\Big[\frac{h(X_t, t)}{h(x, 0)}\,;\, A_t\Big] = \mathbb{E}_x^X\Big[\frac{q_{T-t}(X_t, y)}{q_T(x, y)}\,;\, A_t\Big]. \tag{5}$$

We refer to Chung and Walsh [3] for a general discussion on h-transforms, to Fitzsimmons et al. [5], in particular, Proposition 1, for general Markovian bridges, and to Salminen [13] for some properties of diffusion bridges.

Since our main interest in this paper is focused on the case where the underlying process is a Brownian motion, we make use of the notations $(W_t)_{t\geq 0}$ and $(B_t^{x,T,y})_{0\leq t<T}$ for standard Brownian motion and Brownian bridge from x to y of length T, denoting the corresponding probability measures by $\mathbb{P}_x$ and $\mathbb{P}_{x,T,y}$, and the expectations by $\mathbb{E}_x$ and $\mathbb{E}_{x,T,y}$, respectively. Brownian bridge has in addition to the general h-transform approach a few equivalent specific characterizations. Indeed, Brownian bridge can be viewed as (i) a Gaussian process, (ii) a deterministic time change with an additional drift of standard Brownian motion or (iii) as a solution of a SDE, see e.g. [2] p. 66. In Section 4 we apply, in particular, (ii) and (iii). The second one states that

$$B_t^{x,T,y} \overset{d}{=} \begin{cases} (1-\frac{t}{T})W_{\frac{Tt}{T-t}} + x + \frac{(y-x)t}{T}, & t<T \\ y & t=T, \end{cases} \tag{6}$$

where $\overset{d}{=}$ means that the processes on the left and the right hand side are identical in law. The third one says that

$$B_t^{x,T,y} \overset{d}{=} \begin{cases} (T-t)\int_0^t \frac{dW_s}{T-s} + x + \frac{(y-x)t}{T}, & t<T \\ y & t=T. \end{cases} \tag{7}$$

Here it is assumed that the canonical filtration $(\mathcal{C}_t)_{t\geq 0}$ is augmented with the null sets of $\mathcal{C}$ with respect to $\mathbb{P}_0$ in order to have the usual conditions satisfied.

3 Main results

3.1 Integral representations for the mean of the first exit time

Let $X=(X_t)_{t\geq 0}$ be a regular diffusion taking values on an interval I, and recall from (3) the notation $q_t(x,y)$ for its transition density with respect to the speed measure m^X. For $a<b$, $a,b\in I$, let $\widehat{X}$ denote X killed at $\mathcal{T}_{(a,b)}$. Then $\widehat{X}$ is a regular diffusion on (a,b). The speed measure of $\widehat{X}$ is m^X, and $\widehat{X}$ has a continuous strictly positive transition density $\widehat{q}_t(x,y)$ such that for $x,y\in(a,b)$

$$\begin{aligned} \mathbb{P}_x^{\widehat{X}}(\widehat{X}_t\in dy) &= \widehat{q}_t(x,y)m^X(dy) \\ &= \mathbb{P}_x^X(\mathcal{T}_{(a,b)}>t, X_t\in dy) \\ &= \mathbb{P}_x^X(a<\inf_{0\leq s\leq t} X_s, \sup_{0\leq s\leq t} X_s<b, X_t\in dy). \end{aligned} \tag{8}$$

This yields immediately the following result.

Proposition 3.1. *In case* $\mathbb{P}_x^X(\mathcal{T}_{(a,b)} < \infty) = 1$, *i.e.,* X *does not die inside* (a,b), *it holds*

$$\mathbb{E}_x^X[\mathcal{T}_{(a,b)}] = \int_0^\infty dt \int_a^b m^X(dz)\,\widehat{q}_t(x,z). \tag{9}$$

The next proposition will serve as an important tool for the calculations below.

Proposition 3.2. *For* $x \in (a,b) \subset I$ *and* $y \in I \backslash (a,b)$ *it holds*

$$\mathbb{E}_{x,T,y}^X[\mathcal{T}_{(a,b)}] = \int_0^T dt \int_a^b m^X(dz)\,\widehat{q}_t(x,z)\frac{q_{T-t}(z,y)}{q_T(x,y)}. \tag{10}$$

Proof. Since $x \in (a,b)$ and $y \in I\backslash(a,b)$ we have $\mathbb{P}_{x,T,y}^X(\mathcal{T}_{(a,b)} < T) = 1$. Consider

$$\begin{aligned}
\mathbb{E}_{x,T,y}^X[\mathcal{T}_{(a,b)}] &= \int_0^T dt \int_a^b \mathbb{P}_{x,T,y}^X(\mathcal{T}_{(a,b)} > t,\ X_t^{x,T,y} \in dz) \\
&= \int_0^T dt \int_a^b \mathbb{P}_x^X(\mathcal{T}_{(a,b)} > t,\ X_t \in dz)\frac{q_{T-t}(z,y)}{q_T(x,y)},
\end{aligned} \tag{11}$$

where the Markov property and formula (4) is used. It remains to apply (8).

□

3.2 The limit behaviour for regular diffusions

One of our main issues concerns the limiting behavior of the mean of $\mathcal{T}^X_{(x-h,x+h)}$ as $h \to 0$ for diffusion bridges. For regular diffusions we have the following fairly complete characterization.

Proposition 3.3. *Assume that the differential operator associated with the regular diffusion* X *is given by*

$$\mathcal{G}u(x) := \frac{1}{2}a^2(x)u''(x) + b(x)u'(x), \qquad x \in I,$$

where $x \mapsto a^2(x) > 0$ *and* $x \mapsto b(x)$ *are continuous in* I. *Let* $x_o \in Int(I)$. *Then*

$$\lim_{h\downarrow 0} \frac{\mathbb{E}_{x_o}^X[\mathcal{T}_{(x_o-h,x_o+h)}]}{h^2} = a^{-2}(x_o).$$

Proof. Recall from [2, Part I, Chapter II, No 7, p.17] that the speed measure m^X and the scale function s^X can be taken to be

$$m^X(dx) = m^X(x)dx, \quad \frac{d}{dx}s^X(x) = e^{-B(x)}, \tag{12}$$

where

$$m^X(x) = 2a^{-2}(x)e^{B(x)}, \quad B(x) = \int^x 2a^{-2}(y)b(y)dy. \tag{13}$$

Consider now the process X initiated at x_o and killed when it leaves the interval $(x_o - h, x_o + h)$. We let $(\widehat{X}_t)_{t\geq 0}$ denote this diffusion:

$$\widehat{X}_t = \begin{cases} X_t, & t < \mathcal{T}_{(x_o-h,x_o+h)}, \\ \partial, & t \geq \mathcal{T}_{(x_o-h,x_o+h)}, \end{cases}$$

where ∂ is a cemetary point. It is well known (cf. [2, Part I, Chapter II, No 11]) that the 0-resolvent kernel of $\widehat{X}$ is given by

$$\widehat{G}_0(x,y) := \int_0^\infty \widehat{q}_t(x,y)dt = \begin{cases} \dfrac{(s(x) - s(x_o - h))(s(x_o + h) - s(y))}{s(x_o + h) - s(x_o - h)}, & x < y, \\[2ex] \dfrac{(s(y) - s(x_o - h))(s(x_o + h) - s(x))}{s(x_o + h) - s(x_o - h)}, & x > y, \end{cases}$$

where $s := s^X$ and $\widehat{q}_t(x,y)$ is the transition density w.r.t. the speed measure m^X. Consequently, relation (9) implies that

$$\begin{aligned} \mathbb{E}^X_{x_o}[\mathcal{T}_{(x_o-h,x_o+h)}] &= \int_0^\infty \left(\int_{x_o-h}^{x_o+h} \widehat{q}_t(x_o,y) m^X(dy) \right) dt \\ &= \int_{x_o-h}^{x_o+h} \widehat{G}_0(x_o,y) m^X(dy) \\ &= \frac{s(x_o+h) - s(x_o)}{s(x_o+h) - s(x_o-h)} \int_{x_o-h}^{x_o} \big(s(y) - s(x_o-h)\big) m^X(y)dy \\ &\quad + \frac{s(x_o) - s(x_o-h)}{s(x_o+h) - s(x_o-h)} \int_{x_o}^{x_o+h} \big(s(x_o+h) - s(y)\big) m^X(y)dy. \end{aligned} \tag{14}$$

Since s is assumed to be continuously differentiable we have

$$\lim_{h\downarrow 0} \frac{s(x_o+h) - s(x_o)}{s(x_o+h) - s(x_o-h)} = \lim_{h\downarrow 0} \frac{s(x_o) - s(x_o-h)}{s(x_o+h) - s(x_o-h)} = \frac{1}{2}, \qquad s'(x_o) \neq 0,$$

and l'Hospital's rule yields

$$\begin{aligned} &\lim_{h\downarrow 0} \frac{1}{h^2} \int_{x_o-h}^{x_o} (s(y) - s(x_o - h)) m^X(y)dy \\ &= \lim_{h\downarrow 0} \frac{1}{2h} \int_{x_o-h}^{x_o} \left(-\frac{d}{dh} s(x_o - h) \right) m^X(y)dy \\ &= \lim_{h\downarrow 0} \left(-\frac{d}{dh} s(x_o - h) \right) \frac{1}{2h} \int_{x_o-h}^{x_o} m^X(y)dy \\ &= \frac{1}{2} s'(x_o) m^X(x_o) \\ &= a^{-2}(x_o), \end{aligned}$$

where the last equality follows from relations (12) and (13). Similarly,

$$\lim_{h\downarrow 0}\frac{1}{h^2}\int_{x_o}^{x_o+h}\big(s(x_o+h)-s(y)\big)m^X(y)dy = a^{-2}(x_o).$$

Hence, by (14),

$$\lim_{h\downarrow 0}\frac{\mathbb{E}^X_{x_o}[\mathcal{T}_{(x_o-h,x_o+h)}]}{h^2} = a^{-2}(x_o).$$

□

3.3 The mean of the first exit time for Brownian bridge and 3-dimensional Bessel bridge

Let $p_t(x,y)$ the transition density of the standard Brownian motion,

$$p_t(x,y) = \frac{1}{\sqrt{2\pi t}}e^{-\frac{(y-x)^2}{2t}},$$

and $r_t^{(3)}(x,y)$ the transition density of the 3-dimensional Bessel process,

$$r_t^{(3)}(x,y) = \frac{y}{x}(p_t(x,y)-p_t(x,-y)), \quad x,y>0. \tag{15}$$

We introduce also the following function

$$\begin{aligned}\Delta(z,h,t) &:= \frac{1}{\sqrt{2\pi t}}\sum_{m=-\infty}^{\infty}\left(e^{-\frac{(z+4mh)^2}{2t}} - e^{-\frac{(z+(2m+1)2h)^2}{2t}}\right)\\ &= \frac{1}{\sqrt{2\pi t}}\sum_{m=-\infty}^{\infty}(-1)^m e^{\frac{(z-2mh)^2}{2t}}\\ &= p_t(0,z)\sum_{m=-\infty}^{\infty}(-1)^m e^{\frac{-2mh(mh-z)}{t}}, \quad t>0.\end{aligned} \tag{16}$$

Then, by [2, Part II, Section 1, 1.15.8 (p. 180)] and (16),

$$\begin{aligned}\mathbb{P}_0(\inf_{0\le s\le t}W_s > -h, \sup_{0\le s\le t}W_s < h, W_t\in dz) &= \mathbb{P}_0(\mathcal{T}_{(-h,h)} > t, W_t\in dz)\\ &= \Delta(z,h,t)dz.\end{aligned} \tag{17}$$

In addition, according to [2, Part II, Section 5, 1.15.8] we have for $z\in(x-h,x+h)$ that

$$\mathbb{P}^{(3)}_x(\sup_{0\le s\le t}|X_s-x|<h, X_t\in dz)$$

$$= \frac{z}{x\sqrt{2\pi t}} \sum_{m=-\infty}^{\infty} \left(e^{-\frac{((z-x)+2h(2m))^2}{2t}} - e^{-\frac{((z-x)+2h(2m+1))^2}{2t}} \right) dz$$

$$= \frac{z}{x} \Delta(z - x, h, t) dz, \quad x > h. \tag{18}$$

Finally, let us recall that by [2, Part II, Section 1, 3.0.6 (a)]

$$\mathbb{P}_x(\mathcal{T}_a \in dr, \mathcal{T}_a < \mathcal{T}_b) = \mathrm{ss}_r(b - x, b - a)dr, \quad a < x < b, \tag{19}$$

where the special function ss_t is given by (see [2, Appendix 2, No 13])

$$\mathrm{ss}_t(a, b) := \mathcal{L}^{-1}_{t,\gamma}\Big(\frac{\sinh(a\sqrt{2\gamma})}{\sinh(b\sqrt{2\gamma})}\Big) = \sum_{m=-\infty}^{\infty} \frac{b - a + 2mb}{\sqrt{2\pi t^3}} e^{-\frac{(b-a+2mb)^2}{2t}}, \quad a < b, \tag{20}$$

($\mathcal{L}^{-1}_{t,\gamma}$ denotes the inverse Laplace transform).

Theorem 3.4.

(i) For the Brownian bridge with $|y| \geq h$,

$$\mathbb{E}_{0,T,y}[\mathcal{T}_{(-h,h)}] = \int_0^T \int_{-h}^h \frac{p_{T-t}(z, y)}{p_T(0, y)} \Delta(z, h, t) dz dt \tag{21}$$

and

$$\lim_{h \downarrow 0} \frac{\mathbb{E}_{0,T,y}[\mathcal{T}_{(-h,h)}]}{h^2} = 1, \quad y \neq 0. \tag{22}$$

(ii) For the 3-dimensional Bessel bridge with $x > h$ *and* $y \notin (x - h, x + h)$,

$$\mathbb{E}^{(3)}_{x,T,y}[\mathcal{T}_{(x-h,x+h)}] = \int_0^T \int_{-h}^h \frac{z + x}{x} \Delta(z, h, t) \frac{r^{(3)}_{T-t}(z + x, y)}{r^{(3)}_T(x, y)} dz dt$$

and

$$\lim_{h \downarrow 0} \frac{\mathbb{E}^{(3)}_{x,T,y}[\mathcal{T}_{(x-h,x+h)}]}{h^2} = 1. \tag{23}$$

(iii) If $y \in (-h, h)$ *(or* $y \in (x - h, x + h)$ *with* $x > h$, $y > 0$*) the mean of the first exit time is infinite for both bridges, i.e.*

$$\mathbb{E}_{0,T,y}[\mathcal{T}_{(-h,h)}] = \mathbb{E}^{(3)}_{x,T,y}[\mathcal{T}_{(x-h,x+h)}] = \infty. \tag{24}$$

Proof. (i) According to (11), (8), and (17),

$$\mathbb{E}_{0,T,y}[\mathcal{T}_{(-h,h)}] = \int_0^T \int_{-h}^{h} \frac{p_{T-t}(z,y)}{p_T(0,y)} \Delta(z,h,t) dz dt.$$

To show (22), substitute $z = hu$ and $t = h^2 s$, and notice that $\Delta(hu, h, h^2 s)h = \Delta(u, 1, s)$, so that

$$\mathbb{E}_{0,T,y}[\mathcal{T}_{(-h,h)}] = h^2 \int_0^{T/h^2} \int_{-1}^{1} \frac{p_{T-h^2 s}(hu,y)}{p_T(0,y)} \Delta(u,1,s) du ds.$$

To apply dominated convergence for $h \downarrow 0$ let y be fixed and assume that $h < \frac{|y|}{2}$. Then, for any $u \in (-1,1)$ it holds that $|y - hu| > |y| - h > \frac{|y|}{2}$. Moreover, notice that for $y \neq 0$ there exists a constant $C > 0$ such that

$$\sup_{t>0} p_t(0,y) = \frac{C}{|y|}. \tag{25}$$

This implies that

$$\sup_{u \in (-1,1),\, s \in (0,T/h^2)} \frac{p_{T-h^2 s}(hu,y)}{p_T(0,y)} \leq \frac{2\,C}{|y| p_T(0,y)}.$$

Therefore, since $(u,s) \mapsto \Delta(u,1,s)$ is integrable and $(u,s,h) \mapsto p_{T-h^2 s}(hu,y)/p_T(0,y)$ is bounded on $(-1,1) \times (0, T/h^2) \times (0, |y|/2)$, dominated convergence yields

$$\begin{aligned} \lim_{h \downarrow 0} \frac{\mathbb{E}_{0,T,y}[\mathcal{T}_{(-h,h)}]}{h^2} &= \int_0^\infty \int_{-1}^{1} \Delta(u,1,s) du ds \\ &= \int_0^\infty \int_{-1}^{1} \mathbb{P}_0(\mathcal{T}_{(-1,1)} > s, W_s \in du) ds \\ &= \mathbb{E}_0[\mathcal{T}_{(-1,1)}] \\ &= 1, \end{aligned} \tag{26}$$

where (2) is used for the last equality.
(ii) By (11), (8), and (18),

$$\mathbb{E}^{(3)}_{x,T,y}[\mathcal{T}_{(x-h,x+h)}] = \int_0^T \int_{x-h}^{x+h} \frac{z}{x} \Delta(z-x,h,t) \frac{r^{(3)}_{T-t}(z,y)}{r^{(3)}_T(x,y)} dz dt. \tag{27}$$

For the proof of (23) we substitute $t = h^2 s$ and $z = hu + x$ and recall that $\Delta(hu, h, h^2 s)h = \Delta(u, 1, s)$, so that

$$\mathbb{E}^{(3)}_{x,T,y}[\mathcal{T}_{(x-h,x+h)}] = h^2 \int_0^{T/h^2} \int_{-1}^{1} \frac{hu+x}{x} \Delta(u,1,s) \frac{r^{(3)}_{T-h^2 s}(hu+x,y)}{r^{(3)}_T(x,y)} du ds.$$

Using relation (15) yields

$$\left|\frac{hu+x}{x}\frac{r^{(3)}_{T-h^2s}(hu+x,y)}{r^{(3)}_T(x,y)}\right| = \left|\frac{p_{T-h^2s}(hu+x,y)-p_{T-h^2s}(hu+x,-y)}{p_T(x,y)-p_T(x,-y)}\right|.$$

To see that this expression is bounded in $s \in (0, T/h^2)$ and $u \in (-1,1)$ notice that if $h < |y-x|/2$ with $x > h$ and $y > 0$, then

$$|x-y+hu| > |x-y| - h > |x-y|/2 \quad \text{and} \quad |x+y+hu| > x+y-h > |x-y|/2$$

so that, by (25)

$$|p_{T-h^2s}(hu+x,y) - p_{T-h^2s}(hu+x,-y)| \le \frac{C}{|x-y+hu|} + \frac{C}{|x+y+hu|} < \frac{4C}{|x-y|}.$$

Hence, dominated convergence gives

$$\lim_{h\downarrow 0}\frac{\mathbb{E}^{(3)}_{x,T,y}[\mathcal{T}_{(x-h,x+h)}]}{h^2} = \int_0^\infty\int_{-1}^1 \Delta(u,1,s)duds = 1,$$

where the last equality was shown in (26) .
(iii) For a regular diffusion X, (24) follows from $\mathbb{P}^X_{x,T,y}(\mathcal{T}_{(x-h,x+h)} \ge T) > 0$. It holds

$$\mathbb{P}^X_{x,T,y}(\mathcal{T}_{(x-h,x+h)} \ge T) = \lim_{\varepsilon\downarrow 0}\frac{\mathbb{P}^X_x(\mathcal{T}_{(x-h,x+h)} \ge T, y \le X_T \le y+\varepsilon)}{\mathbb{P}^X_x(y \le X_T \le y+\varepsilon)} \tag{28}$$

and

$$\mathbb{P}^X_x(\mathcal{T}_{(x-h,x+h)} \ge T, y \le X_T \le y+\varepsilon) = \int_y^{y+\varepsilon}\widehat{q}_T(x,z)dz,$$

where $\widehat{q}$ is the transition density of X killed at $\mathcal{T}_{(x-h,x+h)}$. By [7, p.157] the transition density of any regular diffusion is strictly positive. In our case this means

$$\widehat{q}_T(x,z) > 0 \quad \text{for all} \quad |z| < h.$$

Using l'Hospital's rule in (28) yields

$$\mathbb{P}^X_{x,T,y}(\mathcal{T}_{(x-h,x+h)} \ge T) = \frac{\widehat{q}_T(x,y)}{q_T(x,y)} > 0.$$

□

The scaled first exit time $\mathcal{T}_{(-h,h)}/h^2$ of the Brownian bridge $B^{0,T,y}$ and the scaled first exit time $\mathcal{T}_{(x-h,x+h)}/h^2$ of the 3-dimensional Bessel bridge from $x > 0$ to $y > 0$ of length T both converge weakly to the first exit time $\mathcal{T}_{(-1,1)}$ of the standard Brownian motion as $h \downarrow 0$:

Theorem 3.5. *For each $y \neq 0$ and $T > 0$,*

$$\lim_{h \downarrow 0} \mathbb{E}_{0,T,y}\left[e^{-\gamma \mathcal{T}_{(-h,h)}/h^2}\right] = \mathbb{E}_0\left[e^{-\gamma \mathcal{T}_{(-1,1)}}\right], \quad \gamma > 0 \tag{29}$$

and for each $x, y > 0$,

$$\lim_{h \downarrow 0} \mathbb{E}^{(3)}_{x,T,y}\left[e^{-\gamma \mathcal{T}_{(x-h,x+h)}/h^2}\right] = \mathbb{E}_0\left[e^{-\gamma \mathcal{T}_{(-1,1)}}\right], \quad \gamma > 0.$$

Proof. We will sketch the proof of (29) since the other convergence follows similarly. Let

$$\mathcal{T}^X_c := \inf\{t > 0 : X_t = c\}$$

denote the first hitting time of c by a continuous process X. Since $\mathcal{T}_{(a,b)} = \mathcal{T}_a \wedge \mathcal{T}_b = \mathcal{T}_a \mathbb{1}_{\{\mathcal{T}_a < \mathcal{T}_b\}} + \mathcal{T}_b \mathbb{1}_{\{\mathcal{T}_a < \mathcal{T}_b\}}$ it suffices to consider $\mathbb{E}_{x,T,y}\left[\mathbb{1}_{\{\mathcal{T}_a < \mathcal{T}_b\}} e^{-\beta \mathcal{T}_a}\right]$. We recall that by (5),

$$Z_t(\omega) := \left.\frac{d\mathbb{P}_{x,T,y}}{d\mathbb{P}_x}\right|_{\mathcal{C}_{t+}}(\omega) = \frac{p_{T-t}(\omega(t), y)}{p_T(x,y)}, \quad t < T.$$

Since $(Z_t)_{t \in [0,T)}$ is a $(\mathbb{P}_x, (\mathcal{C}_t)_{t \in [0,T)})$-martingale and the random variable $\mathbb{1}_{\{\mathcal{T}_a \le t, \mathcal{T}_a < \mathcal{T}_b\}} e^{-\beta \mathcal{T}_a}$ (with $\beta > 0$) is $\mathcal{C}_{\mathcal{T}_a \wedge t}$ -measurable, we conclude that for $a < x < b$

$$\begin{aligned}
\mathbb{E}_{x,T,y}\left[\mathbb{1}_{\{\mathcal{T}_a \le t, \mathcal{T}_a < \mathcal{T}_b\}} e^{-\beta \mathcal{T}_a}\right] &= \mathbb{E}_x\left[\mathbb{1}_{\{\mathcal{T}_a \le t, \mathcal{T}_a < \mathcal{T}_b\}} e^{-\beta \mathcal{T}_a} Z_t\right] \\
&= \mathbb{E}_x\left[\mathbb{E}_x\left[\mathbb{1}_{\{\mathcal{T}_a \le t, \mathcal{T}_a < \mathcal{T}_b\}} e^{-\beta \mathcal{T}_a} Z_t \mid \mathcal{F}_{\mathcal{T}_a \wedge t}\right]\right] \\
&= \mathbb{E}_x\left[\mathbb{1}_{\{\mathcal{T}_a \le t, \mathcal{T}_a < \mathcal{T}_b\}} e^{-\beta \mathcal{T}_a} Z_{\mathcal{T}_a \wedge t}\right] \\
&= \mathbb{E}_x\left[\mathbb{1}_{\{\mathcal{T}_a \le t, \mathcal{T}_a < \mathcal{T}_b\}} e^{-\beta \mathcal{T}_a} \tfrac{p_{T-\mathcal{T}_a \wedge t}(W_{\mathcal{T}_a \wedge t}, y)}{p_T(x,y)}\right] \\
&= \int_0^t e^{-\beta r} \frac{p_{T-r}(a,y)}{p_T(x,y)} \mathbb{P}_x(\mathcal{T}_a \in dr, \mathcal{T}_a < \mathcal{T}_b).
\end{aligned}$$

Consequently,

$$\begin{aligned}
\mathbb{E}_{0,T,y}\left[e^{-\gamma \mathcal{T}_{(-h,h)}/h^2}\right] = \int_0^T e^{-\gamma r/h^2} \Bigg[&\frac{p_{T-r}(-h,y)}{p_T(0,y)} \mathbb{P}_0(\mathcal{T}_{-h} \in dr, \mathcal{T}_{-h} < \mathcal{T}_h) \\
&+ \frac{p_{T-r}(h,y)}{p_T(0,y)} \mathbb{P}_0(\mathcal{T}_h \in dr, \mathcal{T}_h < \mathcal{T}_{-h}) \Bigg].
\end{aligned}$$

The result follows by the variable transform $t = r/h^2$ and dominated convergence as $h \downarrow 0$. □

Remark 3.6. For the 3-dimensional Bessel bridge starting in 0 with $y > h$ it holds

$$\mathbb{E}^{(3)}_{0,T,y}[\mathcal{T}_h] = \int_0^T \int_0^h \mathbb{P}^{(3)}_0\left(\sup_{0 \le s \le t} X_s < h, X_t \in dz\right) \frac{r^{(3)}_{T-t}(z,y)}{r^{(3)}_T(0,y)} dt. \tag{30}$$

The representation follows from (8) and (10), which can be extended to $x = 0$ since 0 is an entrance boundary point. Similarly as above it can be shown that

$$\lim_{h \downarrow 0} \frac{\mathbb{E}^{(3)}_{0,T,y}[\mathcal{T}_h]}{h^2} = \lim_{h \downarrow 0} \frac{\mathbb{E}^{(3)}_{0}[\mathcal{T}_h]}{h^2} = \frac{1}{3}.$$

Indeed, by [2, Part II, Section 5, 1.1.8 (p. 435)] we have for $0 < x \vee z < h$

$$\mathbb{P}^{(3)}_x(\sup_{0 \le s \le t} X_s < h, X_t \in dz) = \frac{z}{x\sqrt{2\pi t}} \sum_{m=-\infty}^{\infty} \left(e^{-\frac{(z-x+2mh)^2}{2t}} - e^{-\frac{(z+x+2mh)^2}{2t}} \right) dz$$

which implies for $x \downarrow 0$

$$\mathbb{P}^{(3)}_0(\sup_{0 \le s \le t} X_s < h, X_t \in dz) \quad = \quad 2z\mathrm{ss}_t(h-z, h)dz,$$

where ss_t was defined in (20). Then, relations (30) and (19) yield

$$\begin{aligned}
\lim_{h \downarrow 0} \frac{\mathbb{E}^{(3)}_{0,T,y}[\mathcal{T}_h]}{h^2} &= 2\int_0^1 u \int_0^\infty \mathbb{P}_u(\mathcal{T}_0 \in dr, \mathcal{T}_0 < \mathcal{T}_1)du \\
&= 2\int_0^1 u\mathbb{P}_u(\mathcal{T}_0 < \mathcal{T}_1)du \\
&= 2\int_0^1 u(1-u)du \\
&= \frac{1}{3}.
\end{aligned}$$

The other limit follows in the same way.

3.4 The Kolmogorov distribution function and the mean of the first exit time

A classical result due to Doob [4] is that the distribution of the supremum of the absolute value of the standard Brownian bridge is given by

$$F(h) := \mathbb{P}_{0,1,0}\left(\sup_{0 \le s \le 1} \left| B^{0,1,0}_s \right| \le h \right) = \sum_{m=-\infty}^{\infty} (-1)^m e^{-2m^2h^2} = \sqrt{2\pi}\Delta(0, h, 1).$$

The function $h \mapsto F(h)$, $h > 0$, is called the Kolmogorov distribution function due to Kolmogorov's fundamental work [9] (and also Smirnov [14]) on empirical distributions. We refer also to [12] and [10, Section 5.7]. The main result of this section – Theorem 3.8 – provides representations for the mean of the first exit time of the Brownian bridge and of the 3-dimensional Bessel bridge involving the Kolmogorov distribution function. We present now some formulas related to the Kolmogorov distribution which we need later. The following Jacobi's theta function indentity, an instance of the Poisson summation formula, is stated in [1, equation (2.1)]):

$$\sum_{m=-\infty}^{\infty} \cos(2m\pi v) e^{-m^2\pi^2 u} = \frac{1}{\sqrt{\pi u}} \sum_{m=-\infty}^{\infty} e^{-\frac{(m+v)^2}{u}}, \quad u > 0, v \in \mathbb{R}.$$

Putting here $u = 2x^2/\pi^2$ and $v = 1/2$ yields

$$F(x) = \sum_{m=-\infty}^{\infty} (-1)^m e^{-2m^2x^2} = \frac{\sqrt{2\pi}}{x} \sum_{k=1}^{\infty} \exp\left(-\frac{(2k-1)^2\pi^2}{8x^2}\right). \tag{31}$$

Notice also that

$$\begin{aligned} F(h/\sqrt{t}) &= \mathbb{P}_{0,1,0}\left(\sup_{0\le s\le 1} \left|B_s^{0,1,0}\right| \le h/\sqrt{t}\right) \\ &= \mathbb{P}_{0,t,0}\left(\sup_{0\le s\le 1} \left|B_{ts}^{0,t,0}\right| \le h\right) \\ &= \mathbb{P}_{0,t,0}\left(\sup_{0\le s\le t} \left|B_s^{0,t,0}\right| \le h\right), \end{aligned}$$

where it is used that

$$\left(\sqrt{t}B_s^{0,1,0}\right)_{0\le s\le 1} \stackrel{d}{=} \left(\,B_{st}^{0,t,0}\right)_{0\le s\le 1},$$

which can be seen by applying the scaling property of Brownian motion to the representation (6). Consequently,

$$F(h/\sqrt{t}) = \mathbb{P}_{0,t,0}(\mathcal{T}_{(-h,h)} > t) = \mathbb{P}_{0,t,0}(\mathcal{T}_{(-h,h)} = \infty).$$

In the following, we will denote the Laplace transform of a function f by

$$\mathcal{L}_{t,\gamma}(f(t)) := \int_0^\infty e^{-\gamma t} f(t)dt, \quad \gamma > 0.$$

Notice that we also indicate the integration variable t.

Lemma 3.7. *For* $|y| \ge h$,

$$\int_0^T \int_{-h}^h p_{T-t}(x,y)\Delta(x,h,t)dxdt = h\int_0^T p_{T-t}(0,y)\Delta(0,h,t)dt. \tag{32}$$

Proof. We prove Lemma 3.7 by showing that the Laplace transforms of the two sides of (32) coincide. Using Fubini's theorem and the convolution formula, we get for the Laplace transform of the l.h.s. of (32)

$$\mathcal{L}_{T,\gamma}\Big(\int_0^T \int_{-h}^h p_{T-t}(x,y)\Delta(x,h,t)dxdt\Big) = \int_{-h}^h \mathcal{L}_{t,\gamma}(p_t(x,y))\mathcal{L}_{t,\gamma}(\Delta(x,h,t))dx. \tag{33}$$

To compute the second Laplace transform expression of the r.h.s. of (33) we use the series representation (16). For $x \in (-h, h)$ it holds that

$$\left| \sum_{|m| \geq 2} (-1)^m e^{-\frac{2mh(mh-x)}{t}} \right| \leq \sum_{|m| \geq 2} e^{-\frac{2mh(mh-x)}{t}} \leq 2 \sum_{m \geq 2} \left(e^{-\frac{2h^2}{t}} \right)^m \leq \frac{2}{1 - e^{\frac{-2h^2}{t}}}.$$

Hence one may interchange the summation and the Laplace transform, and since

$$\mathcal{L}_{t,\gamma}(p_t(x,z)) = \frac{1}{\sqrt{2\gamma}} e^{-\sqrt{2\gamma}|z-x|},$$

one gets

$$\begin{aligned}
\mathcal{L}_{t,\gamma}(\Delta(x,h,t)) &= \mathcal{L}_{t,\gamma}\left(\sum_{m=-\infty}^{\infty} (-1)^m e^{-\frac{2mh(mh-x)}{t}} \frac{e^{-\frac{x^2}{2t}}}{\sqrt{2\pi t}} \right) \\
&= \sum_{m=-\infty}^{\infty} (-1)^m \mathcal{L}_{t,\gamma}\left(p_t(x, 2mh)\right) \\
&= \frac{1}{\sqrt{2\gamma}} \left(e^{-\sqrt{2\gamma}|x|} + (e^{\sqrt{2\gamma}x} + e^{-\sqrt{2\gamma}x}) \sum_{m=1}^{\infty} (-1)^m e^{-\sqrt{2\gamma}2mh} \right) \\
&= \frac{1}{\sqrt{2\gamma}} \left(e^{-\sqrt{2\gamma}|x|} - (e^{\sqrt{2\gamma}x} + e^{-\sqrt{2\gamma}x}) \frac{e^{-\sqrt{2\gamma}2h}}{1 + e^{-\sqrt{2\gamma}2h}} \right).
\end{aligned}$$

By a straightforward calculation, this yields for (33)

$$\begin{aligned}
\int_{-h}^{h} \mathcal{L}_{t,\gamma}(p_t(x,y)) \mathcal{L}_{t,\gamma}(\Delta(x,h,t)) dx &= \frac{h}{2\gamma} e^{-\sqrt{2\gamma}|y|} \frac{1 - e^{-\sqrt{2\gamma}\,2h}}{1 + e^{-\sqrt{2\gamma}\,2h}} \\
&= \frac{h}{2\gamma} e^{-\sqrt{2\gamma}|y|} \tanh(h\sqrt{2\gamma}). \qquad (34)
\end{aligned}$$

We continue with the Laplace transform of the r.h.s. of (32). From (31) we get

$$\begin{aligned}
\mathcal{L}_{t,\gamma}(\Delta(0,h,t)) &= \frac{1}{h} \sum_{k=1}^{\infty} \int_0^{\infty} \exp\left(-\frac{(2k-1)^2 \pi^2 t}{8h^2} \right) \exp(-\gamma t) dt \\
&= h \sum_{k=1}^{\infty} \frac{8}{(2k-1)^2 \pi^2 + 8h^2 \gamma} \\
&= \frac{\tanh(h\sqrt{2\gamma})}{\sqrt{2\gamma}}, \qquad (35)
\end{aligned}$$

where we use

$$\tanh\left(\frac{\pi x}{2}\right) = \frac{4x}{\pi} \sum_{k=1}^{\infty} \frac{1}{(2k-1)^2 + x^2}, \quad x \in \mathbb{R}$$

(see [6, Subsection 1.421]). From the convolution formula we conclude that

$$\mathcal{L}_{T,\gamma}\left(h\int_0^T p_{T-t}(0,y)\Delta(0,h,t)dt\right) = h\mathcal{L}_{t,\gamma}\left(p_t(0,y)\right)\mathcal{L}_{t,\gamma}(\Delta(0,h,t))$$
$$= \frac{h}{\sqrt{2\gamma}}e^{-\sqrt{2\gamma}|y|}\,\frac{\tanh(h\sqrt{2\gamma})}{\sqrt{2\gamma}}. \tag{36}$$

We see that (34) and (36) coincide which finishes the proof of Lemma 3.7.

□

Theorem 3.8.

(i) For the Brownian bridge with $|y| \geq h$,

$$\mathbb{E}_{0,T,y}[\mathcal{T}_{(-h,h)}] = h\int_0^T \frac{p_{T-t}(0,y)}{p_T(0,y)}\Delta(0,h,t)dt. \tag{37}$$

(ii) For the 3-dimensional Bessel bridge with positive $y \notin (x-h, x+h)$ *and* $x > h$,

$$\mathbb{E}^{(3)}_{x,T,y}[\mathcal{T}_{(x-h,x+h)}] = h\int_0^T \frac{r^{(3)}_{T-t}(x,y)}{r^{(3)}_T(x,y)}\Delta(0,h,t)dt. \tag{38}$$

Proof. (i) We derive (37) from (10), (8) and Lemma 3.7.

(ii) Recall (27). Since by (15) it holds

$$\frac{r^{(3)}_{T-t}(z,y)}{r^{(3)}_T(x,y)} = \frac{x}{z}\,\frac{p_{T-t}(z,y)-p_{T-t}(z,-y)}{p_T(x,y)-p_T(x,-y)}, \tag{39}$$

we have

$$\begin{aligned}
&(p_T(x,y)-p_T(x,-y))\mathbb{E}^{(3)}_{x,T,y}[\mathcal{T}_{(x-h,x+h)}]\\
&= \int_0^T\int_{x-h}^{x+h}\Delta(z-x,h,t)(p_{T-t}(z,y)-p_{T-t}(z,-y))dzdt\\
&= \int_0^T\int_{-h}^{h}\Delta(u,h,t)(p_{T-t}(u,y-x)-p_{T-t}(u,-(y+x)))dudt\\
&= h\int_0^T\Delta(0,h,t)(p_{T-t}(0,y-x)-p_{T-t}(0,y+x))dt,
\end{aligned}$$

where the last equality is implied by (32). Then (39) yields (38). □

Remark 3.9. (i) We have not been able to find a probabilistic explanation for the appearance of the Kolmogorov distribution function in the representations (37) and (38).

(ii) Notice that the integrands w.r.t. t of the expressions in [Theorem 3.4, equation (21)] and in [Theorem 3.8, equation (37)] do not coincide, that is

$$\int_{-h}^{h} p_{T-t}(z,y)\Delta(z,h,t)dz \neq h p_{T-t}(0,y)\Delta(0,h,t).$$

To see this, one can compare the double Laplace transform of

$$h\, p_{T-t}(0,y)\Delta(0,h,t) \quad \text{and} \quad \int_{-h}^{h} p_{T-t}(z,y)\Delta(z,h,t)dz.$$

A similar computation as in the proof of Lemma 3.7 yields for $\lambda > 0$

$$\mathcal{L}_{T,\gamma}\left(h\int_0^T p_{T-t}(0,y)\Delta(0,h,t)dz e^{-\lambda t}dt\right) = \frac{h}{\sqrt{2\gamma}} e^{-\sqrt{2\gamma}|y|}\, \frac{\tanh(h\sqrt{2(\gamma+\lambda)})}{\sqrt{2(\gamma+\lambda)}}.$$

On the other hand, after some calculations,

$$\mathcal{L}_{T,\gamma}\left(\int_0^T \int_{-h}^{h} p_{T-t}(z,y)\Delta(z,h,t)dz e^{-\lambda t}dt\right)$$
$$= \frac{1}{\lambda\sqrt{2\gamma}} e^{-\sqrt{2\gamma}|y|}\left[1 - \frac{e^{-\sqrt{2\gamma}h} + e^{\sqrt{2\gamma}h}}{e^{-\sqrt{2(\gamma+\lambda)}h} + e^{\sqrt{2(\gamma+\lambda)}h}}\right].$$

(iii) It is also possible to use (37) and (38) to show (22) and (23). For example, after the variable transform $t = h^2 r$ one finds an integrable majorant and concludes by dominated convergence that

$$h\int_0^T \frac{p_{T-t}(0,y)}{p_T(0,y)}\Delta(0,h,t)dt \to \int_0^\infty \Delta(0,1,t)dt, \quad h \downarrow 0.$$

Then by monotone convergence, thanks to relation (35) one obtains

$$\int_0^\infty \Delta(0,1,t)dt = \lim_{\gamma\to 0+}\int_0^\infty e^{-\gamma t}\Delta(0,1,t)dt = \lim_{\gamma\to 0+}\frac{\tanh(\sqrt{2\gamma})}{\sqrt{2\gamma}} = 1.$$

We conclude this section by pointing out a connection between the Kolmogorov distribution function and the density of the last visit of 0 by a Brownian motion before $\mathcal{T}_{(-h,h)}$. This connection has been noticed by Knight in [8, Corollary 2.1].

We provide here a different proof for this fact.

Proposition 3.10. *When $\widehat{W}$ is a Brownian motion killed at $\mathcal{T}_{(-h,h)}$, then $t \mapsto \frac{1}{h}\Delta(0,h,t)$ is the density of the last passage time $\lambda_0 := \sup\{t > 0 : \widehat{W}_t = 0\}$.*

Proof. It is well known that (cf. [2, Part I, Chapter II, No 20, p. 26])

$$\mathbb{P}_0^{\widehat{W}}(\lambda_0 \in dt) = \frac{\widehat{q}_t(0,0)}{\widehat{G}_0(0,0)} dt, \tag{40}$$

where

$$\widehat{q}_t(x,y) = \frac{1}{\sqrt{2\pi t}} \sum_{m=-\infty}^{\infty} \left(e^{-\frac{(x-y+4mh)^2}{2t}} - e^{-\frac{(x+y+(2m+1)2h)^2}{2t}} \right)$$

is the transition density (w.r.t. the Lebesgue measure) of $\widehat{W}$ (cf. [2, Part I, Appendix I, No 6, p. 126]), and

$$\widehat{G}_0(x,y) = \begin{cases} \dfrac{(x+h)(h-y)}{h}, & -h \le x \le y \le h, \\[2ex] \dfrac{(y+h)(h-x)}{h}, & -h \le y \le x \le h, \end{cases}$$

denotes the 0-resolvent kernel (see [2, Part I, Appendix I, No 6, p. 126]). Notice that $\widehat{q}_t(0,0) = \Delta(0,h,t)$ and $\widehat{G}_0(0,0) = h$ so that from (40) we get

$$\mathbb{P}_0^{\widehat{W}}(\lambda_0 \in dt)/dt = \frac{\Delta(0,h,t)}{h}.$$

□

4 Application

In this section we apply our previous results to rigorously prove in Corollary 4.3 an estimate needed by Walsh in [15]. The convergence analysis there succeeds to identify the leading constants C_1, C_2 appearing in the error expansion

$$\mathbb{E}[g(X_n^{(n)}) - g(X_T)] = C_1\sqrt{\tfrac{T}{n}} + C_2\tfrac{T}{n} + O(\tfrac{1}{n^{3/2}}), \quad n \text{ even} \tag{41}$$

(cf. [15, equation (14)]) in terms of expressions depending on the function g (which is assumed to be exponentially bounded and piecewise twice continuously differentiable). Here $X_t = \sigma W_t$ for $t \geq 0$ and $(X_k^{(n)})_{k=0}^n$ denotes a symmetric simple random walk with time step T/n and space step size $\sigma\sqrt{T/n}$. For simplicity, we will put $\sigma = 1$ and hence consider

$$\mathbb{E}[g(W_n^{(n)}) - g(W_T)].$$

The central idea for this error analysis is to build the random walk from a given Brownian motion $(W_t)_{t\geq 0}$ so that both processes are on the same probability space: Fix $T > 0$ and $n \in \mathbb{N}$. For $h := \sqrt{T/n}$ define $\tau_0 := 0$ and

$$\tau_k := \inf\{t > \tau_{k-1} : |W_t - W_{\tau_{k-1}}| = h\}, \quad k \geq 1.$$

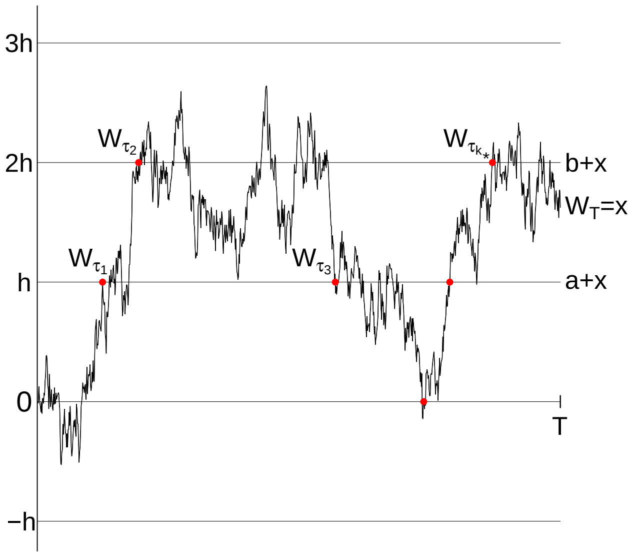
Then $(W_{\tau_k} - W_{\tau_{k-1}})_{k=1}^{\infty}$ is a sequence of i.i.d. random variables with $\mathbb{P}(W_{\tau_k} - W_{\tau_{k-1}} = h) = \mathbb{P}(W_{\tau_k} - W_{\tau_{k-1}} = -h) = \frac{1}{2}$. Let k_* be such that $\tau_{k_*} \leq T < \tau_{k_*+1}$. One central idea to get (41) is to split the error into a local part (exploiting the closeness of W_T and $W_{\tau_{k_*+1}}$) and a global part (using binomial distribution),

$$\mathbb{E}[(g(W_T) - g(W_{\tau_{k_*+1}}))\mathbb{1}_{\{k_* \text{ odd }\}}] + \mathbb{E}[(g(W_{\tau_{k_*+1}}) - g(W_{\tau_n})\mathbb{1}_{\{k_* \text{ odd }\}}],$$

while the splitting is done with $W_{\tau_{k_*+2}}$ if k_* is even. For this, in [15, Section 9] the conditional probability

$$q(x) := \mathbb{P}(k_* \text{ is even } | W_T = x), \quad x \in \mathbb{R}$$

is introduced (there k_* is denoted by L). We want to study and estimate q. The reason to consider only even n and even $k_* + 1$ or $k_* + 2$ is to avoid even/odd fluctuations known to appear in binomial tree schemes. Notice that k_* is an even number if and only if $W_{\tau_{k_*}}$ is an even multiple of h. The process $(W_t)_{0 \leq t \leq T}$ given $W_0 = 0, W_T = x$ is identical in law with a Brownian bridge from 0 to x and of length T. We denote this bridge by $\left(B_t^{0,T,x}\right)_{0 \leq t \leq T}$. By time reversion, we get the Brownian bridge $\left(B_t^{x,T,0}\right)_{0 \leq t \leq T}$.



We fix $k \in \mathbb{Z}$ and assume $x \in ((2k-1)h, (2k+1)h)$. For simplicity put

$$\underline{h} := (2k-1)h, \quad \overline{h} := (2k+1)h, \quad h_e := 2kh \tag{42}$$

for the lower and upper value, and for the 'even' midpoint of the interval. Then, given $W_T = x$, we have that 'k_* is even' is the same as '$W_{\tau_{k_*}} = h_e$'. For the time-reversed bridge associated to $\mathbb{P}_{x,T,0}$ we get

$$q(x) = \mathbb{P}(k_* \text{ is even} | W_T = x) = \begin{cases} \mathbb{P}_{x,T,0}(\mathcal{T}_{h_e} < \mathcal{T}_{\underline{h}}), & \underline{h} < x < h_e, \\ \mathbb{P}_{x,T,0}(\mathcal{T}_{h_e} < \mathcal{T}_{\overline{h}}), & h_e < x < \overline{h}, \end{cases}$$

where $\mathcal{T}_y(\omega) := \inf\{t > 0 : \omega(t) = y\}$ for $\omega \in C[0,T]$. For the time-reversed and by $-x$ shifted bridge this means it hits the 'shifted even line' $h_e - x$ before the shifted odd one:

$$\mathbb{P}(k_* \text{ is even} | W_T = x) = \begin{cases} \mathbb{P}_{0,T,-x}(\mathcal{T}_{h_e - x} < \mathcal{T}_{\underline{h} - x}), & \underline{h} < x < h_e, \\ \mathbb{P}_{0,T,-x}(\mathcal{T}_{h_e - x} < \mathcal{T}_{\overline{h} - x}), & h_e < x < \overline{h}. \end{cases}$$

For any $a < 0 < b$ and $y \notin (a,b)$ it holds

$$\mathbb{E}_{0,T,y}\big[B_{\mathcal{T}_{(a,b)}}\big] \;=\; a(1-\mathbb{P}_{0,T,y}(\mathcal{T}_b < \mathcal{T}_a)) + b\,\mathbb{P}_{0,T,y}(\mathcal{T}_b < \mathcal{T}_a).$$

Consequently,

$$\mathbb{P}_{0,T,y}(\mathcal{T}_b < \mathcal{T}_a) = \frac{-a}{b-a} + \frac{\mathbb{E}_{0,T,y}\big[B_{\mathcal{T}_{(a,b)}}\big]}{b-a}.$$

Hence

$$q(x) = \begin{cases} \dfrac{x-\underline{h}}{h} + \dfrac{\mathbb{E}_{0,T,-x}\big[B_{\mathcal{T}_{(\underline{h}-x,h_e-x)}}\big]}{h}, & \underline{h} < x < h_e, \\ \dfrac{\overline{h}-x}{h} - \dfrac{\mathbb{E}_{0,T,-x}\big[B_{\mathcal{T}_{(h_e-x,\overline{h}-x)}}\big]}{h}, & h_e < x < \overline{h}. \end{cases} \tag{43}$$

Arguing that Brownian bridge and Brownian motion have a similar exit behavior for small h, in [15, equation (20)] it is stated that

$$q(x) = \frac{\operatorname{dist}(x,\mathbb{N}_o^h)}{h} + O(h), \tag{44}$$

where $x \in \mathbb{R}$, $\mathbb{N}_o^h = \{(2k+1)h : k \in \mathbb{Z}\}$ and $\operatorname{dist}(x,\mathbb{N}_o^h) := \inf\{|x-y| : y \in \mathbb{N}_o^h\}$. If one compares (44) with (43) one notices that

$$\operatorname{dist}(x,\mathbb{N}_o^h) \quad \text{is equal to} \quad x-\underline{h} \text{ or } \overline{h}-x,$$

so that we should have

$$q(x) - \frac{\operatorname{dist}(x,\mathbb{N}_o^h)}{h} = \frac{\mathbb{E}_{0,T,-x}[B_{\mathcal{T}(x)}]}{h} = O(h), \tag{45}$$

where (42) was used to rewrite $\mathcal{T}_{(\underline{h}-x,h_e-x)}$ and $\mathcal{T}_{(h_e-x,\overline{h}-x)}$ as

$$\mathcal{T}(x) := \mathcal{T}_{(kh-x,(k+1)h-x)} \text{ if } x \in (kh,(k+1)h). \tag{46}$$

We will show below that (45) is indeed true for a fixed x. However, we will see that this relation does not hold uniformly in x, and therefore it can not be used to estimate the second term on the r.h.s. of (51) below. We will estimate that second term in Corollary 4.3 below.

Lemma 4.1. *Suppose that $a < 0 < b$, $y \notin (a,b)$, $T > 0$ and $h > 0$. Then*

$$\big|\mathbb{E}_{0,T,y}\big[B_{\mathcal{T}_{(a,b)}}\big]\big| \le \frac{\mathbb{E}_{0,T,y}[\mathcal{T}_{(a,b)}]}{T}\left(2(|a|\vee b) + |y| + 3\sqrt{2T}\right). \tag{47}$$

Proof. Restricting the expectation onto the set $\{\mathcal{T}_{(a,b)} > c\}$, (where we choose $c = \frac{T}{2}$ to simplify the computation for the restriction onto $\{\mathcal{T}_{(a,b)} \le c\}$) we derive by Markov's inequality that

$$\Big|\mathbb{E}_{0,T,y}\big[B_{\mathcal{T}_{(a,b)}}\mathbb{1}_{\{\mathcal{T}_{(a,b)}>T/2\}}\big]\Big| \le (|a|\vee b)\mathbb{P}_{0,T,y}(\mathcal{T}_{(a,b)} > T/2) \le \frac{2(|a|\vee b)}{T}\mathbb{E}_{0,T,y}[\mathcal{T}_{(a,b)}].$$

To estimate $\left|\mathbb{E}_{0,T,y}\left[B_{\mathcal{T}_{(a,b)}}\mathbb{1}_{\{\mathcal{T}_{(a,b)}\le T/2\}}\right]\right|$ we let

$$\tilde{B}_t^{0,T,y} := (T-t)\int_0^t \frac{dW_s}{T-s} + \frac{t}{T}y, \quad t\in[0,T), \quad \text{and} \quad \tilde{B}_T^{0,T,y} := y.$$

Then $\left(\tilde{B}_t^{0,T,y}\right)_{0\le t\le T} \stackrel{d}{=} \left(B_t^{0,T,y}\right)_{0\le t\le T}$ (cf. (7)). Setting

$$\tilde{\mathcal{T}} := \inf\{t\in[0,T] : \tilde{B}_t^{0,T,y} \notin (a,b)\}$$

yields $\mathbb{E}\left[\tilde{B}_{\tilde{\mathcal{T}}}^{0,T,y}\mathbb{1}_{\{\tilde{\mathcal{T}}\le T/2\}}\right] = \mathbb{E}_{0,T,y}\left[B_{\mathcal{T}_{(a,b)}}\mathbb{1}_{\{\mathcal{T}_{(a,b)}\le T/2\}}\right]$ and

$$\begin{aligned}
&\left|\mathbb{E}\left[\tilde{B}_{\tilde{\mathcal{T}}}^{0,T,y}\mathbb{1}_{\{\tilde{\mathcal{T}}\le T/2\}}\right]\right| \\
&\le \left|\mathbb{E}\left[(T-\tilde{\mathcal{T}})\int_0^{\tilde{\mathcal{T}}}\frac{dW_s}{T-s}\mathbb{1}_{\{\tilde{\mathcal{T}}\le T/2\}}\right]\right| + \frac{|y|}{T}\mathbb{E}[\tilde{\mathcal{T}}] \\
&\le T\left|\mathbb{E}\left[\int_0^{\tilde{\mathcal{T}}\wedge\frac{T}{2}}\frac{dW_s}{T-s}\,\mathbb{1}_{\{\tilde{\mathcal{T}}\le T/2\}}\right]\right| + \mathbb{E}\left|(\tilde{\mathcal{T}}\wedge(T/2))\int_0^{\tilde{\mathcal{T}}\wedge\frac{T}{2}}\frac{dW_s}{T-s}\right| + \frac{|y|}{T}\mathbb{E}[\tilde{\mathcal{T}}].
\end{aligned}$$

To benefit from Markov's inequality again we notice that the optional stopping theorem yields $\mathbb{E}\int_0^{\tilde{\mathcal{T}}\wedge\frac{T}{2}}\frac{dW_s}{T-s} = 0$ so that

$$\left|\mathbb{E}\left[\int_0^{\tilde{\mathcal{T}}\wedge\frac{T}{2}}\frac{dW_s}{T-s}\,\mathbb{1}_{\{\tilde{\mathcal{T}}\le T/2\}}\right]\right| = \left|\mathbb{E}\left[\int_0^{\tilde{\mathcal{T}}\wedge\frac{T}{2}}\frac{dW_s}{T-s}\,\mathbb{1}_{\{\tilde{\mathcal{T}}> T/2\}}\right]\right|.$$

Then by the inequalities of Cauchy-Schwarz and Markov,

$$\begin{aligned}
\left|\mathbb{E}\left[\mathbb{1}_{\{\tilde{\mathcal{T}}>T/2\}}\int_0^{\tilde{\mathcal{T}}\wedge\frac{T}{2}}\frac{dW_s}{T-s}\right]\right| &\le \left(\frac{2}{T}\mathbb{E}[\tilde{\mathcal{T}}]\right)^{\frac{1}{2}}\left(\mathbb{E}\left[\int_0^{\tilde{\mathcal{T}}\wedge\frac{T}{2}}\frac{ds}{(T/2)^2}\right]\right)^{\frac{1}{2}} \\
&\le \left(\frac{2}{T}\mathbb{E}[\tilde{\mathcal{T}}]\right)^{\frac{1}{2}}\left(\left(\frac{2}{T}\right)^2\mathbb{E}\left(\tilde{\mathcal{T}}\wedge\frac{T}{2}\right)\right)^{\frac{1}{2}} \\
&\le (\sqrt{2/T})^3\,\mathbb{E}[\tilde{\mathcal{T}}].
\end{aligned}$$

Again by the Cauchy-Schwarz inequality,

$$\begin{aligned}
\mathbb{E}\left|\left(\tilde{\mathcal{T}}\wedge\frac{T}{2}\right)\int_0^{\tilde{\mathcal{T}}\wedge\frac{T}{2}}\frac{dW_s}{T-s}\right| &\le \left(\mathbb{E}\left(\tilde{\mathcal{T}}\wedge\frac{T}{2}\right)^2\right)^{\frac{1}{2}}\left(\mathbb{E}\left[\int_0^{\tilde{\mathcal{T}}\wedge\frac{T}{2}}\frac{ds}{(T-s)^2}\right]\right)^{\frac{1}{2}} \\
&\le \left(\frac{T}{2}\mathbb{E}[\tilde{\mathcal{T}}]\right)^{\frac{1}{2}}\left(\left(\frac{2}{T}\right)^{\frac{1}{2}}\mathbb{E}\left(\tilde{\mathcal{T}}\wedge\frac{T}{2}\right)\right)^{\frac{1}{2}} \\
&\le \sqrt{2/T}\,\mathbb{E}[\tilde{\mathcal{T}}].
\end{aligned}$$

From the above estimates and $\mathbb{E}_{0,T,y}[\mathcal{T}_{(a,b)}] = \mathbb{E}[\tilde{\mathcal{T}}]$ we get (47). □

Now we derive estimates for $\mathbb{E}_{0,T,y}[\mathcal{T}_{(a,b)}]$.

Lemma 4.2. *Let $T > 0$ be fixed. Suppose that $a < 0 < b$ and $y \notin (a, b)$.*

(i) It holds

$$\mathbb{E}_{0,T,y}\left[\mathcal{T}_{(a,b)}\right] \;\leq\; \begin{cases} [b(2|a| + y)] \wedge T & \textit{if} \quad y \geq b, \\ [|a|(2b + |y|)] \wedge T & \textit{if} \quad y \leq a. \end{cases} \tag{48}$$

(ii) Define for $|y| \geq h$

$$C(T, h, y) := \frac{\mathbb{E}_{0,T,y}\left[\mathcal{T}_{(-h,h)}\right]}{h^2}. \tag{49}$$

Then

$$\mathbb{E}_{0,T,y}\left[\mathcal{T}_{(a,b)}\right] \;\leq C(b-a)^2 \quad \textit{if } |y| \geq b - a,$$

where $C = C(T, b-a, y)$, and it holds

$$\lim_{b-a \to 0,\, a<0<b} C(T, b-a, y) = 1.$$

Proof. (i) We first assume that $y \geq b$. We let

$$\bar{B}^{0,T,y}_t := \Big(1 - \frac{t}{T}\Big) W_{\frac{Tt}{T-t}} + y\frac{t}{T}, \qquad t \in [0, T)$$

and $\bar{B}^{0,T,y}_T := y$. Then $(\bar{B}^{0,T,y}_t)_{0\leq t\leq T} \overset{d}{=} (B^{0,T,y}_t)_{0\leq t\leq T}$ (cf. (6)). Set also

$$\bar{\mathcal{T}} := \inf\{t \in [0, T] : \bar{B}^{0,T,y}_t \notin (a, b)\}.$$

Since by definition $\mathcal{T}_{(a,b)} \leq T$ if $y \notin (a, b)$ one has $\mathcal{T}_{(a,b)} \overset{d}{=} \bar{\mathcal{T}} \wedge T$. With the aid of the change of variable $u = \frac{Tt}{T-t}$, $t \in [0, T)$, $u \geq 0$, one gets

$$\begin{aligned} \mathbb{E}_{0,T,y}[\mathcal{T}_{(a,b)}] &= \mathbb{E}_0\left[\inf\left\{t \in [0, T] : W_{\frac{Tt}{T-t}} \notin \left(\frac{a - \frac{yt}{T}}{1 - \frac{t}{T}}, \frac{b - \frac{yt}{T}}{1 - \frac{t}{T}}\right)\right\} \wedge T\right] \\ &= \mathbb{E}_0\left[\inf\left\{u \geq 0 : W_u \notin \left(a + (a-y)\frac{u}{T}, b + (b-y)\frac{u}{T}\right)\right\} \wedge T\right]. \end{aligned}$$

Since $y \geq b$, for any $u \in [0, T]$ we have $(a + (a-y)\frac{u}{T}, b + (b-y)\frac{u}{T}) \subseteq (2a - y, b)$ and then by (2)

$$\begin{aligned} \mathbb{E}_{0,T,y}[\mathcal{T}_{(a,b)}] &\leq \mathbb{E}_0\big[\inf\{u \geq 0 : W_u \notin (2a - y, b)\} \wedge T\big] \\ &= \mathbb{E}_0[\mathcal{T}_{(2a-y,b)}] \wedge T \leq \big[b(2|a| + y)\big] \wedge T. \end{aligned}$$

The case $y \le a$ follows similarly.

(ii) For $|y| \ge b-a$ we have $\mathbb{E}_{0,T,y}\left[\mathcal{T}_{(a,b)}\right] \le \mathbb{E}_{0,T,y}\left[\mathcal{T}_{(-(b-a),b-a)}\right]$. Using (49) with $h = b-a$ gives for $|y| \ge h$ that

$$\mathbb{E}_{0,T,y}\left[\mathcal{T}_{(-h,h)}\right] \;=\; h^2\, C(T,h,y),$$

and from Theorem 3.4 (i) we have that $C(T,h,y)$ converges to 1 as $h \to 0$. □

Hence for $x \in (kh,(k+1)h)$ and $k \in \{-1,0\}$ we get by Lemma 4.2 (i) that $\mathbb{E}_{0,T,-x}[\mathcal{T}_{(kh-x,(k+1)h-x)}] \le ch^2$, and for $k \notin \{-1,0\}$ we use (ii). Then Lemma 4.1 implies

$$\mathbb{E}_{0,T,-x}[B_{\mathcal{T}(x)}] = O(h^2).$$

However, this equality does not hold uniformly in x with the consequence that we can not use (44) in integrals like (51) below. For example, for the sequence $x_k := (k+0.5)h$ it holds

$$\mathbb{E}_{0,T,-x_k}[B_{\mathcal{T}(x_k)}] = \mathbb{E}_{0,T,-x_k}[B_{\mathcal{T}_{(-h/2,h/2)}}] \to -h/2, \quad k \to \infty, \tag{50}$$

which contradicts that $\mathbb{E}_{0,T,-x}[B_{\mathcal{T}(x)}] = O(h^2)$ holds uniformly in x. The limit in (50) can be easily seen from the representation $\left(B_t^{0,T,-x_k}\right)_{0\le t\le T} \overset{d}{=} \left(B_t^{0,T,0} - \frac{x_k t}{T}\right)_{0\le t\le T}$. For any path $t \mapsto B_t^{0,T,0}(\omega)$ one can find a sufficiently large x_k, such that the transformed path $t \mapsto B_t^{0,T,0}(\omega) - \frac{x_k t}{T}$ exits $(-h/2,h/2)$ first at $-h/2$.

Nevertheless, one can prove the estimates needed in [15], where q was used inside an integral over the real line. We only discuss here [15, equation (38)], because the calculations for the other cases where the function q appears are similar. For $\sigma = 1$ and denoting $\mathbb{N}_e^h := \{2kh : k \in \mathbb{Z}\}$ the last term in [15, equation (38)] can be written as

$$\begin{aligned}
&\int_{-\infty}^{\infty} \left(2h^2 - \operatorname{dist}^2(x,\mathbb{N}_e^h)\right) q(x) p_T(0,x)dx \\
&\quad = \int_{-\infty}^{\infty} \left(2h^2 - \operatorname{dist}^2(x,\mathbb{N}_e^h)\right) \operatorname{dist}(x,\mathbb{N}_o^h) h^{-1} p_T(0,x)dx \\
&\qquad + \int_{-\infty}^{\infty} \left(2h^2 - \operatorname{dist}^2(x,\mathbb{N}_e^h)\right) \left(q(x) - \operatorname{dist}(x,\mathbb{N}_o^h) h^{-1}\right) p_T(0,x)dx.
\end{aligned} \tag{51}$$

The calculation for the first integral on the r.h.s. is carried out in [15]. To justify (41) it remains to show that the other integral behaves like $O(h^3)$. Since $2h^2 - \operatorname{dist}^2(x,\mathbb{N}_e^h) \le 2h^2$ and

$$|q(x) - \operatorname{dist}(x,\mathbb{N}_o^h)h^{-1}| = \left|\mathbb{E}_{0,T,-x}\left[B_{\mathcal{T}(x)}\right]\right| h^{-1}$$

by (43) and (46), we get the desired estimate from the next corollary.

Corollary 4.3. *For $T > 0$ and $h = \sqrt{T/n}$, there exists a $C = C(T) > 0$ such that*

$$\int_{-\infty}^{\infty} \left|\mathbb{E}_{0,T,-x}\left[B_{\mathcal{T}(x)}\right]\right| p_T(0,x)dx \le Ch^2,$$

where $\mathcal{T}(x)$ is given in (46).

Proof. Since $B^{0,T,-x}_{\mathcal{T}(x)} \stackrel{d}{=} -B^{0,T,x}_{\mathcal{T}(-x)}$, it suffices to estimate the integral over $[0,\infty)$. By (47),

$$\left|\mathbb{E}_{0,T,-x}\left[B_{\mathcal{T}(x)}\right]\right| \leq \frac{\mathbb{E}_{0,T,-x}[\mathcal{T}(x)]}{T}\left(2h+|x|+3\sqrt{2T}\right).$$

If $x \in (0,h)$, then $\mathcal{T}(x) = \mathcal{T}_{(-x,h-x)}$, and estimate (48) gives

$$\mathbb{E}_{0,T,-x}[\mathcal{T}(x)] \leq 2x\left(h-x+\frac{x}{2}\right) \leq h^2.$$

For $x \geq h$ it holds

$$\mathbb{E}_{0,T,-x}[\mathcal{T}(x)] \leq C(T,h,-x)h^2$$

by Lemma 4.2. From the above estimates we get

$$\int_0^h \left|\mathbb{E}_{0,T,-x}\left[B_{\mathcal{T}(x)}\right]\right| p_T(0,x)dx \leq h^2\int_0^h \frac{2h+x+3\sqrt{2T}}{T}p_T(0,x)dx \leq C(T)h^2, \quad (52)$$

and

$$\begin{aligned}\int_h^\infty \left|\mathbb{E}_{0,T,-x}\left[B_{\mathcal{T}(x)}\right]\right| p_T(0,x)dx &\leq h^2\int_h^\infty C(T,h,-x)\frac{2h+x+3\sqrt{2T}}{T}p_T(0,x)dx \\ &\leq C(T)h^2\int_h^\infty (1+x)C(T,h,-x)p_T(0,x)dx \\ &\leq C(T)h^2, \qquad (53)\end{aligned}$$

where the constant $C(T)$ varies from line to line. The last inequality in (53) can be seen as follows. We use the representation (37) for (49) and substitute $t = h^2u$, so that

$$\begin{aligned}C(T,h,-x) = \mathbb{E}_{0,T,-x}[\mathcal{T}_{(-h,h)}]h^{-2} &= h^{-1}\int_0^T \frac{p_{T-t}(0,x)}{p_T(0,x)}\Delta(0,h,t)dt \\ &= \int_0^{T/h^2} \frac{p_{T-h^2u}(0,x)}{p_T(0,x)}\Delta(0,1,u)du.\end{aligned}$$

By Fubini's theorem it holds

$$\int_0^\infty \Delta(0,1,u)\int_h^\infty (1+x)p_{T-h^2u}(0,x)\mathbb{1}_{[0,T/h^2)}(u)dxdu \leq C(T),$$

since $u \mapsto \Delta(0,1,u)\mathbb{1}_{(0,\infty)}(u)$ is a density. The claim then follows by (52) and (53). □

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

References

[1] PH. BIANE, J. PITMAN AND M. YOR (2001). Probability laws related to the Jacobi theta and Riemann zeta functions, and Brownian excursions. *Bull. Amer. Math. Soc. (N.S.)*, **38**, 435–465.

[2] A. N. BORODIN AND P. SALMINEN (2015). *Handbook of Brownian Motion - Facts and Formulae*, Probability and its Applications. Birkhäuser Verlag, Basel, 2nd edition, 2nd printing.

[3] K. L. CHUNG AND J. B. WALSH (2005). *Markov Processes, Brownian Motion, and Time Symmetry*, Grundlehren der Mathematischen Wissenschaften, 249, 2nd edition, Springer, New York.

[4] J. L. DOOB (1949). Heuristic approach to the Kolmogorov-Smirnov theorems. *Ann. Math. Stat.*, **20**, 393–403.

[5] P. FITZSIMMONS, J. PITMAN AND M. YOR (1993). Markovian bridges: construction, Palm interpretation, and splicing. In *Seminar on Stochastic Processes*, Progress in Probability **32**, 101–134.

[6] I. S. GRADSHTEYN AND I. M. RYZHIK (2014). *Table of Integrals, Series, and Products*, 8th edition, Elsevier/Academic Press, Amsterdam.

[7] K. ITÔ AND H.P. MCKEAN (1974). *Diffusion Processes and their Sample Paths*, Die Grundlehren der Mathematischen Wissenschaften, **125**, 2nd printing, corrected, Berlin-New York.

[8] F. B. KNIGHT (1969). Brownian local times and taboo processes. *Trans. Amer. Math. Soc.*, **143**, 173–185.

[9] A. N. KOLMOGOROV (1933). Sulla determinazione empirica delle leggi di probabilita. *Giorn. Ist. Ital. Attuari*, **4**, 1–11.

[10] D. KROESE, T. TAIMRE AND Z. BOTEV (2011). *Handbook of Monte Carlo Methods*, Wiley Series in Probability and Statistics, John Wiley and Sons, New York.

[11] A. LUOTO (2017). Time-dependent weak rate of convergence for functions of generalized bounded variation, *preprint*, https://arxiv.org/abs/1609.05768.

[12] J. PITMAN AND M. YOR (1999). The law of the maximum of a Bessel bridge. *Electr. J. Probab.*, **4**, 1–15.

[13] P. SALMINEN (1997). On last exit decomposition of linear diffusions. *Studia Sci. Math. Hungar.*, **33**, 251–262.

[14] N. V. SMIRNOV (1939). On the estimation of the discrepancy between empirical curves of distribution for two independent samples. *Bul. Math. de l'Univ. de Moscou*, **2**, 3–14 (in Russian).

[15] J. B. WALSH (2003). The rate of convergence of the binomial tree scheme. *Finance Stoch.*, **7**, 337–361.

[C]

RANDOM WALK APPROXIMATION OF BSDES WITH HÖLDER CONTINUOUS TERMINAL CONDITION

C. GEISS, C. LABART, AND A. LUOTO

Random walk approximation of BSDEs with Hölder continuous terminal condition

Christel Geiss[1], Céline Labart[2], Antti Luoto[3]

Abstract

In this paper we consider the random walk approximation of the solution of a Markovian BSDE whose terminal condition is a locally Hölder continuous function of the Brownian motion. We state the rate of the L_2-convergence of the approximated solution to the true one. The proof relies in part on growth and smoothness properties of the solution u of the associated PDE. Here we improve existing results by showing some properties of the second derivative of u in space.

Keywords : Backward stochastic differential equations, numerical scheme, random walk approximation, speed of convergence

MSC codes : 65C30 60H35 60G50 65G99

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 $(\mathcal{F}_t)_{t\geq 0}$ is the augmented natural filtration. We consider the following backward stochastic differential equation (BSDE for short)

$$Y_s = g(B_T) + \int_s^T f(r, B_r, Y_r, Z_r)dr - \int_s^T Z_r dB_r, \qquad 0 \leq s \leq T, \tag{1}$$

where f is Lipschitz continuous and g is a locally α-Hölder continuous and polynomially bounded function (see (3)). In this paper we are interested in the L_2-convergence of the numerical approximation of (1) by using a random walk. First results dealing with the numerical approximation of BSDEs date back to the late 1990s. Bally (see [2]) was the first to consider this problem by introducing random discretization, namely the jump times of a Poisson process. In his PhD thesis, Chevance (see [17]) proposed the following discretization

$$y_k = \mathbb{E}(y_{k+1} + hf(y_{k+1})|\mathcal{F}_k^n), \qquad k = n-1, \ldots, 0, \quad n \in \mathbb{N}^*$$

and proved the convergence of $(Y_t^n)_t := (y_{[t/h]})_t$ to Y. At the same time, Coquet, Mackevičius and Mémin [18] proved the convergence of Y^n by using convergence of filtrations, still in the case of

[1]Department of Mathematics and Statistics, P.O.Box 35 (MaD), FI-40014 University of Jyvaskyla, Finland
`christel.geiss@jyu.fi`
[2]Univ. Grenoble Alpes, Univ. Savoie Mont Blanc, CNRS, LAMA, 73000 Chambéry, France
`celine.labart@univ-smb.fr`
[3]Department of Mathematics and Statistics, P.O.Box 35 (MaD), FI-40014 University of Jyvaskyla, Finland
`antti.k.luoto@student.jyu.fi`

a generator independent from z. The general case (f depends on z, terminal condition $\xi \in L_2$) has been studied by Briand, Delyon and Mémin (see [5]). In that paper the authors define an approximated solution (Y^n, Z^n) based on random walk and prove weak convergence to (Y, Z) using convergence of filtrations. We also refer to [27], [29], [30], [31] for other numerical methods for BSDEs which use a random walk approach. The rate of convergence of this method was left as an open problem.
Introducing instead of random walk an approach based on the dynamic programming equation, Bouchard and Touzi in [8] and Zhang in [35] managed to establish a rate of convergence. However, to be fully implementable, this algorithm requires to have a good approximation of its associated conditional expectation. For this, various methods have been developed (see [24], [19], [15]). Forward methods have also been introduced to approximate (1): a branching diffusion method (see [26]), a multilevel Picard approximation (see [34]) and Wiener chaos expansion (see [7]). Many extensions of (1) have also been considered: high order schemes (see [11], [10]), schemes for reflected BSDEs (see [3], [14]), for fully-coupled BSDEs (see [21], [9]), for quadratic BSDEs (see [13]), for BSDEs with jumps (see [23]) and for McKean-Vlasov BSDEs (see [1], [16], [12]).
From a numerical point of view, the random walk is of course not competitive with recent methods listed above. We emphasize that the aim of this paper is to give the convergence rate of the initial method based on random walk, which, to the best of our knowledge, has not been done so far.
As in [5], let us introduce the following approximation of B, based on a random walk:

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

where $h = \frac{T}{n}$ ($n \in \mathbb{N}^*$) and $(\varepsilon_i)_{i=1,2,\ldots}$ is a sequence of i.i.d. Rademacher random variables. Consider the following approximated solution (Y^n, Z^n) of (Y, Z)

$$Y_{t_k}^n = g(B_T^n) + h \sum_{m=k}^{n-1} f(t_{m+1}, B_{t_m}^n, Y_{t_m}^n, Z_{t_m}^n) - \sqrt{h} \sum_{m=k}^{n-1} Z_{t_m}^n \varepsilon_{m+1}, \quad 0 \le k \le n-1. \tag{2}$$

The main result of our paper gives the rate of convergence in L_2-norm of $Y_v^n - Y_v$ and $Z_v^n - Z_v$ for each $v \in [0, T)$ (see Theorem 3.1). Basically, we get that the L_2-norm of the error on Y is of order $h^{\frac{\alpha}{4}}$ and the L_2-norm of the error on Z is of order $\frac{h^{\frac{\alpha}{4}}}{\sqrt{T-v}}$. The proof of this result is based on several ingredients. In particular, we need some estimates on the bound of the first and second derivatives of the solution of the PDE associated to the BSDE (1). We establish these bounds in the case of a forward backward SDE (FBSDE for short) whose terminal condition satisfies the Hölder continuity condition (3). This result extends Zhang [36, Theorem 3.2].

The rest of the paper is organized as follows. Section 2 introduces notations, assumptions and the representation for Z and Z^n based on the Malliavin weights. Section 3 states the rate of convergence of the error on Y and Z in L_2-norm, which is the main result of the paper. Section 4 presents numerical simulations and Section 5 recalls some properties of Malliavin weights, of the regularity of solutions to FBSDEs with a locally Hölder continuous terminal condition function and states some properties of the solutions to the PDEs associated to these FBSDEs.

2 Preliminaries

This section is dedicated to notations, assumptions and the representation of Z and Z^n using the Malliavin weights.

Notation:

- $\mathcal{G}_k := \sigma(\varepsilon_i : 1 \le i \le k)$ and $\mathcal{G}_0 = \{\emptyset, \Omega\}$. The associated discrete-time random walk $(B^n_{t_k})_{k=0}^n$ is $(\mathcal{G}_k)_{k=0}^n$-adapted.

- $\|\cdot\|_p := \|\cdot\|_{L^p(\mathbb{P})}$ for $p \ge 1$ and for $p = 2$ simply $\|\cdot\|$. constant.

Assumption 2.1.

- *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*

$$\forall (x,y) \in \mathbb{R}^2, \quad |g(x) - g(y)| \le C_g(1 + |x|^{p_0} + |y|^{p_0})|x - y|^\alpha. \tag{3}$$

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

$$|f(t,x,y,z) - f(t',x',y',z')| \le L_f(\sqrt{t-t'} + |x-x'| + |y-y'| + |z-z'|). \tag{4}$$

Notice that (3) implies

$$|g(x)| \le K(1 + |x|^{p_0+1}) =: \Psi(x). \tag{5}$$

In the rest of the paper, the study of the error $(Y^n - Y, Z^n - Z)$ will either rely on (2) or on its integral version:

$$Y^n_s = g(B^n_T) + \int_{(s,T]} f(r, B^n_{r-}, Y^n_{r-}, Z^n_{r-}) d[B^n, B^n]_r - \int_{(s,T]} Z^n_{r-} dB^n_r, \quad 0 \le s \le T, \tag{6}$$

where the backward equation (6) arises from (2) by setting $Y^n_r := Y^n_{t_m}$ and $Z^n_r := Z^n_{t_m}$ for $r \in [t_m, t_{m+1})$. For n large enough, (6) has a unique solution (Y^n, Z^n), and $(Y^n_{t_m}, Z^n_{t_m})_{m=0}^{n-1}$ is adapted to the filtration $(\mathcal{G}_m)_{m=0}^{n-1}$. Let us now introduce the Malliavin representations for Z and Z^n. They are the cornerstone of our study of the error on Z.

2.1 Representations for Z and Z^n

We will use the representation (see Ma and Zhang [28, Theorem 4.2])

$$Z_t = \mathbb{E}_t\left(g(B_T)N^t_T + \int_t^T f(s, B_s, Y_s, Z_s)N^t_s ds\right), \quad 0 \le t \le T, \tag{7}$$

where $\mathbb{E}_t[\cdot] = \mathbb{E}[\cdot|\mathcal{F}_t]$, and for all $s \in (t,T]$ we have

$$N^t_s := \frac{B_s - B_t}{s - t}.$$

Lemma 2.2. *Suppose that Assumption 2.1 holds. Then the process Z^n given by (6) has the representation*

$$Z^n_{t_k} = \mathbb{E}_k\left(g(B^n_T)\frac{B^n_{t_n} - B^n_{t_k}}{t_n - t_k}\right) + \mathbb{E}_k\left(h\sum_{m=k+1}^{n-1} f(t_{m+1}, B^n_{t_m}, Y^n_{t_m}, Z^n_{t_m})\frac{B^n_{t_m} - B^n_{t_k}}{t_m - t_k}\right) \tag{8}$$

for $k = 0, 1, \dots, n-1$, where $\mathbb{E}_k[\,\cdot\,] := \mathbb{E}[\,\cdot\,|\mathcal{G}_k]$.

Proof. We multiply equation (2) by ε_{k+1} and take the conditional expectation with respect to $\mathcal{G}_k$. Since $(Y^n_{t_k}, Z^n_{t_k})$ is $\mathcal{G}_k$-measurable, it holds for $0 \le k \le n-1$ that

$$\begin{aligned}
&\mathbb{E}_k\left(Y^n_{t_k}\varepsilon_{k+1}\right)\\
&= \mathbb{E}_k\left(g(B^n_T)\varepsilon_{k+1}\right) + h\mathbb{E}_k\left(\sum_{m=k}^{n-1} f(t_{m+1}, B^n_{t_m}, Y^n_{t_m}, Z^n_{t_m})\varepsilon_{k+1}\right) - \sqrt{h}\mathbb{E}_k\left(\sum_{m=k}^{n-1} Z^n_{t_m}\varepsilon_{m+1}\varepsilon_{k+1}\right)\\
&= \sqrt{h}\mathbb{E}_k\left(g(B^n_T)\frac{B^n_{t_n} - B^n_{t_k}}{t_n - t_k}\right) + h^{3/2}\sum_{m=k+1}^{n-1} \mathbb{E}_k\left(f(t_{m+1}, B^n_{t_m}, Y^n_{t_m}, Z^n_{t_m})\frac{B^n_{t_m} - B^n_{t_k}}{t_m - t_k}\right) - \sqrt{h}Z^n_{t_k}, \quad (9)
\end{aligned}$$

where the l.h.s. is equal to zero. Indeed, for $m \ge k+1$, we have

$$\mathbb{E}_k(Z^n_{t_m}\varepsilon_{m+1}\varepsilon_{k+1}) = \mathbb{E}_k(Z^n_{t_m}\varepsilon_{k+1}\mathbb{E}_m\varepsilon_{m+1}) = 0,$$

and for $m = k$ it holds $\mathbb{E}_k(Z^n_{t_k}\varepsilon^2_{k+1}) = Z^n_{t_k}$. Moreover, the fact that $B^n_T = \sqrt{h}\sum_{m=0}^{n-1}\varepsilon_{m+1}$, where $(\varepsilon_m)_{m=1,2\ldots}$ are i.i.d., yields

$$\mathbb{E}_k\left(g(B^n_T)\varepsilon_{k+1}\right) = \mathbb{E}_k\left(g(B^n_T)\sum_{m=k}^{n-1}\frac{\varepsilon_{k+1}}{n-k}\right) = \mathbb{E}_k\left(g(B^n_T)\sum_{m=k}^{n-1}\frac{\varepsilon_{m+1}}{n-k}\right) = \sqrt{h}\mathbb{E}_k\left(g(B^n_T)\frac{B^n_{t_n} - B^n_{t_k}}{t_n - t_k}\right).$$

Similarly, for $m \ge k+1$, we get (using [5, Proposition 5.1], where it is stated that both $Y^n_{t_m}$ and $Z^n_{t_m}$ can be represented as functions of t_m and $B^n_{t_m}$)

$$\mathbb{E}_k\left(f(t_{m+1}, B^n_{t_m}, Y^n_{t_m}, Z^n_{t_m})\varepsilon_{k+1}\right) = \sqrt{h}\mathbb{E}_k\left(f(t_{m+1}, B^n_{t_m}, Y^n_{t_m}, Z^n_{t_m})\frac{B^n_{t_m} - B^n_{t_k}}{t_m - t_k}\right).$$

It remains to divide (9) by $\sqrt{h}$ and rearrange. □

3 Main result

This section is devoted to the main result of the paper: the rate of the L_2-convergence of (Y^n, Z^n) to (Y, Z). The proof will rely on the fact that the random walk B^n can be constructed from the Brownian motion B by Skorohod embedding. Let $\tau_0 := 0$ and define

$$\tau_k := \inf\{t > \tau_{k-1} : |B_t - B_{\tau_{k-1}}| = \sqrt{h}\}, \quad k \ge 1.$$

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 $\sqrt{h}\varepsilon_k \stackrel{d}{=} B_{\tau_k} - B_{\tau_{k-1}}$. We will use this random walk for our approximation, i.e. we will require

$$B^n_t = \sum_{k=1}^{[t/h]}(B_{\tau_k} - B_{\tau_{k-1}}), \qquad 0 \le t \le T. \quad (10)$$

Properties satisfied by τ_k and B_{τ_k} are stated in Lemma A.1. We will denote by $\mathbb{E}_{\tau_k}$ the conditional expectation w.r.t. $\mathcal{F}_{\tau_k}$.

Theorem 3.1. *Let Assumption 2.1 hold. If B^n satisfies (10) then we have (for sufficiently large n) that*

$$\mathbb{E}|Y_v - Y_v^n|^2 \le C_0 h^{\frac{\alpha}{2}} \quad \textit{for} \quad v \in [0,T),$$
$$\mathbb{E}|Z_v - Z_v^n|^2 \le C_0 \frac{h^{\frac{\alpha}{2}}}{T-t_k} + C_1 \frac{h^{\frac{\alpha}{2}}}{(T-v)^{1-\frac{\alpha}{2}}} \mathbf{1}_{v \neq t_k} \quad \textit{for} \quad v \in [t_k, t_{k+1}),\ k = 0, ..., n-1,$$

where we have the dependencies $C_0 = C(T, p_0, L_f, C_g, C^y_{5.3}, C^z_{5.3}, K_f, c_{5.4}, \alpha)$, $C_1 = C(T, p_0, C^z_{5.3}, \alpha)$ and $K_f := \sup_{0 \le t \le T} |f(t,0,0,0)|$.

Remark 3.2. *Theorem 3.1 implies that*

$$\sup_{v \in [0,T)} \mathbb{E}|Y_v - Y_v^n|^2 \le C_0 h^{\frac{\alpha}{2}} \qquad \textit{and} \qquad \mathbb{E}\int_0^T |Z_v - Z_v^n|^2 dv \le C(C_0, C_1, \beta)\, h^{\beta} \qquad \textit{for } \beta \in (0, \tfrac{\alpha}{2}).$$

Proof of Theorem 3.1. Let $u : [0,T) \times \mathbb{R} \to \mathbb{R}$ be the solution of the PDE associated to (1). Since by Theorem 5.4

$$Y_s = u(s, B_s), \quad Z_s = u_x(s, B_s), \quad a.s.$$

we introduce

$$F(s,x) := f(s, x, u(s,x), u_x(s,x)),$$

so that $F(s, B_s) = f(s, B_s, Y_s, Z_s)$. We first give some properties satisfied by F.

Lemma 3.3. *If Assumption 2.1 holds then F is a Lipschitz continuous and polynomially bounded function in x:*

$$|F(t,x_1) - F(t,x_2)| \le C(T, L_f, c^{2,3}_{5.4})(1 + |x_1|^{p_0+1} + |x_2|^{p_0+1}) \frac{|x_1 - x_2|}{(T-t)^{1-\frac{\alpha}{2}}},$$
$$|F(t,x)| \le C(T, L_f, c^{1,2}_{5.4}, K_f) \frac{\Psi(x)}{(T-t)^{\frac{1-\alpha}{2}}},$$

where $\Psi(x)$ is given in (5).

Proof of Lemma 3.3. Thanks to the mean value theorem and Theorem 5.4-(ii-c) and (iii-b) we have for $x_1, x_2 \in \mathbb{R}$ that there exist $\xi_1, \xi_2 \in [\min\{x_1, x_2\}, \max\{x_1, x_2\}]$ such that

$$\begin{aligned} |F(t,x_1) - F(t,x_2)| &= |f(t, x_1, u(t,x_1), u_x(t,x_1)) - f(t, x_2, u(t,x_2), u_x(t,x_2))| \\ &\le L_f(|x_1 - x_2| + |u(t,x_1) - u(t,x_2)| + |u_x(t,x_1) - u_x(t,x_2)|) \\ &\le L_f \left(1 + \frac{c^2_{5.4}\Psi(\xi_1)}{(T-t)^{\frac{1-\alpha}{2}}} + \frac{c^3_{5.4}\Psi(\xi_2)}{(T-t)^{1-\frac{\alpha}{2}}}\right) |x_1 - x_2| \\ &\le C(T, L_f, c^{2,3}_{5.4})(1 + |x_1|^{p_0+1} + |x_2|^{p_0+1}) \frac{|x_1 - x_2|}{(T-t)^{1-\frac{\alpha}{2}}}. \end{aligned}$$

The second inequality can be shown similarly. □

For the estimate of $\mathbb{E}|Y_{t_k} - Y^n_{t_k}|^2$ we will use (1) and (2): Since $Y^n_{t_k}$ is $\mathcal{F}_{\tau_k}$-measurable we have

$$\begin{aligned}\|Y_{t_k} - Y^n_{t_k}\| &\leq \|\mathbb{E}_{t_k} g(B_T) - \mathbb{E}_{\tau_k} g(B^n_T)\| \\ &\quad + \left\| \mathbb{E}_{t_k} \int_{t_k}^T f(s, B_s, Y_s, Z_s) ds - h\mathbb{E}_{\tau_k} \sum_{m=k}^{n-1} f(t_{m+1}, B^n_{t_m}, Y^n_{t_m}, Z^n_{t_m}) \right\|. \qquad (11)\end{aligned}$$

We frequently express conditional expectations with the help of an independent copy of B denoted by $\tilde{B}$, for example $\mathbb{E}_t g(B_T) = \tilde{\mathbb{E}} g(B_t + \tilde{B}_{T-t})$.
By (3) and Lemma A.1,

$$\begin{aligned}\|\mathbb{E}_{t_k} g(B_T) - \mathbb{E}_{\tau_k} g(B^n_T)\|^2 &= \mathbb{E}|\tilde{\mathbb{E}} g(B_{t_k} + \tilde{B}_{T-t_k}) - \tilde{\mathbb{E}} g(B_{\tau_k} + \tilde{B}_{\tilde{\tau}_{n-k}})|^2 \\ &\leq (\mathbb{E}\tilde{\mathbb{E}}(\Psi_1)^4)^{\frac{1}{2}} (\mathbb{E}\tilde{\mathbb{E}}|B_{t_k} - B_{\tau_k} + \tilde{B}_{T-t_k} - \tilde{B}_{\tilde{\tau}_{n-k}}|^{4\alpha})^{\frac{1}{2}} \\ &\leq C(C_g, T, p_0)((\mathbb{E}|B_{t_k} - B_{\tau_k}|^{4\alpha})^{\frac{1}{2}} + (\mathbb{E}|B_{T-t_k} - B_{\tau_{n-k}}|^{4\alpha})^{\frac{1}{2}}) \\ &\leq C(C_g, T, p_0) h^{\frac{\alpha}{2}}, \qquad (12)\end{aligned}$$

where $\Psi_1 := C_g(1 + |B_{t_k} + \tilde{B}_{T-t_k}|^{p_0} + |B_{\tau_k} + \tilde{B}_{\tilde{\tau}_{n-k}}|^{p_0})$. To estimate the other term in (11) we consider the decomposition

$$\begin{aligned}&\mathbb{E}_{t_k} f(s, B_s, Y_s, Z_s) - \mathbb{E}_{\tau_k} f(t_{m+1}, B^n_{t_m}, Y^n_{t_m}, Z^n_{t_m}) \\ &= (\mathbb{E}_{t_k} f(s, B_s, Y_s, Z_s) - \mathbb{E}_{t_k} f(t_m, B_{t_m}, Y_{t_m}, Z_{t_m})) + (\mathbb{E}_{t_k} F(t_m, B_{t_m}) - \mathbb{E}_{\tau_k} F(t_m, B_{\tau_m})) \\ &\quad + (\mathbb{E}_{\tau_k} F(t_m, B_{\tau_m}) - \mathbb{E}_{\tau_k} F(t_m, B_{t_m})) + (\mathbb{E}_{\tau_k} f(t_m, B_{t_m}, Y_{t_m}, Z_{t_m}) - \mathbb{E}_{\tau_k} f(t_{m+1}, B^n_{t_m}, Y^n_{t_m}, Z^n_{t_m})) \\ &=: D_1(s, m) + D_2(m) + \ldots + D_4(m)\end{aligned}$$

so that

$$\begin{aligned}&\left\| \mathbb{E}_{t_k} \int_{t_k}^T f(s, B_s, Y_s, Z_s) ds - h\mathbb{E}_{\tau_k} \sum_{m=k}^{n-1} f(t_{m+1}, B^n_{t_m}, Y^n_{t_m}, Z^n_{t_m}) \right\| \\ &\leq \sum_{m=k}^{n-1} \left(\left\| \int_{t_m}^{t_{m+1}} D_1(s, m) ds \right\| + h \sum_{i=2}^{4} \|D_i(m)\| \right).\end{aligned}$$

For D_1 we have by Theorem 5.3 that

$$\begin{aligned}\|D_1(s, m)\| &\leq L_f(\sqrt{s - t_m} + \|B_s - B_{t_m}\| + \|Y_s - Y_{t_m}\| + \|Z_s - Z_{t_m}\|) \\ &\leq C(T, L_f, C^y_{5.3}, C^z_{5.3}, p_0)\, (T-s)^{\frac{\alpha-2}{2}}\, h^{\frac{1}{2}}, \qquad (13)\end{aligned}$$

where the last inequality follows from $\|B_s - B_{t_m}\| = \sqrt{s - t_m} \leq h^{\frac{1}{2}}$ for $s \in [t_m, t_{m+1}]$ and

$$\begin{aligned}\|Y_s - Y_{t_m}\| + \|Z_s - Z_{t_m}\| &\leq (\mathbb{E}\Psi(B_{t_m})^2)^{\frac{1}{2}} \left(C^y_{5.3} \left(\int_{t_m}^s (T-r)^{\alpha-1} dr \right)^{\frac{1}{2}} + C^z_{5.3} \left(\int_{t_m}^s (T-r)^{\alpha-2} dr \right)^{\frac{1}{2}} \right) \\ &\leq C(T, C^y_{5.3}, C^z_{5.3}, p_0)\sqrt{s - t_m}((T-s)^{\frac{\alpha-1}{2}} + (T-s)^{\frac{\alpha-2}{2}}).\end{aligned}$$

We bound D_2 using Lemma 3.3 and Lemma A.1. Similar to (12) we conclude (setting $\Psi_2 := 1 + |B_{t_k} + \tilde{B}_{t_{m-k}}|^{p_0+1} + |B_{\tau_k} + \tilde{B}_{\tilde{\tau}_{m-k}}|^{p_0+1}$) that

$$\|D_2(m)\| = \left(\mathbb{E}\, |\mathbb{E}_{t_k} F(t_m, B_{t_m}) - \mathbb{E}_{\tau_k} F(t_m, B_{\tau_m})|^2 \right)^{\frac{1}{2}}$$

$$\leq C(T, L_f, c^{2,3}_{5.4})(\mathbb{E}\tilde{\mathbb{E}}\Psi_2^4)^{\frac{1}{4}} \frac{1}{(T-t_m)^{1-\frac{\alpha}{2}}} (t_k h + t_{m-k}h)^{\frac{1}{4}}$$
$$\leq C(T, p_0, L_f, c^{2,3}_{5.4}) \frac{1}{(T-t_m)^{1-\frac{\alpha}{2}}} h^{\frac{1}{4}}.$$

For D_3 we apply again Lemma 3.3 and Lemma A.1,

$$\begin{aligned}\|D_3(m)\| \leq \|F(t_m, B_{t_m}) - F(t_m, B_{\tau_m})\| &\leq C(T, L_f, c^{2,3}_{5.4}) \frac{1}{(T-t_m)^{1-\frac{\alpha}{2}}} \|\Psi_3 \, |B_{t_m} - B_{\tau_m}|\| \\ &\leq C(T, p_0, L_f, c^{2,3}_{5.4}) \frac{1}{(T-t_m)^{1-\frac{\alpha}{2}}} h^{\frac{1}{4}},\end{aligned}$$

where $\Psi_3 := 1 + |B_{t_m}|^{p_0+1} + |B_{\tau_m}|^{p_0+1}$. For the last term D_4 we get

$$\|D_4(m)\| \leq L_f(h^{\frac{1}{2}} + \|B_{t_m} - B^n_{t_m}\| + \|Y_{t_m} - Y^n_{t_m}\| + \|Z_{t_m} - Z^n_{t_m}\|).$$

Finally, using the estimates for the terms $D_1(s,m), D_2(m), ..., D_4(m)$ we arrive at

$$\begin{aligned}\|Y_{t_k} - Y^n_{t_k}\| &\leq C(C_g, T, p_0) h^{\frac{\alpha}{4}} + C(T, L_f, C^y_{5.3}, C^z_{5.3}, p_0)\, h^{\frac{1}{2}} \int_{t_k}^T (T-s)^{\frac{\alpha-2}{2}} ds \\ &\quad + C(T, p_0, L_f, c^{2,3}_{5.4}) h^{\frac{1}{4}} \sum_{m=k}^{n-1} \frac{h}{(T-t_m)^{1-\frac{\alpha}{2}}} + hL_f \sum_{m=k}^{n-1} (\|Y_{t_m} - Y^n_{t_m}\| + \|Z_{t_m} - Z^n_{t_m}\|) \\ &\leq C(C_g, T, p_0, L_f, c^{2,3}_{5.4}, C^y_{5.3}, C^z_{5.3}) h^{\frac{\alpha}{4}} + hL_f \sum_{m=k}^{n-1} (\|Y_{t_m} - Y^n_{t_m}\| + \|Z_{t_m} - Z^n_{t_m}\|). \quad (14)\end{aligned}$$

For $\|Z_{t_k} - Z^n_{t_k}\|$ we exploit the representations (7) and (8) and estimate

$$\begin{aligned}\|Z_{t_k} - Z^n_{t_k}\| &\leq \frac{1}{T-t_k} \|\mathbb{E}_{t_k} g(B_T)(B_T - B_{t_k}) - \mathbb{E}_{\tau_k} g(B_{\tau_n})(B_{\tau_n} - B_{\tau_k})\| \\ &\quad + \Big\| \mathbb{E}_{t_k} \left(\int_{t_{k+1}}^T f(s, B_s, Y_s, Z_s) \frac{B_s - B_{t_k}}{s - t_k} ds \right) \\ &\qquad - \mathbb{E}_{\tau_k} \left(h \sum_{m=k+1}^{n-1} f(t_{m+1}, B^n_{t_m}, Y^n_{t_m}, Z^n_{t_m}) \frac{B^n_{t_m} - B^n_{t_k}}{t_m - t_k} \right) \Big\| \\ &\quad + \Big\| \mathbb{E}_{t_k} \int_{t_k}^{t_{k+1}} f(s, B_s, Y_s, Z_s) \frac{B_s - B_{t_k}}{s - t_k} ds \Big\|.\end{aligned}$$

Then, similar to (12), we have for the terminal condition by Lemma A.1 that

$$\begin{aligned}&\|\mathbb{E}_{t_k}[g(B_T)(B_T - B_{t_k})] - \mathbb{E}_{\tau_k}[g(B_{\tau_n})(B_{\tau_n} - B_{\tau_k})]\| \\ &= \|\tilde{\mathbb{E}}[g(B_{t_k} + \tilde{B}_{T-t_k}) - g(B_{t_k})](\tilde{B}_{T-t_k} - \tilde{B}_{\tilde{\tau}_{n-k}}) + \tilde{\mathbb{E}}[g(B_{t_k} + \tilde{B}_{T-t_k}) - g(B_{\tau_k} + \tilde{B}_{\tilde{\tau}_{n-k}})]\tilde{B}_{\tilde{\tau}_{n-k}}\| \\ &\leq C(C_g, T, p_0) h^{\frac{1}{4}} (T-t_k)^{\frac{\alpha}{2}+\frac{1}{4}} + C(C_g, T, p_0) h^{\frac{\alpha}{4}} (T-t_k)^{\frac{1}{2}} \leq C(C_g, T, p_0) h^{\frac{\alpha}{4}} (T-t_k)^{\frac{1}{2}}.\end{aligned}$$

Here we have used that $\tilde{\mathbb{E}}[g(B_{t_k})(\tilde{B}_{T-t_k} - \tilde{B}_{\tilde{\tau}_{n-k}})] = 0$. The term $\tilde{\mathbb{E}}[g(B_{t_k} + \tilde{B}_{T-t_k}) - g(B_{t_k})](\tilde{B}_{T-t_k} - \tilde{B}_{\tilde{\tau}_{n-k}})$ provides us with the factor $(T-t_k)^{\frac{\alpha}{2}}((T-t_k)h)^{\frac{1}{4}}$. For the next term of the estimate of $\|Z_{t_k} - Z^n_{t_k}\|$ we use for $s \in [t_m, t_{m+1})$, where $m \geq k+1$, the decomposition

$$\frac{\mathbb{E}_{t_k} f(s, B_s, Y_s, Z_s)(B_s - B_{t_k})}{s - t_k} - \frac{\mathbb{E}_{\tau_k} f(t_{m+1}, B^n_{t_m}, Y^n_{t_m}, Z^n_{t_m})(B^n_{t_m} - B^n_{t_k})}{t_m - t_k}$$

$$
\begin{aligned}
= \quad & \frac{\mathbb{E}_{t_k} f(s, B_s, Y_s, Z_s)(B_s - B_{t_k})}{s - t_k} - \frac{\mathbb{E}_{t_k} f(t_m, B_{t_m}, Y_{t_m}, Z_{t_m})(B_{t_m} - B_{t_k})}{t_m - t_k} \\
& + \frac{\mathbb{E}_{t_k} F(t_m, B_{t_m})(B_{t_m} - B_{t_k})}{t_m - t_k} - \frac{\mathbb{E}_{\tau_k} F(t_m, B_{\tau_m})(B_{\tau_m} - B_{\tau_k})}{t_m - t_k} \\
& + \mathbb{E}_{\tau_k} \left[[F(t_m, B_{\tau_m}) - F(t_m, B_{t_m})] \frac{B_{\tau_m} - B_{\tau_k}}{t_m - t_k} \right] \\
& + \mathbb{E}_{\tau_k} \left[[f(t_m, B_{t_m}, Y_{t_m}, Z_{t_m}) - f(t_{m+1}, B^n_{t_m}, Y^n_{t_m}, Z^n_{t_m})] \frac{B^n_{t_m} - B^n_{t_k}}{t_m - t_k} \right] \\
=: \quad & T_1(s, m) + T_2(m) + \ldots + T_4(m).
\end{aligned}
$$

Then by the conditional Hölder inequality and by (13) as well as by Lemma 3.3 we have

$$
\begin{aligned}
\|T_1(s,m)\| \quad \le \quad & \|D_1(s,m)\| \frac{\|B_s - B_{t_k}\|}{s - t_k} + \|f(t_m, B_{t_m}, Y_{t_m}, Z_{t_m})\| \left\| \frac{B_s - B_{t_k}}{s - t_k} - \frac{B_{t_m} - B_{t_k}}{t_m - t_k} \right\| \\
\le \quad & C(T, L_f, C^y_{5.3}, C^z_{5.3}, p_0)\,(T-s)^{\frac{\alpha-2}{2}} \, \frac{h^{\frac{1}{2}}}{\sqrt{s - t_k}} \\
& + C(T, L_f, c^{1,2}_{5.4}, K_f) \frac{(\mathbb{E}\Psi(B_{t_m})^2)^{\frac{1}{2}}}{(T - t_m)^{\frac{1-\alpha}{2}}} \\
& \times \left(\frac{\|B_s - B_{t_m}\|}{s - t_k} + \|B_{t_m} - B_{t_k}\| \left| \frac{1}{s - t_k} - \frac{1}{t_m - t_k} \right| \right) \\
\le \quad & C(T, L_f, K_f, C^y_{5.3}, C^z_{5.3}, c^{1,2}_{5.4}, p_0)(T-s)^{\frac{\alpha-2}{2}} \, \frac{h^{\frac{1}{4}}}{(s - t_k)^{\frac{3}{4}}}.
\end{aligned}
$$

Indeed,

$$
\frac{\|B_s - B_{t_m}\|}{s - t_k} + \|B_{t_m} - B_{t_k}\| \left| \frac{1}{s - t_k} - \frac{1}{t_m - t_k} \right| \quad \le \quad \frac{\sqrt{s - t_m}}{s - t_k} + \frac{\sqrt{t_m - t_k}(s - t_m)}{(s - t_k)(t_m - t_k)} \le C \frac{h^{\frac{1}{4}}}{(s - t_k)^{\frac{3}{4}}},
$$

where the last inequality follows from $s - t_m \le t_{m+1} - t_m = h$ and $h \le t_m - t_k \le s - t_k$. We estimate T_2 with the help of Lemma 3.3 and Lemma A.1 as follows :

$$
\begin{aligned}
\|T_2(m)\| \quad \le \quad & \|\widehat{D}_2(m)\| \frac{\|B_{t_m} - B_{t_k}\|}{t_m - t_k} + \|F(t_m, B_{\tau_m})\| \frac{\|B_{t_{m-k}} - B_{\tau_{m-k}}\|}{t_m - t_k} \\
\le \quad & C(T, p_0, L_f, K_f, c_{5.4}) \frac{1}{(T - t_m)^{1 - \frac{\alpha}{2}}} \, \frac{h^{\frac{1}{4}}}{(t_m - t_k)^{\frac{3}{4}}}.
\end{aligned}
$$

Here $\widehat{D}_2(m) := (\tilde{\mathbb{E}}|F(t_m, B_{t_k} + \tilde{B}_{t_{m-k}}) - F(t_m, B_{\tau_k} + \tilde{B}_{\tilde{\tau}_{m-k}})|^2)^{\frac{1}{2}}$ which can be estimated as $D_2(m)$. For T_3 the conditional Hölder inequality and Lemma A.1 yield

$$
\|T_3(m)\| \quad \le \quad \|\widehat{D}_3(m)\| \left\| \frac{B_{\tau_m} - B_{\tau_k}}{t_m - t_k} \right\| \le C(T, p_0, L_f, c^{2,3}_{5.4}) \frac{1}{(T - t_m)^{1 - \frac{\alpha}{2}}} \, \frac{h^{\frac{1}{4}}}{(t_m - t_k)^{\frac{1}{2}}},
$$

where $\widehat{D}_3(m) := F(t_m, B_{\tau_m}) - F(t_m, B_{t_m})$ is estimated as $D_3(m)$. Finally,

$$
\|T_4(m)\| \quad \le \quad L_f(h^{\frac{1}{2}} + \|B_{t_m} - B^n_{t_m}\| + \|Y_{t_m} - Y^n_{t_m}\| + \|Z_{t_m} - Z^n_{t_m}\|) \frac{1}{\sqrt{t_m - t_k}}.
$$

For the estimate of $\left\|\mathbb{E}_{t_k}\int_{t_k}^{t_{k+1}} f(s,B_s,Y_s,Z_s)\frac{B_s-B_{t_k}}{s-t_k}ds\right\|$ one notices that by the conditional Hölder inequality,

$$\begin{aligned}\|\mathbb{E}_{t_k} f(s,B_s,Y_s,Z_s)\tfrac{B_s-B_{t_k}}{s-t_k}\| &= \|\mathbb{E}_{t_k}[(f(s,B_s,Y_s,Z_s)-f(s,B_{t_k},Y_{t_k},Z_{t_k}))\tfrac{B_s-B_{t_k}}{s-t_k}]\| \\ &\le \|f(s,B_s,Y_s,Z_s)-f(s,B_{t_k},Y_{t_k},Z_{t_k})\|\frac{1}{\sqrt{s-t_k}} \\ &\le C(T,L_f,C^y_{5.3},C^z_{5.3},p_0)\,(T-s)^{\frac{\alpha-2}{2}}\,\frac{h^{\frac12}}{\sqrt{s-t_k}},\end{aligned}$$

where the last inequality follows in the same way as in (13). Consequently, we have

$$\begin{aligned}\|Z_{t_k}-Z^n_{t_k}\| \le\ & \frac{C(C_g,T,p_0)}{(T-t_k)^{\frac12}}\,h^{\frac{\alpha}{4}} + C(T,L_f,K_f,C^y_{5.3},C^z_{5.3},c^{1,2}_{5.4},p_0)\int_{t_k}^T \frac{ds}{(T-s)^{1-\frac{\alpha}{2}}(s-t_k)^{\frac34}}\,h^{\frac14} \\ & + C(T,p_0,L_f,K_f,c_{5.4})\,h\sum_{m=k+1}^{n-1}\frac{1}{(T-t_m)^{1-\frac{\alpha}{2}}}\frac{h^{\frac14}}{(t_m-t_k)^{\frac34}} \\ & + L_f h\sum_{m=k+1}^{n-1}(\|B_{t_m}-B^n_{t_m}\|+\|Y_{t_m}-Y^n_{t_m}\|+\|Z_{t_m}-Z^n_{t_m}\|)\frac{1}{\sqrt{t_{m-k}}}.\end{aligned}$$

Lemma A.2 enables to bound the second and third term of the r.h.s. by $C\frac{h^{\frac14}}{(T-t_k)^{\frac34-\frac{\alpha}{2}}}B(\frac{\alpha}{2},\frac14)$, which is bounded by $C\frac{h^{\frac{\alpha}{4}}}{(T-t_k)^{\frac12-\frac{\alpha}{4}}}$. Thus we get

$$\|Z_{t_k}-Z^n_{t_k}\| \le \frac{C_0\,h^{\frac{\alpha}{4}}}{(T-t_k)^{\frac12}} + L_f h\sum_{m=k+1}^{n-1}(\|Y_{t_m}-Y^n_{t_m}\|+\|Z_{t_m}-Z^n_{t_m}\|)\frac{1}{\sqrt{t_{m-k}}}.$$

Then we use (14) and the above estimate to get

$$\|Y_{t_k}-Y^n_{t_k}\|+\|Z_{t_k}-Z^n_{t_k}\| \le \frac{C_0\,h^{\frac{\alpha}{4}}}{(T-t_k)^{\frac12}} + C(L_f)\,h\sum_{m=k+1}^{n-1}(\|Y_{t_m}-Y^n_{t_m}\|+\|Z_{t_m}-Z^n_{t_m}\|)\frac{1}{\sqrt{t_{m-k}}}.$$

If this inequality is iterated, one gets a shape where the Gronwall lemma applies. Indeed, setting $a_m := (\|Y_{t_m}-Y^n_{t_m}\|+\|Z_{t_m}-Z^n_{t_m}\|)$ one has to consider the double sum

$$\sum_{m=k+1}^{n-1}\left(\sum_{l=m+1}^{n-1}a_l\frac{h}{\sqrt{t_{l-m}}}\right)\frac{h}{\sqrt{t_{m-k}}} = h\sum_{l=k+1}^{n-1}\left(\sum_{m=k+1}^{l-1}\frac{h}{\sqrt{t_{m-k}}\sqrt{t_{l-m}}}\right)a_l \le Ch\sum_{l=k+1}^{n-1}a_l.$$

Consequently,

$$\|Y_{t_k}-Y^n_{t_k}\|+\|Z_{t_k}-Z^n_{t_k}\| \le \frac{C_0\,h^{\frac{\alpha}{4}}}{(T-t_k)^{\frac12}}$$

which gives the bound on the error on Z. Moreover, (14) yields

$$\|Y_{t_k}-Y^n_{t_k}\| \le C_0\,h^{\frac{\alpha}{4}}.$$

If $v\in[t_k,t_{k+1})$, we have by Theorem 5.3 that

$$\|Y_v-Y^n_v\|\le\|Y_v-Y_{t_k}\|+\|Y_{t_k}-Y^n_{t_k}\|\le C(C^y_{5.3},T,p_0)\left(\int_{t_k}^v(T-r)^{\alpha-1}dr\right)^{\frac12}+\|Y_{t_k}-Y^n_{t_k}\|,$$

$$\|Z_v - Z_v^n\| \leq \|Z_v - Z_{t_k}\| + \|Z_{t_k} - Z_{t_k}^n\| \leq C(C_{5.3}^z, T, p_0)\left(\int_{t_k}^v (T-r)^{\alpha-2} dr\right)^{\frac{1}{2}} + \|Z_{t_k} - Z_{t_k}^n\|,$$

where

$$\int_{t_k}^v (T-r)^{\alpha-1} dr \leq \frac{1}{\alpha}(v-t_k)^\alpha \leq \frac{1}{\alpha} h^\alpha$$

and

$$\int_{t_k}^v (T-r)^{\alpha-2} dr \leq \frac{1}{(T-v)^{1-\frac{\alpha}{2}}} \int_{t_k}^v (T-r)^{\frac{\alpha}{2}-1} dr \leq \frac{1}{(T-v)^{1-\frac{\alpha}{2}}} \frac{2}{\alpha}(v-t_k)^{\frac{\alpha}{2}} \leq \frac{2}{\alpha} \frac{h^{\frac{\alpha}{2}}}{(T-v)^{1-\frac{\alpha}{2}}}.$$

□

4 Numerical simulations

This section deals with the algorithm used to compute $(Y_{t_k}^n, Z_{t_k}^n)_{k=0,\dots,n}$ and numerical experiments for three different terminal conditions. In each case the exact solution is available and we are able to compute the error $(Y^n - Y, Z^n - Z)$ in L_2-norm.

4.1 Simulation of $(\tau_1, \dots, \tau_n)$ and B^n

In order to simulate $(\tau_1, \dots, \tau_n)$, we use the fact that

$$\tau_0 = 0 \qquad \text{and} \qquad \forall k \geq 1, \qquad \tau_k = \tau_{k-1} + \sigma_k,$$

where $(\sigma_k)_{1 \leq n}$ is an i.i.d. sequence whose common law σ represents the first exit time of the Brownian motion B of the interval $[-\sqrt{h}, \sqrt{h}]$,

$$\sigma := \inf\{t > 0 : |B_t| = \sqrt{h}\}$$

From the book of Borodin and Salminen [4], we have that the Laplace transform of σ is given by $\mathbb{E}(e^{-\lambda\sigma}) = \frac{1}{\cosh(\sqrt{2\lambda h})}$.

Let F denote the cumulative distribution function of σ. It holds $\mathbb{E}(e^{-\lambda\sigma}) = \lambda \hat{F}(\lambda)$, where $\hat{F}$ is the Laplace transform of F. Then, to obtain F, it remains to inverse numerically its Laplace transform. Once we have F, we simulate the sequence $(\sigma_k)_{1 \leq k \leq n}$ by following the steps of Algorithm 1.

Algorithm 1 Simulation of the sequence $(\tau_1, \dots, \tau_n)$

Simulate one vector with uniform law $(U_1, \dots, U_n)$

$\tau_0 = 0$
for $k = 1 : n$ **do**
  Compute $\sigma_k := F^{-1}(U_k)$
  Define $\tau_k = \tau_{k-1} + \sigma_k$
end for

4.2 Simulation of B^n

In order to get the trajectory $B^n_{t_1}, \dots, B^n_{t_n}$ ($B^n_{t_0} = 0$), we simulate an i.i.d. Bernoulli sequence $(\xi_k)_{1\leq k\leq n}$ i.e. $\mathbb{P}(\xi_k = \pm 1) = \frac{1}{2}$. Then

$$B^n_{t_{k+1}} = \begin{cases} B^n_{t_k} + \sqrt{h} & \text{if } \xi_k = 1 \\ B^n_{t_k} - \sqrt{h} & \text{otherwise.} \end{cases} \tag{15}$$

4.3 Simulation of (Y^n, Z^n)

Since B^n is built using the random walk (15), it can be represented by a recombining binomial tree. Both $(Y^n_{t_k})_{0\leq k\leq n}$ and $(Z^n_{t_k})_{0\leq k\leq n-1}$ can then also be represented as a recombining binomial tree. Since $Y^n_{t_n} = g(B^n_{t_n})$, we solve backward in time the BSDE by following these equalities, ensuing from (2) ($Y^n_{t_k}$ has been replaced by $Y^n_{t_{k+1}}$ in the generator term, but the error induced by this modification is smaller than the ones we consider)

$$\begin{aligned} Z^n_{t_k} &= \frac{1}{\sqrt{h}} \mathbb{E}_{\tau_k}(Y^n_{t_{k+1}} \varepsilon_{k+1}), \\ Y^n_{t_k} &= \mathbb{E}_{\tau_k}(Y^n_{t_{k+1}} + h f(t_{k+1}, B^n_{t_k}, Y^n_{t_{k+1}}, Z^n_{t_k})). \end{aligned}$$

4.4 Study of the error $\mathbb{E}|Y^n_{t_k} - Y_{t_k}|^2$ and $\mathbb{E}|Z^n_{t_k} - Z_{t_k}|^2$

In this subsection we assume that we are able to compute the exact solution (Y, Z). We want to study numerically the convergence in n of $\mathbb{E}|Y^n_{t_k} - Y_{t_k}|^2$ and $\mathbb{E}|Z^n_{t_k} - Z_{t_k}|^2$, where (Y, Z) solves (1) and (Y^n, Z^n) solves (6). To do so, we approximate the error $\mathbb{E}|A^n_{t_k} - A_{t_k}|^2$ ($A = Y$ or $A = Z$) by Monte Carlo:

$$\mathbb{E}|A^n_{t_k} - A_{t_k}|^2 \sim \frac{1}{M} \sum_{m=1}^{M} |A^{n,m}_{t_k} - A^m_{t_k}|^2 := E_A \tag{16}$$

1. For each Monte Carlo simulation, we pick at random one sequence $(\xi_1, \dots, \xi_n)$ (which gives the value of $(B^n_{t_1}, \dots, B^n_{t_n})$) and one sequence $(\tau_1, \dots, \tau_n)$.
2. From the sequence $(\xi_1, \dots, \xi_n)$ we get the trajectory of Y^n, including $Y^n_{t_k}$.
3. From the sequence $(B_{\tau_1}, \dots, B_{\tau_n})$ (which is equal to $(B^n_{t_1}, \dots, B^n_{t_n})$), we compute B_{t_k} by using the Brownian bridge method. We deduce (Y_{t_k}, Z_{t_k}) as functions of B_{t_k}.

In the following experiments, we plot the logarithm of the errors E_Y and E_Z (defined in (16)) w.r.t. $\log(n)$. From Theorem 3.1, we get that $\log(E_Y)$ and $\log(E_Z)$ decrease as $-\frac{\alpha}{2}\log(n)$. By using a linear regression, we compute the slope of the line solving the least square problem and compare it to $-\frac{\alpha}{2}$.

4.5 Numerical Experiment

4.5.1 Case $g(x) = e^{T+x}$ and $f(y, z) = y + z$

We consider the BSDE with terminal condition $g(x) = e^{T+x}$ and driver $f(y, z) = y + z$. In this case, we know that $Y_t = e^{T+B_t+\frac{5}{2}(T-t)}$. We run $M = 20000$ Monte Carlo simulations.

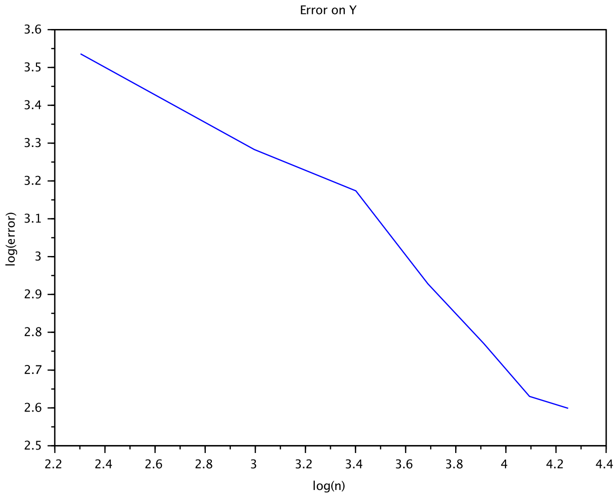


Figure 1: log(error on Y) w.r.t. $\log(n)$ for $f(y,z) = y + z$ and $g(x) = e^{T+x}$.

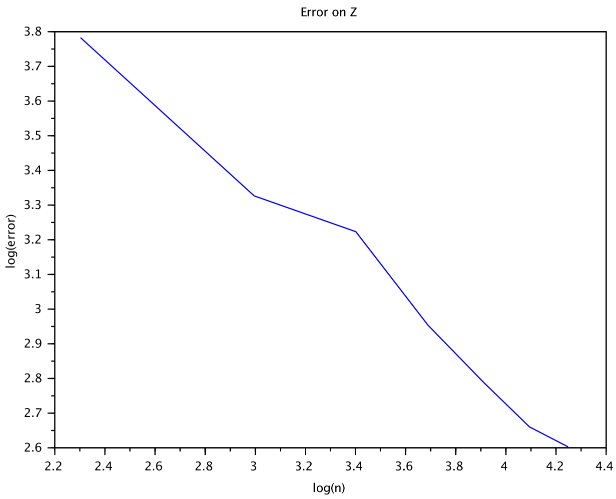


Figure 2: log(error on Z) w.r.t. $\log(n)$ for $f(y,z) = y + z$ and $g(x) = e^{T+x}$.

Figure 1 (resp. Figure 2) represents log(error on Y) (the error is defined by (16)) (resp. log(error on Z)) with respect to $\log(n)$. For the Y case, the slope ensuing from the linear regression is -0.53. Even though $g(x) = e^{T+x}$ does not satisfy (3), g is locally Lipschitz continuous, and the outcome seems to be consistent with Theorem 3.1 for $\alpha = 1$. For the Z case, we get the slope -0.61.

4.5.2 Case $g(x) = x^2$ and $f(y,z) = y + z$

In that case, we know that $Y_t = e^{T-t}((B_t - (T-t))^2 + T - t)$ and $Z_t = 2e^{T-t}(B_t - (T-t))$. We run $M = 20000$ Monte Carlo simulations.

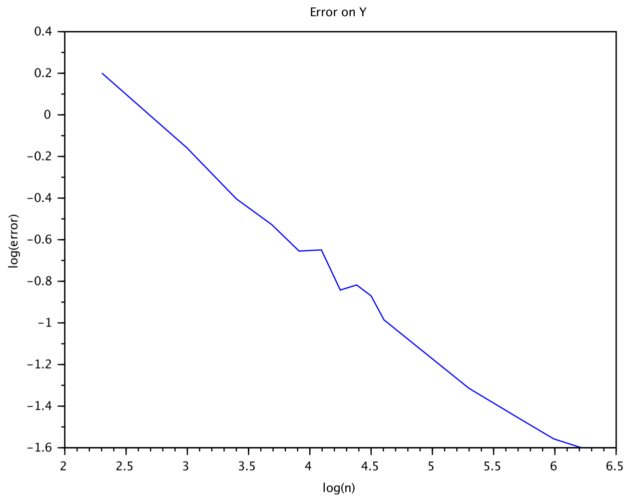


Figure 3: log(error on Y) w.r.t. $\log(n)$ for $f(y,z) = y + z$ and $g(x) = x^2$.

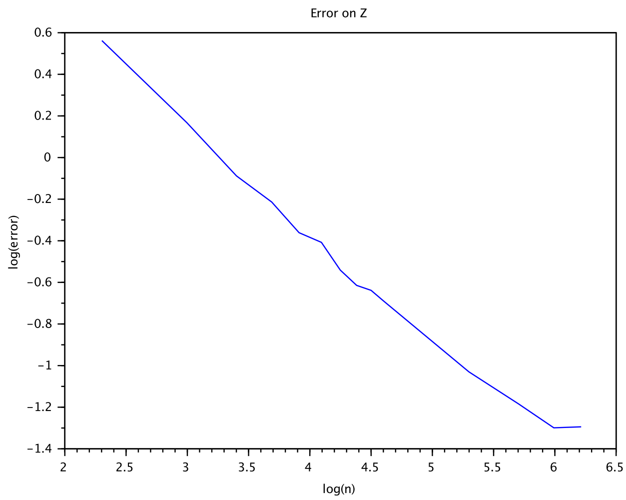


Figure 4: log(error on Z) w.r.t. $\log(n)$ for $f(y,z) = y + z$ and $g(x) = x^2$.

Figure 3 represents log(error on Y) with respect to $\log(n)$. The slope of the linear regression is -0.465. Figure 4 represents log(error on Z) with respect to $\log(n)$. The slope of the linear regression is -0.48. The results are then consistent with Theorem 3.1.

4.5.3 Case $g(x) = \sqrt{|x|}$ and $f(y,z) = y + z$

In that case, we know that $Y_t = e^{\frac{T-t}{2}} \tilde{\mathbb{E}}(\sqrt{|\tilde{B}_{T-t} + B_t|} e^{\tilde{B}_{T-t}})$. We run $M = 20000$ Monte Carlo simulations.

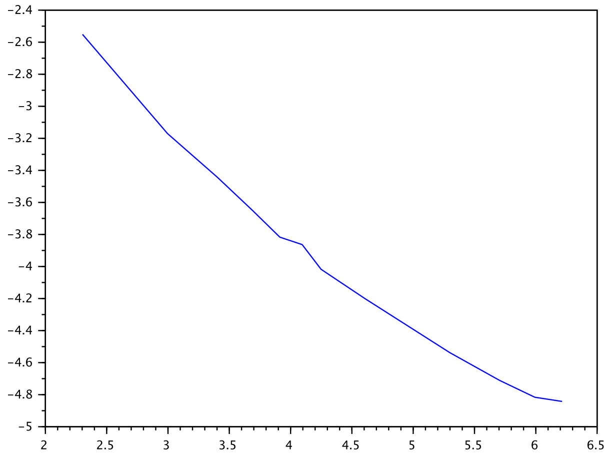


Figure 5: log(error on Y) w.r.t. $\log(n)$ for $f(y,z) = y + z$ and $g(x) = \sqrt{|x|}$.

Figure 5 represents log(error on Y) with respect to $\log(n)$. The slope of the linear regression -0.56. Here we notice that the modulus of the slope we get is larger than $\frac{1}{4}$, the upper bound obtained in that case in Theorem 3.1.

5 Some properties of solutions to PDEs and BSDEs

In the following we recall and prove results for FBSDEs with a general forward process, even though we apply them in the present paper only for the case where the forward process is just the Brownian motion. Restricting ourselves to the case of Brownian motion would not shorten the proofs considerably. Let us consider the following SDE started in (t,x),

$$X_s^{t,x} = x + \int_t^s b(r, X_r^{t,x})dr + \int_t^s \sigma(r, X_r^{t,x})dB_r, \quad 0 \le t \le s \le T, \tag{17}$$

where b and σ satisfy

Assumption 5.1.

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

2. *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* $d((x,t),(x',t')) = (|x-x'|^2 + |t-t'|)^{\frac{1}{2}}$ *on all compact subsets of* $[0,T] \times \mathbb{R}$,

3. b, σ *are* $\frac{1}{2}$*-Hölder continuous in time, uniformly in space,*

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

5.1 Malliavin weights

In this section we recall the Malliavin weights and their properties from [22, Subsection 1.1 and Remark 3].

Lemma 5.2. *Let $H : \mathbb{R} \to \mathbb{R}$ be a polynomially bounded Borel function. If Assumption 5.1 holds and $X^{t,x}$ is given by (17), then setting*

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

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

$$\partial_x G(r, X_r^{t,x}) = \mathbb{E}[H(X_R^{t,x}) N_R^{r,1,(t,x)} | \mathcal{F}_r^t], \quad \text{and} \quad \partial_x^2 G(r, X_r^{t,x}) = \mathbb{E}[H(X_R^{t,x}) N_R^{r,2,(t,x)} | \mathcal{F}_r^t],$$

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

$$N_R^{r,1,(t,x)} = \frac{1}{R-r} \int_r^R \frac{\nabla X_s^{t,x}}{\sigma(s, X_s^{t,x}) \nabla X_r^{t,x}} dB_s \text{ and } N_R^{r,2,(t,x)} = \frac{N_R^{\rho,1,(t,x)} \nabla X_R^{t,x} N_\rho^{r,1,(t,x)} + \nabla N_R^{\rho,1,(t,x)}}{\nabla X_r^{t,x}},$$

with $\rho := \frac{r+R}{2}$. Moreover, for $q \in (0,\infty)$ it holds a.s.

$$(\mathbb{E}[|N_R^{r,i,(t,x)}|^q | \mathcal{F}_r^t])^{\frac{1}{q}} \le \frac{\kappa_q}{(R-r)^{\frac{i}{2}}}, \tag{18}$$

and $\mathbb{E}[N_R^{r,i,(t,x)} | \mathcal{F}_r^t] = 0$ a.s. for $i = 1,2$. Finally, we have

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

and

$$\|\partial_x^2 G(r, X_r^{t,x})\|_{L_p(\mathbb{P})} \le \kappa_q \frac{\|H(X_R^{t,x}) - \mathbb{E}[H(X_R^{t,x}) | \mathcal{F}_r^t]\|_{L_p(\mathbb{P})}}{R-r}$$

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

5.2 Regularity of solutions to BSDEs

Let us now consider the FBSDE

$$Y_s^{t,x} = g(X_T^{t,x}) + \int_s^T f(r, X_r^{t,x}, Y_r^{t,x}, Z_r^{t,x}) dr - \int_s^T Z_r^{t,x} dB_r, \qquad 0 \le t \le s \le T, \tag{19}$$

where $X^{t,x}$ is the process satisfying (17). The following result is taken from [22, Theorem 1]. We reformulate it here for the simple situation where we need it. On the other hand, we will use $\mathbb{P}_{t,x}$ and are interested in an estimate for all $(t,x) \in [0,T) \times \mathbb{R}$.

Theorem 5.3. *Let Assumption 2.1 and 5.1 hold. Then for any $p \in [2,\infty)$ the following assertions are true.*

(i) There exists a constant $C_{5.3}^y > 0$ such that for $0 \le t < s < T$ and $x \in \mathbb{R}$,

$$\|Y_s - Y_t\|_{L_p(\mathbb{P}_{t,x})} \le C_{5.3}^y \Psi(x) \left(\int_t^s (T-r)^{\alpha-1} dr \right)^{\frac{1}{2}},$$

(ii) There exists a constant $C^z_{5.3} > 0$ such that for $0 \le t < s < T$ and $x \in \mathbb{R}$,

$$\|Z_s - Z_t\|_{L_p(\mathbb{P}_{t,x})} \le C^z_{5.3}\Psi(x)\left(\int_t^s (T-r)^{\alpha-2}dr\right)^{\frac{1}{2}}.$$

The constants $C^y_{5.3}$ and $C^z_{5.3}$ depend on $K_f, L_f, C_g, c^{1,2}_{5.4}, T, p_0, b, \sigma, \kappa_q$ and p.

Proof of Theorem 5.3. (i) First we follow the step [22, Theorem 1, proof of $(C2_l) \implies (C3_l)$]. We conclude from the linear growth $|f(r,x,y,z)| \le L_f(|x|+|y|+|z|) + K_f$ and from the Burkholder-Davis-Gundy inequality with constant $a_p > 0$ that

$$\begin{aligned}
&\|Y_s - Y_t\|_{L_p(\mathbb{P}_{t,x})} \\
&= \left\| \int_t^s f(r, X_r, Y_r, Z_r)dr - \int_t^s Z_r dB_r \right\|_{L_p(\mathbb{P}_{t,x})} \\
&\le K_f(s-t) + L_f \int_t^s \|X_r\|_{L_p(\mathbb{P}_{t,x})} + \|Y_r\|_{L_p(\mathbb{P}_{t,x})} + \|Z_r\|_{L_p(\mathbb{P}_{t,x})}dr + a_p \left(\int_t^s \|Z_r\|^2_{L_p(\mathbb{P}_{t,x})}dr\right)^{\frac{1}{2}}.
\end{aligned}$$

We then use (i) and (ii) of Theorem 5.4 below to get

$$\begin{aligned}
&\|Y_s - Y_t\|_{L_p(\mathbb{P}_{t,x})} \\
&\le K_f(s-t) + C(T, L_f, c^{1,2}_{5.4}, p, b, \sigma, p_0)\Psi(x)\left[\int_t^s \left(1 + (T-r)^{\frac{\alpha-1}{2}}\right)dr + \left(\int_t^s (T-r)^{\alpha-1}dr\right)^{\frac{1}{2}}\right].
\end{aligned}$$

(ii) Here one can follow [22, Theorem 1, proof of $(C4_l) \implies (C1_l)$].
Step 1: We first assume additionally that $f : [0,T] \times \mathbb{R}^3 \to \mathbb{R}$ is continuously differentiable in x, y, and z with uniformly bounded derivatives as it was assumed for [22, Theorem 1]. To take the dependency on x into consideration which arises since we use $\mathbb{P}_{t,x}$, it suffices to replace everywhere in the proof in [22] the constant $c_{B^{\Theta}_{p,\infty}}$ by $C(T, C_g, \sigma, b, p, p_0)\Psi(x)$. The constant $C^z_{5.3}$ depends moreover on L_f and κ_q.
Step 2: Now let f be as in Assumption 5.1. In [22, Theorem 1, proof of $(C4_l) \implies (C1_l)$] a linear BSDE is used which describes the behaviour of the process Z minus its counterpart where the generator is identically 0. Here the partial derivatives of f_x, f_y, f_z appear but only their uniform bound is needed in the estimates. Hence if f satisfies (4), we can use mollifying as explained in (25) below (one may choose $N = \infty$). Since $|\partial_x f^\varepsilon(t,x,y,z)|, |\partial_y f^\varepsilon(t,x,y,z)|$ and $|\partial_z f^\varepsilon(t,x,y,z)|$ are bounded by L_f we conclude from Step 1 that for all $\varepsilon > 0$ the process Z^ε corresponding to f^ε satisfies

$$\|Z^\varepsilon_s - Z^\varepsilon_t\|_{L_p(\mathbb{P}_{t,x})} \le C^z_{5.3}\Psi(x)\left(\int_t^s (T-r)^{\alpha-2}dr\right)^{\frac{1}{2}} \tag{20}$$

for $p \ge 2$. Especially, the family $\{|Z^\varepsilon_s - Z^\varepsilon_t|^q : \varepsilon > 0\}$ is then uniformly integrable provided that $q < p$. By an a priori estimate (cf. [6, Lemma 3.1]) we have that

$$\mathbb{E}\int_0^T |Z_r - Z^\varepsilon_r|^2 dr \le C\int_0^T \sup_{x,y,z} |f(r,x,y,z) - f^\varepsilon(r,x,y,z)|^2 dr \le C\varepsilon^2 T L_f^2.$$

Fubini's theorem implies that there exists a sequence $\varepsilon_m \to 0$ and a measurable set $N \subseteq [0,T]$ of Lebesgue measure zero, such that $\lim_{m\to\infty} \mathbb{E}|Z_r - Z^{\varepsilon_m}_r|^2 = 0$ for all $r \in [0,T] \setminus N$. Consequently, for any $q < p$ and all $t, s \in [0,T] \setminus N$ with $t < s$,

$$\|Z_s - Z_t\|_{L_q(\mathbb{P}_{t,x})} \le C^z_{5.3}\Psi(x)\left(\int_t^s (T-r)^{\alpha-2}dr\right)^{\frac{1}{2}}.$$

The assertion follows for all $q \geq 2$ since (20) holds for all $p \in [2, \infty)$. Since by Theorem 5.4 (ii) the process Z does have a continuous version, we finally get the assertion for all $t < s$. □

5.3 Properties of the associated PDE

We collect in the theorem below properties of the solution to the PDE which are mainly known. The new part concerns $\partial_x^2 u$. For Lipschitz continuous g, the behaviour of $\partial_x^2 u$ has been studied in [37]. General results related to this topic can be found in [20].

Theorem 5.4. *Consider the FBSDE (19) and let Assumptions 2.1 and 5.1 hold. Then for the solution u of the associated PDE*

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

we have

(i) $Y_t = u(t, X_t)$ *where* $u(t,x) = \mathbb{E}_{t,x}\left(g(X_T) + \int_t^T f(r, X_r, Y_r, Z_r)dr\right)$ *and* $|u(t,x)| \leq c^1_{5.4}\Psi(x)$ *with* Ψ *given in (5), where* $c^1_{5.4}$ *depends on* C_g, T, p_0, L_f, K_f *and on the bounds and Lipschitz constants of* b *and* σ.

(ii) u_x *exists,*

$$u_x(t,x) = \mathbb{E}_{t,x}\left(g(X_T)N_T^{t,1} + \int_t^T f(r, X_r, Y_r, Z_r)N_r^{t,1}dr\right), \tag{21}$$

and

(a) u_x *is continuous in* $[0,T) \times \mathbb{R}$,

(b) $Z_s^{t,x} = u_x(s, X_s^{t,x})\sigma(s, X_s^{t,x})$,

(c) $|u_x(t,x)| \leq \frac{c^2_{5.4}\Psi(x)}{(T-t)^{\frac{1-\alpha}{2}}}$,

where $c^2_{5.4}$ *depends on* $C_g, T, p_0, \kappa_2, L_f, K_f$ *and on the bounds and Lipschitz constants of* b *and* σ.

(iii) u_{xx} *exists,*

$$u_{xx}(t,x) = \mathbb{E}_{t,x}\left(g(X_T)N_T^{t,2} + \int_t^T [f(r, X_r, Y_r, Z_r) - f(r, X_t, Y_t, Z_t)]N_r^{t,2}dr\right), \tag{22}$$

and

(a) u_{xx} *is continuous in* $[0,T) \times \mathbb{R}$,

(b) $|u_{xx}(t,x)| \leq \frac{c^3_{5.4}\Psi(x)}{(T-t)^{1-\frac{\alpha}{2}}}$,

where $c^3_{5.4}$ *depends on* $C_g, T, p_0, \kappa_2, L_f, C^y_{5.3}, C^z_{5.3}$ *and on the bounds and Lipschitz constants of* b *and* σ.

In the following $c_{5.4}$ *represents* $(c^1_{5.4}, c^2_{5.4}, c^3_{5.4})$ *and* $c^{i,j}_{5.4}$ $(i \neq j)$ *represents* $(c^i_{5.4}, c^j_{5.4})$, $(i,j) \in \{1,2,3\}$.

Proof. (i): This follows from [36, Theorem 3.2].
(ii): From the proof of [36, Theorem 3.2], we get (21). The points (ii)(a) and (b) ensue from [36, Theorem 3.2 (i)]. It remains to prove (c).
Proof of (ii) **(c):** We show the assertion for a generator not depending on X, since the terms arising from that dependency would be easy to treat. Since $\mathbb{E}_{t,x}(\mathbb{E}_{t,x}(g(X_T))N_T^{t,1}) = 0$ we can subtract it from the right hand side of (21) and get

$$\partial_x u(t,x) = \mathbb{E}_{t,x}\left([g(X_T) - \mathbb{E}_{t,x}(g(X_T))]N_T^{t,1} + \int_t^T f(r, Y_r, Z_r)N_r^{t,1}dr\right).$$

It holds

$$\mathbb{E}_{t,x}|g(X_T) - \mathbb{E}_{t,x}g(X_T)|^2 = \mathbb{E}_{t,x}|g(X_T) - \tilde{\mathbb{E}}g(\tilde{X}_T^{t,X_t})|^2 \le \mathbb{E}_{t,x}\tilde{\mathbb{E}}|g(X_T) - g(\tilde{X}_T^{t,X_t})|^2,$$

and thanks to the Cauchy-Schwarz inequality with $\Psi_1 = C_g(1 + |X_T|^{p_0} + |\tilde{X}_T^{t,X_t}|^{p_0})$ and equation (3),

$$\begin{aligned}\mathbb{E}_{t,x}\tilde{\mathbb{E}}|g(X_T) - g(\tilde{X}_T^{t,X_t})|^2 &\le \mathbb{E}_{t,x}\tilde{\mathbb{E}}(\Psi_1^2|X_T - \tilde{X}_T^{t,X_t}|^{2\alpha})\\ &\le \left[\mathbb{E}_{t,x}\tilde{\mathbb{E}}\Psi_1^4\right]^{\frac{1}{2}}\left[\mathbb{E}_{t,x}\tilde{\mathbb{E}}|X_T - \tilde{X}_T^{t,X_t}|^{4\alpha}\right]^{\frac{1}{2}}\\ &\le C(C_g, T, p_0, b, \sigma)\Psi^2(x)(T-t)^{\alpha}. \end{aligned} \tag{23}$$

Relation (18) and the Lipschitz continuity of f imply

$$\begin{aligned}|\partial_x u(t,x)| \le &\frac{C(C_g, T, p_0, \kappa_2, b, \sigma)\Psi(x)}{(T-t)^{\frac{1-\alpha}{2}}}\\ &+ C(L_f, K_f)\mathbb{E}_{t,x}\int_t^T (1 + |u(r, X_r)| + |\partial_x u(r, X_r)\sigma(r, X_r)|)|N_r^{t,1}|dr. \end{aligned}\tag{24}$$

Since we have $|g(x)| \le \Psi(x)$, [36, Theorem 3.2 (ii)] gives $|u(t,x)| \le c\Psi(x)$ and $|\partial_x u(t,x)| \le c\Psi(x)(T-t)^{-1/2}$, where c depends on $T, L_f, K_f, \kappa_2, b, \sigma$ and p_0. Hence inequality (24) becomes

$$\begin{aligned}|\partial_x u(t,x)| &\le \frac{C(C_g, T, p_0, \kappa_2, b, \sigma)\Psi(x)}{(T-t)^{\frac{1-\alpha}{2}}} + C(L_f, K_f, c, \sigma)\mathbb{E}_{t,x}\left(\int_t^T \left(1 + \Psi(X_r) + \frac{\Psi(X_r)}{(T-r)^{\frac{1}{2}}}\right)|N_r^{t,1}|dr\right)\\ &\le \frac{C(C_g, T, p_0, \kappa_2, b, \sigma)\Psi(x)}{(T-t)^{\frac{1-\alpha}{2}}} + C(T, L_f, K_f, \kappa_2, b, \sigma, p_0)\int_t^T \frac{\Psi(x)}{(T-r)^{\frac{1}{2}}(r-t)^{\frac{1}{2}}}dr\\ &\le \frac{C(C_g, T, p_0, \kappa_2, L_f, K_f, b, \sigma)\Psi(x)}{(T-t)^{\frac{1-\alpha}{2}}}.\end{aligned}$$

(iii): We start with an approximation of g and f by smooth and bounded functions. Let ϕ be a non-negative C^∞ function with support $[-1,1]$, such that $\int_{\mathbb{R}}\phi(u)du = 1$, and $\varepsilon \in (0,1]$. For $N \in \mathbb{N}$ let $b_N : \mathbb{R} \to [-N-1, N+1]$ be a monotone C^∞ function such that $0 \le b_N'(x) \le 1$ and

$$b_N(x) := \begin{cases} N+1, & x > N+2,\\ x, & |x| \le N,\\ -N-1, & x < -N-2.\end{cases}$$

Define

$$g^{\varepsilon,N}(x)=\int_{-1}^{1}\phi(u)g(b_N(x)-\varepsilon u)du$$

and

$$f^{\varepsilon,N}(r,y,z)=\int_{-1}^{1}\int_{-1}^{1}\phi(u)\phi(v)f(r,b_N(y)-\varepsilon u,b_N(z)-\varepsilon v)dudv. \tag{25}$$

Lemma 5.5. *$g^{\varepsilon,N}$ and $f^{\varepsilon,N}$ satisfy*

(a) $\|g^{\varepsilon,N}\|_\infty+\|f^{\varepsilon,N}\|_\infty\le C=C(\varepsilon,N)$ for some $C(\varepsilon,N)>0$,

(b) $g^{\varepsilon,N}$ and $f^{\varepsilon,N}$ are C^∞ functions, with bounded derivatives (the bounds depend on ε and N). Moreover, $f^{\varepsilon,N}$ is a Lipschitz function in y and z, with Lipschitz constant L_f,

(c) $g^{\varepsilon,N}$ satisfies (3), uniformly in $\varepsilon\in(0,1)$ and $N\ge 1$,

(d) for all $x\in\mathbb{R}$ and $\varepsilon\in[0,1]$, we have $|g^{\varepsilon,N}(x)-g(x)|\le C(C_g)\Psi(x)(\varepsilon^\alpha+\frac{|x|^{\alpha+1}}{N})$,

(e) for all $r\in[0,T]$ and for all $(y,z)\in\mathbb{R}^2$, we have

$$|f^{\varepsilon,N}(r,y,z)-f(r,y,z)|\le L_f(2\varepsilon+|b_N(y)-y|+|b_N(z)-z|).$$

Proof. (a) Since g is locally Hölder continuous in the sense of (3), $|g(x)|\le C_g(1+|x|^{p_0+1})$. Then, we get $|g^{\varepsilon,N}(x)|\le C_g(1+(N+1+\varepsilon)^{p_0+1})$, and for f being Lipschitz continuous in y and z, uniformly in time, the same type of result applies.

(b) Since ϕ is a C^∞ function and f and g are of polynomial growth, we get the result.

(c) Since g is locally Hölder continuous, we get

$$\begin{aligned}|g^{\varepsilon,N}(x)-g^{\varepsilon,N}(y)|&\le\int_{-1}^{1}|\phi(u)|C_g(1+|b_N(x)-\varepsilon u|^{p_0}+|b_N(y)-\varepsilon u|^{p_0})|b_N(x)-b_N(y)|^\alpha du\\&\le\int_{-1}^{1}C_g|\phi(u)|(1+(|x|+\varepsilon)^{p_0}+(|y|+\varepsilon)^{p_0})|x-y|^\alpha du\\&\le C(C_g)(1+|x|^{p_0}+|y|^{p_0})|x-y|^\alpha.\end{aligned}$$

(d) We have

$$\begin{aligned}|g^{\varepsilon,N}(x)-g(x)|&=\left|\int_{-1}^{1}\phi(u)(g(b_N(x)-\varepsilon u)-g(x))du\right|\\&\le C_g\int_{-1}^{1}|\phi(u)|(1+|b_N(x)|^{p_0}+\varepsilon^{p_0}+|x|^{p_0})(|b_N(x)-x|^\alpha+\varepsilon^\alpha)du\\&\le C(C_g)(1+|x|^{p_0})(\varepsilon^\alpha+|x|^\alpha\mathbf{1}_{|x|\ge N}),\end{aligned}$$

and the result follows.

(e) We simply have to apply the Lipschitz property of f to get the result.

□

We put now $\varepsilon := \frac{1}{N}$ and write (g^N, f^N) instead of $(g^{\frac{1}{N},N}, f^{\frac{1}{N},N})$ in order to simplify the notation and consider the BSDE

$$Y_t^N = g^N(X_T) + \int_t^T f^N(r, Y_r^N, Z_r^N)dr - \int_t^T Z_r^N dB_r.$$

Representation for $\partial_x^2 u^N(t,x)$.
By (i) we have that

$$u^N(t,x) = \mathbb{E}_{t,x} g^N(X_T^{t,x}) + \int_t^T \mathbb{E}_{t,x} f^N(r, Y_r^N, Z_r^N)dr.$$

According to Lemma 5.2 it holds that $\partial_x^2 \mathbb{E}_{t,x} g^N(X_T) = \mathbb{E}_{t,x}[g^N(X_T)N_T^{t,2}]$ and

$$\partial_x^2 \mathbb{E}_{t,x} f^N(r, Y_r^N, Z_r^N) = \mathbb{E}_{t,x}[f^N(r, Y_r^N, Z_r^N)N_r^{t,2}],$$

because

$$f^N(r, Y_r^N, Z_r^N) = f^N(r, u^N(r, X_r), \sigma(r, X_r)u_x^N(r, X_r)),$$

and $f^N(r,y,z)$ is continuous and bounded. Moreover, [25, Proposition 4] (or [21, Theorem 2.1]) implies that $u^N(r,x)$ is $C^{1,2}$ and it holds that $|u^N(r,x)| + |\partial_x u^N(r,x)| + |\partial_x^2 u^N(r,x)| \le C^N$ for some $C^N > 0$. Since σ is continuous,

$$(r,x) \mapsto f^N(r, u^N(r,x), \sigma(r,x)u_x^N(r,x))$$

is a bounded Borel function. Notice that by Lemma 5.2

$$\mathbb{E}_{t,x}[N_r^{t,2}] = 0 \quad \text{and} \quad \mathbb{E}_{t,x}[(N_r^{t,2})^2] \le \frac{\kappa_2^2}{(r-t)^2}, \tag{26}$$

so that

$$\mathbb{E}_{t,x}[f^N(r, Y_r^N, Z_r^N)N_r^{t,2}] = \mathbb{E}_{t,x}([f^N(r, Y_r^N, Z_r^N) - f^N(r, Y_t^N, Z_t^N)]N_r^{t,2}).$$

Using the Lipschitz continuity of f^N (see Lemma 5.5), the inequality of Cauchy-Schwarz and Theorem 5.3 one can derive the upper bound

$$\begin{aligned}
&|\partial_x^2 \mathbb{E}_{t,x} f^N(r, Y_r^N, Z_r^N)| \\
&\quad\le \mathbb{E}_{t,x}[|f^N(r, Y_r^N, Z_r^N) - f^N(r, Y_t^N, Z_t^N)||N_r^{t,2}|] \\
&\quad\le C(L_f, \kappa_2)(\mathbb{E}_{t,x}(|Y_r^N - Y_t^N|^2 + |Z_r^N - Z_t^N)|^2))^{\frac{1}{2}} \frac{1}{r-t} \\
&\quad\le C(L_f, \kappa_2, C_{5.3}^y, C_{5.3}^z) \left[\left(\int_t^r (T-s)^{\alpha-1} ds \right)^{\frac{1}{2}} + \left(\int_t^r (T-s)^{\alpha-2} ds \right)^{\frac{1}{2}} \right] \frac{\Psi(x)}{r-t} \\
&\quad\le C(T, L_f, \kappa_2, C_{5.3}^y, C_{5.3}^z)\Psi(x) \frac{1}{(T-r)^{1-\frac{\alpha}{2}}(r-t)^{\frac{1}{2}}}.
\end{aligned} \tag{27}$$

By this we do have an integrable bound for the derivative, and by dominated convergence we get

$$\begin{aligned}
\partial_x^2 \int_t^T \mathbb{E}_{t,x} f^N(r, Y_r^N, Z_r^N)dr &= \int_t^T \partial_x^2 \mathbb{E}_{t,x} f^N(r, Y_r^N, Z_r^N)dr \\
&= \int_t^T \mathbb{E}_{t,x}\{[f^N(r, Y_r^N, Z_r^N) - f^N(r, Y_t^N Z_t^N)]N_r^{t,2}\}dr.
\end{aligned}$$

Hence we can write (using Fubini's theorem for the integral)

$$\partial_x^2 u^N(t,x) = \mathbb{E}_{t,x}\left(g^N(X_T) N_T^{t,2} + \int_t^T [f^N(r, Y_r^N, Z_r^N) - f^N(r, Y_t^N, Z_t^N)] N_r^{t,2} dr \right).$$

Convergence of $\partial_x^2 u^N(t,x)$. Since $\mathbb{E}_{t,x}[\mathbb{E}_{t,x}(g^N(X_T))N_T^{t,2}] = 0$, Cauchy-Schwarz's inequality and the local Hölder continuity of g^N (see Lemma 5.5) give like in (23) that

$$\begin{aligned}
|\mathbb{E}_{t,x}(g^N(X_T)N_T^{t,2})| &= \left| \mathbb{E}_{t,x}\left([g^N(X_T) - \mathbb{E}_{t,x}(g^N(X_T))] N_T^{t,2} \right) \right| \\
&\le \left(\mathbb{E}_{t,x}(|g^N(X_T) - \mathbb{E}_{t,x}(g^N(X_T))|^2) \right)^{\frac{1}{2}} \frac{\kappa_2}{T-t} \\
&\le C(C_g, T, p_0, \kappa_2, b, \sigma) \frac{\Psi(x)}{(T-t)^{1-\frac{\alpha}{2}}},
\end{aligned}$$

for all $N \in \mathbb{N}$. For the second term we can use the upper bound (27) and Lemma A.2 to get

$$\begin{aligned}
&\mathbb{E}_{t,x} \int_t^T \left| [f^N(r, Y_r^N, Z_r^N) - f^N(r, Y_t^N, Z_t^N)] N_r^{t,2} \right| dr \\
&\qquad \le C(T, L_f, \kappa_2, C_{5.3}^y, C_{5.3}^z) \int_t^T \frac{\Psi(x)}{(T-r)^{1-\frac{\alpha}{2}} (r-t)^{\frac{1}{2}}} dr, \\
&\qquad \le C(T, L_f, \kappa_2, C_{5.3}^y, C_{5.3}^z) \Psi(x) \frac{B(\frac{\alpha}{2}, \frac{1}{2})}{(T-t)^{\frac{1}{2}-\frac{\alpha}{2}}},
\end{aligned}$$

which implies

$$|\partial_x^2 u^N(t,x)| \le C(C_g, T, L_f, p_0, \kappa_2, C_{5.3}^y, C_{5.3}^z, b, \sigma) \frac{\Psi(x)}{(T-t)^{1-\frac{\alpha}{2}}}. \tag{28}$$

According to [21, Theorem 2.1] $\partial_x^2 u^N(t,x)$ is continuous. Let

$$v(t,x) := \mathbb{E}_{t,x}\left(g(X_T) N_T^{t,2} + \int_t^T [f(r, Y_r, Z_r) - f(r, Y_t, Z_t)] N_r^{t,2} dr \right).$$

We show that for any $(t,x) \in [0,T) \times \mathbb{R}$ it holds $\partial_x^2 u^N(t,x) \to v(t,x)$ if $N \to \infty$, and that v is continuous on $[0,T) \times \mathbb{R}$. The idea to show continuity of v is as follows: If $(t_n, x_n) \to (t,x)$, then we may assume that we can find a $\delta > 0$ such that $x_n \in (x-\delta, x+\delta)$ and $t_n \in (t-\delta, t+\delta) \subseteq [0,T)$ for each sufficiently large n. We consider

$$\begin{aligned}
|v(t_n, x_n) - v(t,x)| \quad \le \quad & |v(t_n, x_n) - \partial_x^2 u^N(t_n, x_n)| + |\partial_x^2 u^N(t_n, x_n) - \partial_x^2 u^N(t,x)| \\
& + |\partial_x^2 u^N(t,x) - v(t,x)|.
\end{aligned}$$

Since $\partial_x^2 u^N$ is continuous, the term $|\partial_x^2 u^N(t_n, x_n) - \partial_x^2 u^N(t,x)|$ is small for large n. Hence it suffices to show that $\sup_{s \in (t-\delta, t+\delta), y \in (x-\delta, x+\delta)} |\partial_x^2 u^N(s,y) - v(s,y)|$ is small for large N. Let $(s,y) \in (t-\delta, t+\delta) \times (x-\delta, x+\delta)$. It holds

$$|\partial_x^2 u^N(s,y) - v(s,y)| \le \mathbb{E}_{s,y} |[g^N(X_T) - g(X_T)] N_T^{s,2}| + \int_s^T D^{\frac{1}{2}}(r,s) \frac{\kappa_2}{r-s} dr := D_1 + D_2,$$

where (setting $\| \cdot \|_{\mathbb{P}_{s,y}} := \| \cdot \|_{L_2(\mathbb{P}_{s,y})}$)

$$D(r,s) \quad := \quad \| f^N(r, Y_r^N, Z_r^N) - f^N(r, Y_s^N, Z_s^N) - [f(r, Y_r, Z_r) - f(r, Y_s, Z_s)] \|_{\mathbb{P}_{s,y}}^2$$

$$\leq \quad L_f(\|Y_r^N - Y_s^N\|_{\mathbb{P}_{s,y}} + \|Z_r^N - Z_s^N\|_{\mathbb{P}_{s,y}} + \|Y_r - Y_s\|_{\mathbb{P}_{s,y}} + \|Z_r - Z_s\|_{\mathbb{P}_{s,y}})$$
$$\times(\|f^N(r, Y_r^N, Z_r^N) - f(r, Y_r, Z_r)\|_{\mathbb{P}_{s,y}} + \|f^N(r, Y_s^N, Z_s^N) - f(r, Y_s, Z_s)\|_{\mathbb{P}_{s,y}}).$$

First, let us bound D_1. According to Cauchy-Schwarz's inequality, (30) below and (26) we get

$$D_1 \leq \delta_1 \sqrt{\mathbb{E}_{s,y}(|N_T^{s,2}|^2)} \leq \frac{\delta_1 \kappa_2}{T-s} \leq \frac{\delta_1 \kappa_2}{T-t-\delta}.$$

Now let us bound D_2. According to Theorem 5.3 it holds

$$D^{\frac{1}{2}}(r,s) \leq C(T, L_f, C^y_{5.3}, C^z_{5.3}) \Psi^{\frac{1}{2}}(y) \frac{(r-s)^{\frac{1}{4}}}{(T-r)^{\frac{1}{2}-\frac{\alpha}{4}}}$$
$$\times (\|f^N(r, Y_r^N, Z_r^N) - f(r, Y_r, Z_r)\|_{\mathbb{P}_{s,y}} + \|f^N(r, Y_s^N, Z_s^N) - f(r, Y_s, Z_s)\|_{\mathbb{P}_{s,y}})^{\frac{1}{2}}.$$

Then, using (31), (33), (34) and Proposition 5.6 below gives

$$D^{\frac{1}{2}}(r,s) \leq C(T, L_f, \kappa_2, C^y_{5.3}, C^z_{5.3}) \Psi(y) \frac{(r-s)^{\frac{1}{4}}}{(T-r)^{\frac{1}{2}-\frac{\alpha}{4}}} \frac{\delta_1}{(T-r)^{\frac{1}{4}}}.$$

Hence we have shown that

$$D_2 \leq C(T, L_f, \kappa_2, C^y_{5.3}, C^z_{5.3}) \Psi(y) \int_s^T \frac{\delta_1}{(r-s)^{\frac{3}{4}}(T-r)^{\frac{1}{2}-\frac{\alpha}{4}+\frac{1}{4}}} dr$$
$$\leq C(T, L_f, \kappa_2, C^y_{5.3}, C^z_{5.3}) \Psi(y) \frac{\delta_1}{(T-s)^{\frac{1}{2}-\frac{\alpha}{4}}},$$
$$\leq C(T, L_f, \kappa_2, C^y_{5.3}, C^z_{5.3}) \Psi(x+\delta) \frac{\delta_1}{(T-t-\delta)^{\frac{1}{2}-\frac{\alpha}{4}}} \quad \forall (s,y) \in (t-\delta, t+\delta) \times (x-\delta, x+\delta).$$

Consequently, $\sup_{y \in (x-\delta, x+\delta), s \in (t-\delta, t+\delta)} |\partial_x^2 u^N(s,y) - v(s,y)|$ is small for large N, hence v is continuous. Since

$$\partial_x u^N(t,x) - \partial_x u^N(t,y) = \int_y^x \partial_x^2 u^N(t,z) dz$$

converges to

$$\partial_x u(t,x) - \partial_x u(t,y) = \int_y^x v(t,z) dz,$$

it follows that $\partial_x^2 u(t,x) = v(t,x)$. Then point (iii-a) and (22) are proved. Since $\partial_x^2 u^N$ converges to v for $N \to \infty$, we deduce point (iii-b) from (28). □

Proposition 5.6. *Let Assumptions 5.1 and 2.1 hold. Then for any $(s,y) \in (t-\delta, t+\delta) \times (x-\delta, x+\delta)$ with $t+\delta < T$ and r such that $s \leq r < T$ we have*

$$\|Y_r^N - Y_r\|_{L_2(\mathbb{P}_{s,y})} + \|Z_r^N - Z_r\|_{L_2(\mathbb{P}_{s,y})} \leq \frac{\delta_1}{\sqrt{T-r}},$$

where δ_1 denotes a generic constant which tends to 0 when N tends to $+\infty$.

Proof. Let here $\|\cdot\|$ stand for $\|\cdot\|_{L_2(\mathbb{P}_{s,y})}$. We will use for the Y differences the inequality

$$\|Y_r^N - Y_r\| \leq \|g^N(X_T) - g(X_T)\| + \int_r^T \|f^N(w, Y_w^N, Z_w^N) - f(w, Y_w, Z_w)\|dw.$$

For the Z differences we get by (21) and (ii-b)

$$\begin{aligned}
&\|Z_r^N - Z_r\| \\
\leq\ & C(\sigma)\left(\left\|\mathbb{E}_r\left(g^N(X_T) - g(X_T)\right)N_T^{r,1}\right\| + \left\|\mathbb{E}_r\int_r^T (f^N(w, Y_w^N, Z_w^N) - f(w, Y_w, Z_w))N_w^{r,1}dw\right\|\right) \\
\leq\ & C(\kappa_2, \sigma)\left(\frac{\|g^N(X_T) - g(X_T)\|}{\sqrt{T-r}} + \int_r^T \|f^N(w, Y_w^N, Z_w^N) - f(w, Y_w, Z_w)\|\frac{1}{\sqrt{w-r}}dw\right).
\end{aligned}$$

Let $S(r) := \|Y_r^N - Y_r\| + \|Z_r^N - Z_r\|$. Using the inequality $(1 + \frac{1}{\sqrt{w-r}}) \leq C(T)\frac{1}{\sqrt{w-r}}$ for $r < w \leq T$ gives

$$S(r) \leq C(T, \kappa_2, \sigma)\left(\|g^N(X_T) - g(X_T)\|\frac{1}{\sqrt{T-r}} + \int_r^T \|f^N(w, Y_w^N, Z_w^N) - f(w, Y_w, Z_w)\|\frac{1}{\sqrt{w-r}}dw\right). \tag{29}$$

Let us bound $\|g^N(X_T) - g(X_T)\|$. By Lemma 5.5 we get the estimate

$$\begin{aligned}
\mathbb{E}_{s,y}|g^N(X_T) - g(X_T)|^2 &\leq C(C_g)\mathbb{E}_{s,y}\left(\Psi^4(X_T)\right)^{\frac{1}{2}}\left(\mathbb{E}_{s,y}\left(\frac{1}{N^\alpha} + \frac{|X_T|^{\alpha+1}}{N}\right)^4\right)^{\frac{1}{2}} \\
&\leq C(C_g, T, b, \sigma, p_0)\Psi^2(y)\left(\frac{1}{N^{2\alpha}} + \frac{|y|^{2\alpha+2}}{N^2}\right) \\
&\leq C(C_g, T, b, \sigma, p_0)\Psi^2(|x| + \delta)\left(\frac{1}{N^{2\alpha}} + \frac{(|x| + \delta)^{2\alpha+2}}{N^2}\right) \leq \delta_1^2,
\end{aligned} \tag{30}$$

for any arbitrarily small $\delta_1 > 0$, provided that N is sufficiently large. Let us now bound $\|f^N(w, Y_w^N, Z_w^N) - f(w, Y_w, Z_w)\|$. Using again Lemma 5.5 yields to

$$\begin{aligned}
&\|f^N(w, Y_w^N, Z_w^N) - f(w, Y_w, Z_w)\| \\
&\quad\leq \|f^N(w, Y_w^N, Z_w^N) - f^N(w, Y_w, Z_w)\| + \|f^N(w, Y_w, Z_w) - f(w, Y_w, Z_w)\| \\
&\quad\leq L_f(\|Y_w^N - Y_w\| + \|Z_w^N - Z_w\| + \tfrac{2}{N} + \|b_N(Y_w) - Y_w\| + \|b_N(Z_w) - Z_w\|).
\end{aligned} \tag{31}$$

Then, plugging (30) and (31) into (29) gives

$$\begin{aligned}
S(r) \leq& \frac{C(T, \kappa_2, \sigma)\delta_1}{\sqrt{T-r}} + C(T, \kappa_2, \sigma)L_f\int_r^T \frac{S(w)}{\sqrt{w-r}}dw \\
&+ C(T, \kappa_2, \sigma)L_f\int_r^T \frac{\frac{1}{N} + \|b_N(Y_w) - Y_w\| + \|b_N(Z_w) - Z_w\|}{\sqrt{w-r}}dw.
\end{aligned} \tag{32}$$

To estimate $\|b_N(Z_w) - Z_w\|$ we use $Z_w = \sigma(w, X_w)u_x(w, X_w)$ and choose a small $a > 0$ such that $\beta := \frac{(2+a)(1-\alpha)}{2} < 1$. Then

$$\|b_N(Z_w) - Z_w\|^2 = \mathbb{E}_{s,y}|b_N(Z_w) - Z_w|^2\mathbf{1}_{|Z_w|\geq N} \leq \frac{\mathbb{E}_{s,y}|Z_w|^{2+a}}{N^a} = \frac{\mathbb{E}_{s,y}|\sigma(w, X_w)u_x(w, X_w)|^{2+a}}{N^a}.$$

Using Theorem 5.4 (ii-c) yields

$$\mathbb{E}_{s,y}|b_N(Z_w) - Z_w|^2 \le \frac{C(c_{5.4}^2, \sigma)\mathbb{E}_{s,y}\Psi^{2+a}(X_w)}{(T-w)^{\frac{(2+a)(1-\alpha)}{2}} N^a} \le \frac{C(T, p_0, c_{5.4}^2, \sigma, b)\,\Psi^{2+a}(y)}{(T-w)^{\frac{(2+a)(1-\alpha)}{2}} N^a}$$
$$\le \frac{\delta_1}{(T-w)^{\frac{(2+a)(1-\alpha)}{2}}}, \quad \forall (s,y) \in (t-\delta, t+\delta) \times (x-\delta, x+\delta). \tag{33}$$

Similarly,

$$\mathbb{E}_{s,y}|b_N(Y_w) - Y_w|^2 \le \frac{C(T, p_0, c_{5.4}^1, b, \sigma)\,\Psi^{2+a}(y)}{N^a} \le \delta_1, \quad \forall (s,y) \in (t-\delta, t+\delta) \times (x-\delta, x+\delta). \tag{34}$$

Plugging (33) and (34) into (32) gives

$$S(r) \le \frac{C(T,\kappa_2)\delta_1}{\sqrt{T-r}} + C(T,\kappa_2)L_f \int_r^T \frac{S(w)}{\sqrt{w-r}}dw$$
$$+ C(T,\kappa_2)L_f \int_r^T \frac{\frac{1}{N}+\delta_1}{\sqrt{w-r}} + \frac{\delta_1}{(T-w)^{\frac{(2+a)(1-\alpha)}{2}}\sqrt{w-r}}dw$$
$$\le C(T,\kappa_2,L_f)\left(\frac{\delta_1}{\sqrt{T-r}} + \int_r^T \frac{S(w)}{\sqrt{w-r}}dw\right),$$

where the last inequality comes from Lemma A.2 ($\beta < 1$). It remains to apply a version of Gronwall's Lemma (see e.g. [32, Lemma 3.1]) to see that $S(r) \le \frac{C(T,\kappa_2,L_f)\delta_1}{\sqrt{T-r}}$. Since $C(T,\kappa_2,L_f)\delta_1$ becomes arbitrarily small for N large, we will slightly abuse the notation and write $S(r) \le \frac{\delta_1}{\sqrt{T-r}}$. □

A Technical results and estimates

Lemma A.1. *For all $0 \le k \le m \le n$ and $p > 0$, it holds for $h = \frac{T}{n}$ that*

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

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

(iii) $\mathbb{E}|B_{\tau_m} - B_{\tau_k}|^2 = t_m - t_k$,

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

Proof. The strong Markov property of the Brownian motion implies that $(\tau_i - \tau_{i-1})_{i=1}^\infty$ is an i.i.d. sequence. According to [33, Proposition 11.1 (iii)], we have that $\mathbb{E}\tau_1 = \frac{T}{n}$, and (i) follows. Item (ii) follows by [33, Proposition 11.1 (iv)] and Jensen's inequality. To prove item (iii), recall that $(B_{\tau_i} - B_{\tau_{i-1}})_{i=1}^\infty$ is a centered i.i.d. sequence with $\mathbb{E}(B_{\tau_i} - B_{\tau_{i-1}})^2 = h$, $i \ge 1$. (iv): The BDG inequality implies that for each $p > 0$,

$$\mathbb{E}|B_{\tau_k} - B_{t_k}|^p = \mathbb{E}\left|\int_0^{\tau_k \vee t_k} (\mathbf{1}_{[0,\tau_k]}(r) - \mathbf{1}_{[0,t_k]}(r))dB_r\right|^p$$
$$\le C(p)\ \mathbb{E}\left(\int_0^{\tau_k \vee t_k} \mathbf{1}_{[0,\tau_k]\Delta[0,t_k]}(r)dr\right)^{p/2} = \mathbb{E}|\tau_k - t_k|^{p/2}.$$

To prove the second inequality of (iv), a generalization of [33, Proposition 11.1 (iv)], we first assume that $p \geq 1$. Let us rewrite $\tau_k - t_k = \sum_{i=1}^{k} \eta_i$ where $(\eta_i)_{1\leq i\leq k}$ is an i.i.d. centered sequence of random variables distributed as $\tau_1 - h$. Burkholder's and Hölder's inequalities, and finally item (ii) yield

$$\mathbb{E}|\tau_k - t_k|^p \leq C(p)\, \mathbb{E}\Big(\sum_{i=1}^{k} \eta_i^2\Big)^{\frac{p}{2}} \leq k^{\frac{p}{2}-1} \sum_{i=1}^{k} \mathbb{E}(\eta_i^p) \leq C(p)(t_k h)^{\frac{p}{2}},$$

which proves the claim for $p \geq 1$. The case $p < 1$ follows from this result by Jensen's inequality. □

Lemma A.2. *For all $t \in [0, T)$ and for all $\alpha < 1$, $\beta < 1$ we have*

$$\int_t^T \frac{1}{(T-r)^{\alpha}(r-t)^{\beta}} dr = \frac{1}{(T-t)^{\alpha+\beta-1}} B(1-\alpha, 1-\beta),$$

where B denotes the beta function.

References

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

[2] V. Bally. *Approximation scheme for solutions of BSDE.* Backward stochastic differential equations (N. El Karoui and L. Mazliak, eds), Pitmal Res, Notes Math. Ser., vol. 364, Longman, Harlow, 1997, pp. 177–191.

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

[4] A. Borodin and P. Salminen. *Handbook of Brownian Motion: Facts and Formulae.* Probability and its Applications. Birkhauser-Verlag, second edition, 2015.

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

[6] Ph. Briand, B. Delyon, Y. Hu, E. Pardoux, L. Stoica. *L_p solutions of backward stochastic differential equations.* Stochastic Process. Appl. 108: 109–129, 2003.

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

[8] 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.

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

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

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

[12] J.-F. Chassagneux, D. Crisan, and F. Delarue. *Numerical Method for FBSDEs of McKean-Vlasov type.* https://arxiv.org/abs/1703.02007, 2017.

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

[14] J.-F. Chassagneux and A. Richou. *Rate of convergence for discrete-time approximation of reflected BSDEs arising in switching problems.* https://arxiv.org/abs/1602.00015, 2016.

[15] 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.

[16] 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.

[17] D. Chevance. *Résolution numérique des équations différentielles stochastiques rétrogrades.* PhD. thesis, Université de Provence-Aix-Marseille I, Marseille, 1997.

[18] F. Coquet and V. Mackevičius and J. Mémin. Corrigendum to *Stability in* $\mathbf{D}$ *of martingales and backward equations under discretization of filtration.* [Stochastic Process. Appl., 75, 1998], Stochastic Process. Appl., 82(2), 335–338, 1999.

[19] 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.

[20] D. Crisan and F. Delarue. *Sharp derivative bounds for solutions of degenerate semi-linear partial differential equations.* J. Funct. Anal., 263(10):3024–3101, 2012.

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

[22] 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.

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

[24] 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.

[25] E. Gobet and C. Labart. *Error expansion for the discretization of backward stochastic differential equations.* Stochastic Process. Appl., 117(7):803–829, 2007.

[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] 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.

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

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

[30] 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.

[31] 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.

[32] P. Sundar, H. Yin. *Existence and uniqueness of solutions to the backward 2D stochastic Navier-Stokes equations*. Stochastic Process. Appl., 119(4):1216–1234, 2009.

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

[34] 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.* https://arxiv.org/abs/1708.03223, 2017.

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

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

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

[D]

L_2-APPROXIMATION RATE OF FORWARD-BACKWARD SDES USING RANDOM WALK

C. GEISS, C. LABART, AND A. LUOTO

L_2-Approximation rate of forward-backward SDEs using random walk

Christel Geiss[1], Céline Labart[2], Antti Luoto[3]

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

MSC codes : 60H10, 60H35, 60G50, 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 $(\mathcal{F}_t)_{t\geq 0}$ is the augmented natural filtration. Let (Y,Z) be the solution of the forward-backward SDE (FBSDE)

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

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

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

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

$$X_s^n = x + \int_{(0,s]} b(r,X_{r-}^n)d[B^n]_r + \int_{(0,s]} \sigma(r,X_{r-}^n)dB_r^n,$$

[1]Department of Mathematics and Statistics, P.O.Box 35 (MaD), FI-40014 University of Jyvaskyla, Finland
christel.geiss@jyu.fi

[2]Univ. Grenoble Alpes, Univ. Savoie Mont Blanc, CNRS, LAMA, 73000 Chambéry, France
celine.labart@univ-smb.fr

[3]Department of Mathematics and Statistics, P.O.Box 35 (MaD), FI-40014 University of Jyvaskyla, Finland
antti.k.luoto@student.jyu.fi

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

In this paper, we study the rate of the L_2-approximation of (Y_t^n, Z_t^n) to (Y_t, Z_t). This extends the results of [19] where this question was considered for the special case $X = B$.

The approximation of BSDEs using random walk has been investigated by many authors, also numerically (see, for example, [4], [23], [27], [29], [30], [31]). In 2001, Briand et al. [4] 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 [6] and Zhang in [36] 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 ([21], [15], [13]). Also forward methods have been introduced to approximate (1): a branching diffusion method ([24]), a multilevel Picard approximation ([35]) and Wiener chaos expansion ([5]). Many extensions of (1) have been considered, among them schemes for reflected BSDEs ([2], [12]), high order schemes ([9], [8]), fully-coupled BSDEs ([16], [7]), quadratic BSDEs ([11]), BSDEs with jumps ([20]) and McKean-Vlasov BSDEs ([1], [14], [10]).

In [19], under the assumption that the forward process X is the Brownian motion itself and given a locally α-Hölder continuous terminal function g and a Lipschitz continuous generator, a rate of convergence of order $h^{\frac{\alpha}{4}}$ was obtained for the L_2-norm of $Y_t^n - Y_t$, and for the L_2-norm of $Z_t^n - Z_t$ the rate of convergence is of order $\frac{h^{\frac{\alpha}{4}}}{\sqrt{T-t}}$. In the present paper, where we assume that X is a solution of the SDE in (1), we need rather strong conditions on the smoothness and boundedness on f and g and also on b and σ. In Theorem 3.2, 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.

One reason behind the strong smoothness requirements on the coefficients is that the discretized Malliavin derivative, which describes the relation between Y^n and Z^n, is not compatible with the variational equations related to Y^n and Z^n. This problem becomes visible in Subsection 2.3 where we introduce a discretized Malliavin weight to obtain a representation $\hat{Z}^n$ for Z^n. While the continuous-time representation of Z is exact, $\hat{Z}^n$ does not coincide with Z^n, but the difference converges to 0 in L_2 as $n \to \infty$. To prove our main result we also need strong smoothness conditions on the solution u^n of the difference equation associated to the discretized FBSDE (3). We sketch the proof by applying methods known for Lévy driven BSDEs.

The paper is organized as follows: Section 2 contains the setting, main assumptions and the approximative representation 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 of a Lipschitz continuous generator 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 numerical 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 $(B^n_{t_k})^n_{k=0}$ is $(\mathcal{G}_k)^n_{k=0}$-adapted. Recall (2) and $h = \frac{T}{n}$. If we extend the sequence $(X^n_{t_k})_{k\ge 0}$ 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) we do not use explicitely for our estimates but it is required for Theorem 4.2 below.

Assumption 2.1.

(i) $b, \sigma \in C^{0,2}_b([0, T] \times \mathbb{R})$, *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* $d((t, x), (\bar{t}, \bar{x})) = (|t-\bar{t}|+|x-\bar{x}|^2)^{\frac{1}{2}}$) *on all compact subsets of* $[0, T] \times \mathbb{R}$.

(iii) b, σ *are* $\frac{1}{2}$*-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

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

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

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

Notation:

- $\|\cdot\|_p := \|\cdot\|_{L^p(\mathbb{P})}$ for $p \ge 1$ 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.

2.2 The FBSDE and its numerical scheme

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

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

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 $r \in [t_m, t_{m+1})$.

For n large enough, the BSDE (3) has a unique solution (Y^n, Z^n) (see [32, Proposition 1.2]), and $(Y^n_{t_m}, Z^n_{t_m})_{m=0}^{n-1}$ is adapted to the filtration $(\mathcal{G}_m)_{m=0}^{n-1}$.

2.3 Representations for Z and Z^n

We will use the representation (see Ma and Zhang [28, Theorem 4.2])

$$Z_t = \mathbb{E}_t\left(g'(X_T)\nabla X_T + \int_t^T f(s, X_s, Y_s, Z_s)N^t_s ds\right)\sigma(t, X_t), \quad 0 \le t \le T \tag{9}$$

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

$$N^t_s = \frac{1}{s-t}\int_t^s \frac{\nabla X_r}{\sigma(r, X_r)\nabla X_t} dB_r, \tag{10}$$

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

$$\nabla X_s = 1 + \int_0^s b_x(r, X_r)\nabla X_r dr + \int_0^s \sigma_x(r, X_r)\nabla X_r dB_r, \tag{11}$$

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

2.3.1 Approximation for Z^n

A counterpart to (9) for Z^n does in general only exist approximatively. In particular for $f \neq 0$ stronger smoothness assumptions are required:

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).*

We shortly introduce the discretized Malliavin derivative and refer the reader to [3] for more information on this topic. We first define for any function $F : \{-1, 1\}^n \to \mathbb{R}$ the mappings $T_{m,+}$ and $T_{m,-}$ by

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

and for any $\xi = F(\varepsilon_1, \dots, \varepsilon_n)$ the discretized Malliavin derivative

$$\mathcal{D}^n_m \xi := \frac{\mathbb{E}[\xi\varepsilon_m|\sigma((\varepsilon_l)_{l\in\{1,\dots,n\}\setminus\{m\}})]}{\sqrt{h}} = \frac{T_{m,+}\xi - T_{m,-}\xi}{2\sqrt{h}}, \quad 1 \le m \le n. \tag{12}$$

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 processes

$$\begin{aligned} \nabla X^{n,t_k,x}_{t_m} &= 1 + h \sum_{l=k+1}^{m} b_x(t_l, X^{n,t_k,x}_{t_{l-1}}) \nabla X^{n,t_k,x}_{t_{l-1}} + \sqrt{h} \sum_{l=k+1}^{m} \sigma_x(t_l, X^{n,t_k,x}_{t_{l-1}}) \nabla X^{n,t_k,x}_{t_{l-1}} \varepsilon_l, \;\; 0 \le k \le m \le n, \\ \mathcal{D}^n_k X^n_{t_m} &= \sigma(t_k, X^n_{t_{k-1}}) + h \sum_{l=k+1}^{m} b_x^{(k,l)} \mathcal{D}^n_k X^n_{t_{l-1}} + \sqrt{h} \sum_{l=k+1}^{m} \sigma_x^{(k,l)} (\mathcal{D}^n_k X^n_{t_{l-1}}) \varepsilon_l, \quad 0 < k \le m \le n, \end{aligned} \tag{13}$$

where for the latter we use for $\phi = b$ and $\phi = \sigma$ the notation (if $\mathcal{D}^n_k X^n_{t_{\ell-1}} \neq 0$ the second $' :=' $ holds as an identity)

$$\phi_x^{(k,l)} := \frac{\mathcal{D}^n_k \phi(t_l, X^n_{t_{l-1}})}{\mathcal{D}^n_k X^n_{t_{l-1}}} := \int_0^1 \phi_x(t_l, \vartheta T_{k,+} X^n_{t_{l-1}} + (1-\vartheta) T_{k,-} X^n_{t_{l-1}}) d\vartheta. \tag{14}$$

However, we can show convergence of $\nabla X^{n,t_k,X^n_{t_k}}_{t_m} - \frac{\mathcal{D}^n_{k+1} X^n_{t_m}}{\sigma(t_{k+1}, X^n_{t_k})}$ in L_p.

Lemma 2.4. *Under Assumption 2.1, and for $p \ge 2$, we have*

(i) $\mathbb{E}|X^n_{t_l} - T_{m,\pm} X^n_{t_l}|^p \le C(b,\sigma,T,p) h^{\frac{p}{2}}, \quad 1 \le l, m \le n,$

(ii) $\mathbb{E}\left| \nabla X^{n,t_k,X^n_{t_k}}_{t_m} - \frac{\mathcal{D}^n_{k+1} X^n_{t_m}}{\sigma(t_{k+1}, X^n_{t_k})} \right|^p \le C(b,\sigma,T,p) h^{\frac{p}{2}}, \quad 0 \le k < m \le n.$

(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

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

It remains to apply Gronwall's lemma.

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

$$\mathbb{E}\left|\nabla X_{t_m}^{n,t_k,X_{t_k}^n} - \frac{\mathcal{D}_{k+1}^n X_{t_m}^n}{\sigma(t_{k+1},X_{t_k}^n)}\right|^p \le C(p,T)\Bigg(|b_x(t_{k+1},X_{t_k}^n)h + \sigma_x(t_{k+1},X_{t_k}^n)\sqrt{h}\varepsilon_{k+1}|^p$$
$$+ h\sum_{l=k+2}^{m} \mathbb{E}\left|b_x(t_l,X_{t_{l-1}}^n)\nabla X_{t_{l-1}}^{n,t_k,X_{t_k}^n} - b_x^{(k+1,l)}\frac{\mathcal{D}_{k+1}^n X_{t_{l-1}}^n}{\sigma(t_{k+1},X_{t_k}^n)}\right|^p$$
$$+ h\sum_{l=k+2}^{m} \mathbb{E}\left|\sigma_x(t_l,X_{t_{l-1}}^n)\nabla X_{t_{l-1}}^{n,t_k,X_{t_k}^n} - \sigma_x^{(k+1,l)}\frac{\mathcal{D}_{k+1}^n X_{t_{l-1}}^n}{\sigma(t_{k+1},X_{t_k}^n)}\right|^p\Bigg).$$

Since by Lemma 2.4 (i) we conclude that

$$\mathbb{E}|b_x^{(k+1,l)} - b_x(t_l,X_{t_{l-1}}^n)|^{2p} + \mathbb{E}|\sigma_x^{(k+1,l)} - \sigma_x(t_l,X_{t_{l-1}}^n)|^{2p} \le C(b,\sigma,T,p)h^p,$$

and Lemma 5.2 implies that

$$\mathbb{E}\sup_{k+1\le l\le m}\left|\nabla X_{t_{l-1}}^{n,t_k,X_{t_k}^n}\right|^{2p} \le C(b,\sigma,T,p),$$

the assertion follows by Gronwall's lemma.

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

□

We introduce a discrete counterpart to the Malliavin weight given in (10) letting

$$N_{t_\ell}^{n,t_k} := \sqrt{h}\sum_{m=k+1}^{\ell}\frac{\nabla X_{t_{m-1}}^{n,t_k,X_{t_k}^n}}{\sigma(t_m,X_{t_{m-1}}^n)}\frac{\varepsilon_m}{t_\ell - t_k}, \quad k<\ell\le n. \tag{15}$$

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

$$\left(\mathbb{E}_k|N_{t_\ell}^{n,t_k}|^2\right)^{\frac{1}{2}} \le \frac{\widehat{\kappa}_2}{(t_\ell - t_k)^{\frac{1}{2}}}, \quad 0\le k<\ell\le n, \tag{16}$$

where $\mathbb{E}_k := \mathbb{E}(\cdot|\mathcal{G}_k)$. We define a process $\hat{Z}^n = (\hat{Z}_{t_k}^n)_{k=0}^{n-1}$ by

$$\hat{Z}_{t_k}^n := \mathbb{E}_k\left(\mathcal{D}_{k+1}^n g(X_T^n)\right) + \mathbb{E}_k\left(h\sum_{m=k+1}^{n-1} f(t_{m+1},X_{t_m}^n,Y_{t_m}^n,Z_{t_m}^n)N_{t_m}^{n,t_k}\right)\sigma(t_{k+1},X_{t_k}^n), \tag{17}$$

and compare it with $Z^n = (Z_{t_k}^n)_{k=0}^{n-1}$ given by

$$Z_{t_k}^n = \mathbb{E}_k\left(\mathcal{D}_{k+1}^n g(X_T^n)\right) + \mathbb{E}_k\left(\sqrt{h}\sum_{m=k+1}^{n-1} f(t_{m+1},X_{t_m}^n,Y_{t_m}^n,Z_{t_m}^n)\varepsilon_{k+1}\right). \tag{18}$$

The latter equation follows if one multiplies (8) by ε_{k+1} and takes the conditional expectation w.r.t. $\mathcal{G}_k$. In (17) we could have used also the approximate expression $\mathbb{E}_k(g(X_T^n)N_{t_n}^{n,t_k}\sigma(t_{k+1},X_{t_k}^n))$, but since we will assume that g'' exists, we work with the correct term.

Proposition 2.5. *If Assumption 2.3 holds, then*

$$\mathbb{E}_{0,x}|Z^n_{t_k} - \hat{Z}^n_{t_k}|^2 \le C_{2.5}\hat{\Psi}^2(x)h^\alpha,$$

where $\mathbb{E}_{0,x} := \mathbb{E}(\cdot|X_0 = x)$*, the function* $\hat{\Psi}$ *is defined in* (59) *below, and* $C_{2.5}$ *depends on* $b, \sigma, f, g, T,$ p_0 *and* δ*.*

Proof. According to [4, Proposition 5.1] one has the representations

$$Y^n_{t_m} = u^n(t_m, X^n_{t_m}), \quad \text{and} \quad Z^n_{t_m} = \mathcal{D}^n_{m+1}u^n(t_{m+1}, X^n_{t_{m+1}}), \tag{19}$$

where u^n is the solution of the 'discretised' PDE (41) with terminal condition $u^n(t_n, x) = g(x)$. Notice that by the definition of $\mathcal{D}^n_{m+1}$ in (12) the expression $\mathcal{D}^n_{m+1}u^n(t_{m+1}, X^n_{t_{m+1}})$ depends in fact on $X^n_{t_m}$. Hence we can put

$$\begin{aligned} f(t_{m+1}, X^n_{t_m}, Y^n_{t_m}, Z^n_{t_m}) &= f(t_{m+1}, X^n_{t_m}, u^n(t_m, X^n_{t_m}), \mathcal{D}^n_{m+1}u^n(t_{m+1}, X^n_{t_{m+1}})) \\ &=: F^n(t_{m+1}, X^n_{t_m}). \end{aligned}$$

By Proposition 4.5 we conclude that $u^n_x(t_m, x)$ (as function of x, uniformly in time) satisfies (5) with p_0 replaced by $2p_0 + 2$, and the function $x \mapsto \partial_x \mathcal{D}^n_{m+1}u^n(t_{m+1}, X^{n,t_m,x}_{t_{m+1}})$ satisfies the local α-Hölder continuity relation (58). By Assumption 2.3 on f we derive the latter bound also for $x \mapsto F^n_x(t_{m+1}, x)$. From (17), (18) and (12) we conclude that (we use $\mathbb{E} := \mathbb{E}_{0,x}$)

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

With the notation introduced in (14) applied to F^n,

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

For A_1 we use (14) again and exploit the fact that $x \mapsto F^n_x(t, x)$ is locally α-Hölder continuous. By Hölder's inequality and Lemma 2.4 (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 $F_x^{n,(\ell,m+1)}$ is bounded by $C\Psi^2(x)$, so that it suffices to estimate

$$\begin{aligned}
&\left\|(\mathcal{D}^n_{k+1}X^n_{t_m}) - (\mathcal{D}^n_\ell X^n_{t_m})\frac{\sigma(t_{k+1},X^n_{t_k})\nabla X^{n,t_k,X^n_{t_k}}_{t_{\ell-1}}}{\sigma(t_\ell,X^n_{t_{\ell-1}})}\right\|_4\\
&\le \left\|(\mathcal{D}^n_{k+1}X^n_{t_m}) - \frac{\sigma(t_{k+1},X^n_{t_k})\,\mathcal{D}^n_\ell X^n_{t_m}}{\sigma(t_\ell,X^n_{t_{\ell-1}})}\frac{\mathcal{D}^n_{k+1}X^n_{t_{\ell-1}}}{\sigma(t_{k+1},X^n_{t_k})}\right\|_4\\
&\qquad+\left\|\frac{\sigma(t_{k+1},X^n_{t_k})\,\mathcal{D}^n_\ell X^n_{t_m}}{\sigma(t_\ell,X^n_{t_{\ell-1}})}\left(\nabla X^{n,t_k,X^n_{t_k}}_{t_{\ell-1}} - \frac{\mathcal{D}^n_{k+1}X^n_{t_{\ell-1}}}{\sigma(t_{k+1},X^n_{t_k})}\right)\right\|_4 .
\end{aligned}\tag{20}$$

The second expression on the r.h.s. of (20) is bounded by $C(b,\sigma,T,\delta)h^{\frac{1}{2}}$ as a consequence of Lemma 2.4 (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 (13) and get

$$\begin{aligned}
&\left|\frac{\mathcal{D}^n_\ell X^n_{t_m}}{\sigma(t_\ell,X^n_{t_{\ell-1}})}\mathcal{D}^n_{k+1}X^n_{t_{\ell-1}} - \mathcal{D}^n_{k+1}X^n_{t_m}\right|\\
&=\Bigg|\left(1+\sum_{l=\ell+1}^{m}\frac{\mathcal{D}^n_\ell X^n_{t_{l-1}}}{\sigma(t_\ell,X^n_{t_{\ell-1}})}(b_x^{(\ell,l)}h+\sigma_x^{(\ell,l)}\sqrt{h}\varepsilon_l)\right)\\
&\qquad\times\left(\sigma(t_{k+1},X^n_{t_k})+\sum_{l=k+2}^{\ell-1}\mathcal{D}^n_{k+1}X^n_{t_{l-1}}(b_x^{(k+1,l)}h+\sigma_x^{(k+1,l)}\sqrt{h}\varepsilon_l)\right)\\
&\qquad-\left(\sigma(t_{k+1},X^n_{t_k})+\left(\sum_{l=k+2}^{\ell-1}+\sum_{l=\ell}^{m}\right)\mathcal{D}^n_{k+1}X^n_{t_{l-1}}(b_x^{(k+1,l)}h+\sigma_x^{(k+1,l)}\sqrt{h}\varepsilon_l)\right)\Bigg|\\
&\le|\mathcal{D}^n_{k+1}X^n_{t_{\ell-1}}(b_x^{(k+1,\ell)}h+\sigma_x^{(k+1,\ell)}\sqrt{h}\varepsilon_\ell)|\\
&\quad+\left|\sum_{l=\ell+1}^{m}\left[\frac{\mathcal{D}^n_\ell X^n_{t_{l-1}}}{\sigma(t_\ell,X^n_{t_{\ell-1}})}\mathcal{D}^n_{k+1}X^n_{t_{\ell-1}}-\mathcal{D}^n_{k+1}X^n_{t_{l-1}}\right]\left(b_x^{(\ell,l)}h+\sigma_x^{(\ell,l)}\sqrt{h}\varepsilon_l\right)\right|\\
&\quad+\left|\sum_{l=\ell+1}^{m}\mathcal{D}^n_{k+1}X^n_{t_{l-1}}\left[b_x^{(\ell,l)}h+\sigma_x^{(\ell,l)}\sqrt{h}\varepsilon_l-\left(b_x^{(k+1,l)}h+\sigma_x^{(k+1,l)}\sqrt{h}\varepsilon_l\right)\right]\right|.
\end{aligned}\tag{21}$$

We take the L_4-norm of (21) and apply the BDG inequality and Hölder's inequality. The second term on the r.h.s. of (21) 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 2.4-(iii). For the last term we also use the Lipschitz continuity of b_x and σ_x in space and Lemma 2.4-(i). □

3 Main results

The following approximations will rely on the fact that the random walk B^n can be constructed from the Brownian motion B by Skorohod embedding. Let $\tau_0 := 0$ and define

$$\tau_k := \inf\{t>\tau_{k-1} : |B_t - B_{\tau_{k-1}}| = \sqrt{h}\},\quad k\ge 1.\tag{22}$$

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 $\sqrt{h}\varepsilon_k \overset{d}{=} B_{\tau_k} - B_{\tau_{k-1}}$. In this case we also use the notation $\mathcal{X}_{\tau_k} := X^n_{t_k}$ for all $k = 0, \ldots, n$, so that (4) turns into

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

and (3) holds for B^n given by

$$B^n_t = \sum_{k=1}^{[t/h]} (B_{\tau_k} - B_{\tau_{k-1}}), \qquad 0 \le t \le T. \tag{23}$$

We will denote by $\mathbb{E}_{\tau_k}$ the conditional expectation w.r.t. $\mathcal{F}_{\tau_k}$.

3.1 Approximation rates for the zero generator case

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

$$\tilde{\mathcal{X}}^{\tau_k, \mathcal{X}_{\tau_k}}_{\tau_n} = \mathcal{X}_{\tau_k} + \sum_{j=k+1}^{n} b(t_j, \tilde{\mathcal{X}}^{\tau_k, \mathcal{X}_{\tau_k}}_{\tau_{j-1}})h + \sum_{j=k+1}^{n} \sigma(t_j, \tilde{\mathcal{X}}^{\tau_k, \mathcal{X}_{\tau_k}}_{\tau_{j-1}})(\tilde{B}_{\tilde{\tau}_{j-k}} - \tilde{B}_{\tilde{\tau}_{j-k-1}}), \tag{24}$$

(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 $j \ge 1$ and $\tau_n := \tau_k + \tilde{\tau}_{n-k}$ for $n \ge k$). 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

$$B'_t = B_{t\wedge t_k} + \tilde{B}_{(t-t_k)^+} \quad \text{and} \quad B''_t = B_{t\wedge \tau_k} + \tilde{B}_{(t-\tau_k)^+}, \quad t \ge 0. \tag{25}$$

Proposition 3.1. *Let Assumption 2.1 and (23) 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*

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

where $C^y_{3.1} = C(T, p_0, C_g, C^y_{4.2}, \sigma, b)$ and $C^z_{3.1} = C(T, p_0, C_{g'}, \sigma, b, \delta)$.

Proof. 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 $Y^n_v = Y^n_{t_k}$. Then

$$\mathbb{E}|Y_v - Y^n_v|^2 \le 2(\mathbb{E}|Y_v - Y_{t_k}|^2 + \mathbb{E}|Y_{t_k} - Y^n_{t_k}|^2).$$

Using Theorem 4.2 we bound $\|Y_v - Y_{t_k}\|$ by $C^y_{4.2}\Psi(x)(v - t_k)^{\frac{1}{2}}$ (since $\alpha = 1$ can be chosen when g is locally Lipschitz continuous). It remains to bound

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

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

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

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

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

Let us then deal with the error of Z. We use $\|Z_v - Z^n_v\| \le \|Z_v - Z_{t_k}\| + \|Z_{t_k} - Z^n_{t_k}\|$ 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 4.3), where

$$\nabla\tilde{X}^{t,x}_s = 1 + \int_t^s b_x(r, \tilde{X}^{t,x}_r)\nabla\tilde{X}^{t,x}_r dr + \int_t^s \sigma_x(r, \tilde{X}^{t,x}_r)\nabla\tilde{X}^{t,x}_r d\tilde{B}_{r-t}, \quad 0 \le t \le s \le T.$$

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

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

We compute the second term using $Z^n_{t_k}$ as given in (18). Hence, with the notation from (14),

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

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

$$\begin{aligned}
&\left|\tilde{\mathbb{E}}(g'(\tilde{X}^{t_k,X_{t_k}}_{t_n})\nabla\tilde{X}^{t_k,X_{t_k}}_{t_n}) - \tilde{\mathbb{E}}\left(g_x^{(k+1,n+1)}\frac{\mathcal{D}^n_{k+1}\tilde{\mathcal{X}}^{\tau_k,\mathcal{X}_{\tau_k}}_{\tau_n}}{\sigma(t_k, X_{t_k})}\right)\right|^2 \\
&\le 2\tilde{\mathbb{E}}|g'(\tilde{X}^{t_k,X_{t_k}}_{t_n}) - g_x^{(k+1,n+1)}|^2\tilde{\mathbb{E}}|\nabla\tilde{X}^{t_k,X_{t_k}}_{t_n}|^2 + 2\tilde{\mathbb{E}}|g_x^{(k+1,n+1)}|^2\tilde{\mathbb{E}}\left|\nabla\tilde{X}^{t_k,X_{t_k}}_{t_n} - \frac{\mathcal{D}^n_{k+1}\tilde{\mathcal{X}}^{\tau_k,\mathcal{X}_{\tau_k}}_{\tau_n}}{\sigma(t_k, X_{t_k})}\right|^2.
\end{aligned} \tag{26}$$

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

$$\tilde{\mathbb{E}}|g'(\tilde{X}^{t_k,X_{t_k}}_{t_n}) - g_x^{(k+1,n+1)}|^2$$

$$\leq \int_0^1 \tilde{\mathbb{E}}\Big|g'(\tilde{X}_{t_n}^{t_k,X_{t_k}}) - g'(\vartheta T_{k+1,+}\tilde{\mathcal{X}}_{\tau_n}^{\tau_k,\mathcal{X}_{\tau_k}} + (1-\vartheta)T_{k+1,-}\tilde{\mathcal{X}}_{\tau_n}^{\tau_k,\mathcal{X}_{\tau_k}})\Big|^2 d\vartheta$$

$$\leq \int_0^1 (\tilde{\mathbb{E}}\Psi_1^4)^{\frac{1}{2}}\left[\tilde{\mathbb{E}}\left|\tilde{X}_{t_n}^{t_k,X_{t_k}} - \vartheta T_{k+1,+}\tilde{\mathcal{X}}_{\tau_n}^{\tau_k,\mathcal{X}_{\tau_k}} - (1-\vartheta)T_{k+1,-}\tilde{\mathcal{X}}_{\tau_n}^{\tau_k,\mathcal{X}_{\tau_k}}\right|^{4\alpha}\right]^{\frac{1}{2}} d\vartheta,$$

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

$$\tilde{\mathbb{E}}\left|\tilde{X}_{t_n}^{t_k,X_{t_k}} - (\vartheta T_{k+1,+}\tilde{\mathcal{X}}_{\tau_n}^{\tau_k,\mathcal{X}_{\tau_k}} + (1-\vartheta)T_{k+1,-}\tilde{\mathcal{X}}_{\tau_n}^{\tau_k,\mathcal{X}_{\tau_k}})\right|^{4\alpha}$$

$$\leq 8\left(\vartheta^{2\alpha}\tilde{\mathbb{E}}\left|\tilde{X}_{t_n}^{t_k,X_{t_k}} - T_{k+1,+}\tilde{\mathcal{X}}_{\tau_n}^{\tau_k,\mathcal{X}_{\tau_k}}\right|^{4\alpha} + (1-\vartheta)^{2\alpha}\tilde{\mathbb{E}}\left|\tilde{X}_{t_n}^{t_k,X_{t_k}} - T_{k+1,-}\tilde{\mathcal{X}}_{\tau_n}^{\tau_k,\mathcal{X}_{\tau_k}}\right|^{4\alpha}\right)$$

$$\leq C(b,\sigma,T)h^{2\alpha} + C(b,\sigma,T)(|X_{t_k} - \mathcal{X}_{\tau_k}|^{4\alpha} + h^{\alpha}),$$

we use Lemma 2.4 and Lemma 5.2-(v). For the last term in (26) we notice that

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

By Lemma 5.2 we have $\mathbb{E}\tilde{\mathbb{E}}|\nabla\tilde{X}_{t_n}^{t_k,X_{t_k}} - \nabla\tilde{\mathcal{X}}_{\tau_n}^{\tau_k,\mathcal{X}_{\tau_k}}|^p \leq C(b,\sigma,T,p)h^{\frac{p}{4}}$, and by Lemma 2.4,

$$\mathbb{E}\tilde{\mathbb{E}}\left|\nabla\tilde{\mathcal{X}}_{\tau_n}^{\tau_k,\mathcal{X}_{\tau_k}} - \frac{\mathcal{D}_{k+1}^n\tilde{\mathcal{X}}_{\tau_n}^{\tau_k,\mathcal{X}_{\tau_k}}}{\sigma(t_k,X_{t_k})}\right|^p$$

$$\leq C(p)\mathbb{E}\left|\nabla X_{t_n}^{n,t_k,X_{t_k}^n} - \frac{\mathcal{D}_{k+1}^n X_{t_n}^n}{\sigma(t_{k+1},X_{t_k}^n)}\right|^p + C(p)\mathbb{E}\left|\frac{\mathcal{D}_{k+1}^n X_{t_n}^n}{\sigma(t_{k+1},X_{t_k}^n)} - \frac{\mathcal{D}_{k+1}^n X_{t_n}^n}{\sigma(t_k,X_{t_k})}\right|^p \leq C(b,\sigma,T,p,\delta)h^{\frac{p}{4}}.$$

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

3.2 Approximation rates for the general case

Theorem 3.2. *Let Assumptions 2.3 be satisfied and B^n be given by (23). Then for all $v \in [0,T)$ and large enough n, we have*

$$\mathbb{E}_{0,x}|Y_v - Y_v^n|^2 + \mathbb{E}_{0,x}|Z_v - Z_v^n|^2 \leq C_{3.2}\hat{\Psi}^2(x)h^{\frac{1}{2}\wedge\alpha}$$

with $C_{3.2} = C(b,\sigma,f,g,T,p_0,\delta,\kappa_2,c_{4.3}^{2,3},C_{4.2}^y,C_{4.4})$ and $\hat{\Psi}$ is given in (59).

Proof. Let $u : [0,T)\times\mathbb{R} \to \mathbb{R}$ be the solution of the PDE (35) 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 4.3 and define

$$F(s,x) := f(s,x,u(s,x),\sigma(s,x)u_x(s,x)). \tag{27}$$

From (1) and (3) we conclude

$$\|Y_{t_k} - Y_{t_k}^n\| \leq \|\mathbb{E}_{t_k}g(X_T) - \mathbb{E}_{\tau_k}g(X_T^n)\| + \left\|\mathbb{E}_{t_k}\int_{t_k}^T f(s,X_s,Y_s,Z_s)ds - h\mathbb{E}_{\tau_k}\sum_{m=k}^{n-1} f(t_{m+1},X_{t_m}^n,Y_{t_m}^n,Z_{t_m}^n)\right\|,$$

where Proposition 3.1 provides the estimate for the terminal condition. The generator terms we decompose as follows:

$$\mathbb{E}_{t_k}f(s,X_s,Y_s,Z_s) - \mathbb{E}_{\tau_k}f(t_{m+1},X_{t_m}^n,Y_{t_m}^n,Z_{t_m}^n)$$

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

We use

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

and estimate the expressions on the right hand side. For the function F defined in (27) we use Assumption 2.3 (which implies that (5) holds for $\alpha = 1$) to derive by Theorem 4.3 and the mean value theorem that for $x_1, x_2 \in \mathbb{R}$ there exist $\xi_1, \xi_2 \in [\min\{x_1,x_2\}, \max\{x_1,x_2\}]$ such that

$$
\begin{aligned}
|F(t,x_1) - F(t,x_2)| &= |f(t,x_1,u(t,x_1),\sigma(t,x_1)u_x(t,x_1)) - f(t,x_2,u(t,x_2),\sigma(t,x_2)u_x(t,x_2))| \\
&\le C(L_f,\sigma)\left(1 + c^2_{4.3}\Psi(\xi_1) + \frac{c^3_{4.3}\Psi(\xi_2)}{(T-t)^{\frac{1}{2}}}\right)|x_1 - x_2| \\
&\le C(T,L_f,\sigma,c^{2,3}_{4.3})(1 + |x_1|^{p_0+1} + |x_2|^{p_0+1})\frac{|x_1-x_2|}{(T-t)^{\frac{1}{2}}}. \qquad (28)
\end{aligned}
$$

By (6), standard estimates on (X_s), Theorem 4.2-(i) and Proposition 4.4 we immediately get

$$
\|d_1(s,m)\| \le C(L_f,b,\sigma,T,C^y_{4.2},C_{4.4})\Psi(x)\,h^{\frac{1}{2}}.
$$

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 (28) and Lemma 5.2-(v). This gives

$$
\|d_2(m)\| \le C(L_f,c^{2,3}_{4.3},b,\sigma,T,p_0)\Psi(x)\frac{1}{(T-t_m)^{\frac{1}{2}}}h^{\frac{1}{4}}.
$$

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

$$
\|d_3(m)\| \le \|F(t_m,X^n_{t_m}) - F(t_m,X_{t_m})\| \le C(L_f,c^{2,3}_{4.3},b,\sigma,T,p_0)\Psi(x)\frac{1}{(T-t_m)^{\frac{1}{2}}}h^{\frac{1}{4}},
$$

and for the last term we get

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

This implies

$$
\|Y_{t_k} - Y^n_{t_k}\| \le C\Psi(x)h^{\frac{1}{4}} + hL_f\sum_{m=k}^{n-1}(\|Y_{t_m} - Y^n_{t_m}\| + \|Z_{t_m} - Z^n_{t_m}\|), \qquad (29)
$$

where $C = C(L_f,C^y_{3.1},c^{2,3}_{4.3},C^y_{4.2},C_{4.4},b,\sigma,T,p_0)$.

For $\|Z_{t_k} - Z^n_{t_k}\|$ we use the representations (9), (18) and the approximation (17) as well as Proposition 2.5. 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

$$\begin{aligned}
\|Z^n_{t_k} - Z_{t_k}\| &\leq \|Z^n_{t_k} - \hat{Z}^n_{t_k}\| + \|Z_{t_k} - \hat{Z}^n_{t_k}\| \\
&\leq C_{2.5}\hat{\Psi}(x)h^{\frac{\alpha}{2}} + \|\sigma(t_k, X_{t_k})\tilde{\mathbb{E}}g'(\tilde{X}^{t_k,X_{t_k}}_{t_n})\nabla\tilde{X}^{t_k,X_{t_k}}_{t_n} - \tilde{\mathbb{E}}\mathcal{D}^n_{k+1}g(\tilde{X}^{n,t_k,X^n_{t_k}}_{t_n})\| \\
&\quad + \Big\|\mathbb{E}_{t_k}\int_{t_{k+1}}^T f(s, X_s, Y_s, Z_s)N^{t_k}_s ds\, \sigma(t_k, X_{t_k}) \\
&\qquad - \mathbb{E}_{\tau_k} h \sum_{m=k+1}^{n-1} f(t_{m+1}, X^n_{t_m}, Y^n_{t_m}, Z^n_{t_m})N^{n,\tau_k}_{\tau_m}\sigma(t_{k+1}, X^n_{t_k})\Big\| \\
&\quad + \Big\|\mathbb{E}_{t_k}\int_{t_k}^{t_{k+1}} f(s, X_s, Y_s, Z_s)N^{t_k}_s ds\, \sigma(t_k, X_{t_k})\Big\|. \qquad (30)
\end{aligned}$$

For the terminal condition Proposition 3.1 provides

$$\|\sigma(t_k, X_{t_k})\tilde{\mathbb{E}}g'(\tilde{X}^{t_k,X_{t_k}}_{t_n})\nabla\tilde{X}^{t_k,X_{t_k}}_{t_n} - \tilde{\mathbb{E}}\mathcal{D}^n_{k+1}g(\tilde{X}^{n,t_k,X^n_{t_k}}_{t_n})\| \leq (C^z_{3.1})^{\frac{1}{2}}\Psi(x)h^{\frac{1}{4}}. \qquad (31)$$

We continue with the generator terms and use F defined in (27) to decompose the difference

$$\begin{aligned}
&\mathbb{E}_{t_k}f(s, X_s, Y_s, Z_s)N^{t_k}_s\sigma(t_k, X_{t_k}) - \mathbb{E}_{\tau_k}f(t_{m+1}, X^n_{t_m}, Y^n_{t_m}, Z^n_{t_m})N^{n,\tau_k}_{\tau_m}\sigma(t_{k+1}, X^n_{t_k}) \\
&= \mathbb{E}_{t_k}f(s, X_s, Y_s, Z_s)N^{t_k}_s\sigma(t_k, X_{t_k}) - \mathbb{E}_{t_k}f(t_m, X_{t_m}, Y_{t_m}, Z_{t_m})N^{t_k}_{t_m}\sigma(t_k, X_{t_k}) \\
&\quad + \mathbb{E}_{t_k}F(t_m, X_{t_m})N^{t_k}_{t_m}\sigma(t_k, X_{t_k}) - \mathbb{E}_{\tau_k}F(t_m, X^n_{t_m})N^{n,\tau_k}_{\tau_m}\sigma(t_{k+1}, X^n_{t_k}) \\
&\quad + \mathbb{E}_{\tau_k}\Big[[F(t_m, X^n_{t_m}) - F(t_m, X_{t_m})]N^{n,\tau_k}_{\tau_m}\sigma(t_{k+1}, X^n_{t_k})\Big] \\
&\quad + \mathbb{E}_{\tau_k}\Big[[f(t_m, X_{t_m}, Y_{t_m}, Z_{t_m}) - f(t_{m+1}, X^n_{t_m}, Y^n_{t_m}, Z^n_{t_m})]N^{n,\tau_k}_{\tau_m}\sigma(t_{k+1}, X^n_{t_k})\Big] \\
&=: t_1(s,m) + t_2(m) + t_3(m) + t_4(m)
\end{aligned}$$

where $s \in [t_m, t_{m+1})$. For t_1 we use that $\mathbb{E}_{t_k}f(t_m, X_{t_k}, Y_{t_k}, Z_{t_k})(N^{t_k}_s - N^{t_k}_{t_m}) = 0$, so that

$$\begin{aligned}
\|t_1(s,m)\| &\leq \|\mathbb{E}_{t_k}f(s, X_s, Y_s, Z_s)N^{t_k}_s\sigma(t_k, X_{t_k}) - \mathbb{E}_{t_k}f(t_m, X_{t_m}, Y_{t_m}, Z_{t_m})N^{t_k}_s\sigma(t_k, X_{t_k})\| \\
&\quad + \|\mathbb{E}_{t_k}(f(t_m, X_{t_m}, Y_{t_m}, Z_{t_m}) - f(t_m, X_{t_k}, Y_{t_k}, Z_{t_k}))(N^{t_k}_s - N^{t_k}_{t_m})\sigma(t_k, X_{t_k})\|.
\end{aligned}$$

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

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

and

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

We apply the conditional Hölder inequality, and from the estimates (34) and $\tilde{\mathbb{E}}|\tilde{N}^{t_k}_s - \tilde{N}^{t_k}_{t_m}|^2 \leq C(b, \sigma, T, \delta)\frac{h}{(s-t_k)^2}$ we get

$$\|t_1(s,m)\| \leq \frac{\kappa_2\|\sigma\|_\infty}{(s-t_k)^{\frac{1}{2}}}\|f(s, X_s, Y_s, Z_s) - f(t_m, X_{t_m}, Y_{t_m}, Z_{t_m})\|$$

$$
\begin{aligned}
&+C(b,\sigma,T,\delta)\frac{h^{\frac{1}{2}}}{s-t_k}\ \|f(t_m,X_{t_m},Y_{t_m},Z_{t_m})-f(t_k,X_{t_k},Y_{t_k},Z_{t_k})\| \\
\leq\ & C(L_f,C^y_{4.2},C_{4.4},\kappa_2,b,\sigma,T,\delta,p_0)\Psi(x)\frac{h^{\frac{1}{2}}}{(s-t_k)^{\frac{1}{2}}},
\end{aligned}
$$

since for $0\leq t<s\leq T$ we have by Theorem 4.2 and Proposition 4.4 that

$$
\|f(s,X_s,Y_s,Z_s)-f(t,X_t,Y_t,Z_t)\| \ \leq\ C(L_f,C^y_{4.2},C_{4.4},b,\sigma,T,p_0)\Psi(x)(s-t)^{\frac{1}{2}}. \tag{32}
$$

For the estimate of t_2 Lemma 5.2, Lemma 5.3, (28) and (34) yield

$$
\begin{aligned}
\|t_2(m)\| &= \|\tilde{\mathbb{E}}F(t_m,\tilde{X}^{t_k,X_{t_k}}_{t_m})\tilde{N}^{t_k}_{t_m}\sigma(t_k,X_{t_k})-\tilde{\mathbb{E}}F(t_m,\tilde{\mathcal{X}}^{\tau_k,\mathcal{X}_{\tau_k}}_{\tau_m})\tilde{N}^{n,\tau_k}_{\tau_m}\sigma(t_{k+1},\mathcal{X}_{\tau_k})\| \\
&\leq \frac{C(\sigma,\kappa_2)}{(t_m-t_k)^{\frac{1}{2}}}\left(\mathbb{E}\tilde{\mathbb{E}}(F(t_m,\tilde{X}^{t_k,X_{t_k}}_{t_m})-F(t_m,\tilde{\mathcal{X}}^{\tau_k,\mathcal{X}_{\tau_k}}_{\tau_m}))^2\right)^{\frac{1}{2}} \\
&\quad +(\mathbb{E}\tilde{\mathbb{E}}|F(t_m,\tilde{\mathcal{X}}^{\tau_k,\mathcal{X}_{\tau_k}}_{\tau_m})-F(t_m,\mathcal{X}_{\tau_k})|^2\tilde{\mathbb{E}}|\tilde{N}^{t_k}_{t_m}\sigma(t_k,X_{t_k})-\tilde{N}^{n,\tau_k}_{\tau_m}\sigma(t_{k+1},\mathcal{X}_{\tau_k})|^2)^{\frac{1}{2}} \\
&\leq C(L_f,c^{2,3}_{4.3},b,\sigma,T,p_0,\delta,\kappa_2)\frac{\Psi(x)}{(T-t_m)^{\frac{1}{2}}}\frac{h^{\frac{1}{4}}}{(t_m-t_k)^{\frac{1}{2}}}.
\end{aligned}
$$

For t_3 we use the conditional Hölder inequality, (28), (16) and Lemma 5.2:

$$
\begin{aligned}
\|t_3(m)\| &= \|\mathbb{E}_{\tau_k}\left[[F(t_m,X^n_{t_m})-F(t_m,X_{t_m})]N^{n,\tau_k}_{\tau_m}\sigma(t_{k+1},\mathcal{X}_{\tau_k})\right]\| \\
&\leq \frac{C(\sigma,\widehat{\kappa}_2)}{(t_m-t_k)^{\frac{1}{2}}}\|F(t_m,X^n_{t_m})-F(t_m,X_{t_m})\| \\
&\leq C(L_f,c^{2,3}_{4.3},b,\sigma,T,p_0,\delta)\frac{\Psi(x)}{(T-t_m)^{\frac{1}{2}}}\frac{h^{\frac{1}{4}}}{(t_m-t_k)^{\frac{1}{2}}}.
\end{aligned}
$$

The term t_4 can be estimated as follows:

$$
\begin{aligned}
\|t_4(m)\| &= \|\mathbb{E}_{\tau_k}\left[[f(t_m,X_{t_m},Y_{t_m},Z_{t_m})-f(t_{m+1},X^n_{t_m},Y^n_{t_m},Z^n_{t_m})]N^{n,\tau_k}_{\tau_m}\sigma(t_{k+1},\mathcal{X}_{\tau_k})\right]\| \\
&\leq \frac{C(L_f,b,\sigma,T,\delta)}{(t_m-t_k)^{\frac{1}{2}}}(h^{\frac{1}{2}}+\|X_{t_m}-X^n_{t_m}\|+\|Y_{t_m}-Y^n_{t_m}\|+\|Z_{t_m}-Z^n_{t_m}\|).
\end{aligned}
$$

Finally, for the remaining term of the estimate of $\|Z_{t_k}-Z^n_{t_k}\|$, we use (32) and (34) to get

$$
\begin{aligned}
\left\|\mathbb{E}_{t_k}f(s,X_s,Y_s,Z_s)N^{t_k}_s\sigma(t_k,X_{t_k})\right\| &= \|\mathbb{E}_{t_k}[(f(s,X_s,Y_s,Z_s)-f(s,X_{t_k},Y_{t_k},Z_{t_k}))N^{t_k}_s]\sigma(t_k,X_{t_k})\| \\
&\leq C(L_f,C^y_{4.2},C_{4.4},b,\sigma,T,p_0,\kappa_2)\Psi(x).
\end{aligned}
$$

Consequently, from (30), (31), the estimates for the remaining term and for $t_1,...,t_4$ it follows that

$$
\begin{aligned}
\|Z_{t_k}-Z^n_{t_k}\| \leq\ & C_{2.5}\hat{\Psi}(x)h^{\frac{\alpha}{2}}+(C^z_{3.1})^{\frac{1}{2}}\Psi(x)h^{\frac{1}{4}}+C(L_f,C^y_{4.2},C_{4.4},b,\sigma,T,p_0,\kappa_2)\Psi(x)h \\
&+C(L_f,C^y_{4.2},C_{4.4},\kappa_2,b,\sigma,T,\delta,p_0)\Psi(x)h^{\frac{1}{2}}\int_{t_k}^T\frac{ds}{(s-t_k)^{\frac{1}{2}}} \\
&+C(L_f,c^{2,3}_{4.3},b,\sigma,T,p_0,\delta,\kappa_2)h\sum_{m=k+1}^{n-1}\frac{\Psi(x)}{(T-t_m)^{\frac{1}{2}}}\frac{h^{\frac{1}{4}}}{(t_m-t_k)^{\frac{1}{2}}} \\
&+C(L_f,b,\sigma,T,\delta)h\sum_{m=k+1}^{n-1}(\|Y_{t_m}-Y^n_{t_m}\|+\|Z_{t_m}-Z^n_{t_m}\|)\frac{1}{(t_m-t_k)^{\frac{1}{2}}}
\end{aligned}
$$

$$
\begin{aligned}
\leq\quad & C(C_{2.5}, C^z_{3.1})\hat{\Psi}(x)h^{\frac{\alpha}{2}\wedge\frac14} + C(L_f, c^{2,3}_{4.3}, C^y_{4.2}, C_{4.4}, \kappa_2, b, \sigma, T, p_0, \delta)\Psi(x)h^{\frac14} \\
& +C(L_f, b, \sigma, T, \delta)\sum_{m=k+1}^{n-1}(\|Y_{t_m} - Y^n_{t_m}\| + \|Z_{t_m} - Z^n_{t_m}\|)\frac{1}{(t_m - t_k)^{\frac12}}h.
\end{aligned}
$$

Then we use (29) and the above estimate to get

$$
\begin{aligned}
& \|Y_{t_k} - Y^n_{t_k}\| + \|Z_{t_k} - Z^n_{t_k}\| \\
\leq\quad & C(C_{2.5}, C^z_{3.1})\hat{\Psi}(x)h^{\frac{\alpha}{2}\wedge\frac14} + C(L_f, C^y_{3.1}, c^{2,3}_{4.3}, C^y_{4.2}, C_{4.4}, b, \sigma, T, p_0, \kappa_2, \delta)\Psi(x)h^{\frac14} \\
& +C(L_f, b, \sigma, T, \delta)\sum_{m=k+1}^{n-1}(\|Y_{t_m} - Y^n_{t_m}\| + \|Z_{t_m} - Z^n_{t_m}\|)\frac{1}{(t_m - t_k)^{\frac12}}h.
\end{aligned}
$$

Consequently,

$$
\|Y_{t_k} - Y^n_{t_k}\| + \|Z_{t_k} - Z^n_{t_k}\| \quad\leq\quad C_{3.2}\hat{\Psi}(x)h^{\frac{\alpha}{2}\wedge\frac14}.
$$

By Theorem 4.2 it follows that

$$
\|Y_v - Y^n_v\| \leq \|Y_v - Y_{t_k}\| + \|Y_{t_k} - Y^n_{t_k}\| \leq C(C_{3.2}, C^y_{4.2})\hat{\Psi}(x)h^{\frac{\alpha}{2}\wedge\frac14},
$$

while Proposition 4.4 below implies that

$$
\|Z_v - Z_{t_k}\| \leq C_{4.4}\Psi(x)h^{\frac12}.
$$

□

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),

$$
X^{t,x}_s = x + \int_t^s b(r, X^{t,x}_r)dr + \int_t^s \sigma(r, X^{t,x}_r)dB_r, \quad 0 \leq t \leq s \leq T \tag{33}
$$

and recall the Malliavin weight and its properties from [18, 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 (33) then setting*

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

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

$$
\partial_x G(r, X^{t,x}_r) = \mathbb{E}[H(X^{t,x}_T)N^{r,(t,x)}_T|\mathcal{F}^t_r],
$$

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

$$
N^{r,(t,x)}_T = \frac{1}{T - r}\int_r^T \frac{\nabla X^{t,x}_s}{\sigma(s, X^{t,x}_s)\nabla X^{t,x}_r}dB_s,
$$

and $\nabla X_s^{t,x}$ *is given in* (11)*. Moreover, for* $q \in (0,\infty)$ *there exists a* $\kappa_q > 0$ *such that a.s.*

$$(\mathbb{E}[|N_T^{r,(t,x)}|^q|\mathcal{F}_r^t])^{\frac{1}{q}} \le \frac{\kappa_q}{(T-r)^{\frac{1}{2}}} \quad \text{and} \quad \mathbb{E}[N_T^{r,(t,x)}|\mathcal{F}_r^t] = 0 \ a.s. \tag{34}$$

and we have

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

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 [18, 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 [19].

Theorem 4.2. *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 $C^y_{4.2} > 0$ *such that for* $0 \le t < s \le T$ *and* $x \in \mathbb{R}$*,*

$$\|Y_s - Y_t\|_{L_p(\mathbb{P}_{t,x})} \le C^y_{4.2}\Psi(x) \left(\int_t^s (T-r)^{\alpha-1} dr\right)^{\frac{1}{2}},$$

(ii) there exists a constant $C^z_{4.2} > 0$ *such that for* $0 \le t < s < T$ *and* $x \in \mathbb{R}$*,*

$$\|Z_s - Z_t\|_{L_p(\mathbb{P}_{t,x})} \le C^z_{4.2}\Psi(x) \left(\int_t^s (T-r)^{\alpha-2} dr\right)^{\frac{1}{2}}.$$

The constants $C^y_{4.2}$ *and* $C^z_{4.2}$ *depend on* $K_f, L_f, C_g, c^{1,2}_{4.3}, T, p_0, b, \sigma, \kappa_q$ *and* p*.*

4.3 Properties of the associated PDE

Theorem 4.3 ([19], Theorem 5.6)**.** *Consider the FBSDE* (1) *and let Assumptions 2.1 and 2.2 hold. Then for the solution* u *of the associated PDE*

$$\begin{cases} u_t(t,x) + \frac{\sigma^2(t,x)}{2} u_{xx}(t,x) + b(t,x)u_x(t,x) + f(t,x,u(t,x),\sigma(t,x)u_x(t,x)) = 0, \\ \hfill t \in [0,T), x \in \mathbb{R}, \\ u(T,x) = g(x), \quad x \in \mathbb{R} \end{cases} \tag{35}$$

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* $|u(t,x)| \le c^1_{4.3}\Psi(x)$ *with* Ψ *given in* (7)*.*

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

(b) $Z_s^{t,x} = u_x(s, X_s^{t,x})\sigma(s, X_s^{t,x})$ *a.s.,*

(c) $|u_x(t,x)| \le \frac{c^2_{4.3}\Psi(x)}{(T-t)^{\frac{1-\alpha}{2}}}$*.*

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

(b) $|\partial_x^2 u(t,x)| \le \frac{c_{4.3}^3 \Psi(x)}{(T-t)^{1-\frac{\alpha}{2}}}$.

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 4.2.

Proposition 4.4. *If Assumption 2.3 holds, then there exists a constant $C_{4.4} > 0$ such that for $0 \le t < s \le T$ and $x \in \mathbb{R}$,*

$$\|Z_s - Z_t\|_{L_p(\mathbb{P}_{t,x})} \le C_{4.4}\Psi(x)(s-t)^{\frac{1}{2}},$$

where $C_{4.4}$ depends on $b, \sigma, T, p_0, g, f, p, c_{4.3}^{2,3}$.

Proof. From $Z_s^{t,x} = u_x(s, X_s^{t,x})\sigma(s, X_s^{t,x})$ and $\nabla Y_s^{t,x} = \partial_x u(s, X_s^{t,x}) = u_x(s, X_s^{t,x})\nabla X_s^{t,x}$ we conclude

$$Z_s^{t,x} = \frac{\nabla Y_s^{t,x}}{\nabla X_s^{t,x}}\sigma(s, X_s^{t,x}), \quad 0 \le t \le s \le T. \tag{36}$$

It is well-known (see e.g. [17]) that the solution ∇Y of the linear BSDE

$$\nabla Y_s = g'(X_T)\nabla X_T + \int_s^T f_x(\Theta_r)\nabla X_r + f_y(\Theta_r)\nabla Y_r + f_z(\Theta_r)\nabla Z_r dr - \int_s^T \nabla Z_r dB_r, \quad 0 \le s \le T, \tag{37}$$

can be represented as

$$\begin{aligned}\frac{\nabla Y_s}{\nabla X_s} &= \mathbb{E}_s\Big[g'(X_T)\nabla X_T \Gamma_T^s + \int_s^T f_x(\Theta_r)\nabla X_r \Gamma_r^s dr\Big]\frac{1}{\nabla X_s} \\ &= \tilde{\mathbb{E}}\Big[g'(\tilde{X}_T^{s,X_s})\nabla \tilde{X}_T^{s,X_s}\tilde{\Gamma}_T^{s,X_s} + \int_s^T f_x(\tilde{\Theta}_r^{s,X_s})\nabla \tilde{X}_r^{s,X_s}\tilde{\Gamma}_r^{s,X_s} dr\Big], \quad 0 \le t \le s \le T,\end{aligned} \tag{38}$$

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

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

and

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

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 (36),

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

Since $(\nabla Y_s, \nabla Z_s)$ is the solution to the linear BSDE (37) with bounded f_x, f_y, f_z, we have that $\|\nabla Y_t\|_{L_{2p}(\mathbb{P}_{t,x})} \le C(b,\sigma,T,p,f,g)$. 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 (38) 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 (25), we define the processes

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

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 $\tilde{\mathbb{E}}|\tilde{\Gamma}^{s,X_s}_T - \tilde{\Gamma}^{t,x}_T|^q$ below. From the above relations we get for ($X_s := X^{t,x}_s$)

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

Since g' is Lipschitz continuous and of polynomial growth, the estimate $J_1 \leq C(b,\sigma,g,T,p)\Psi(x)(s-t)^{\frac{1}{2}}$ 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 $\tilde{X}^{s,X_s}_T - \tilde{X}^{t,x}_T$ and $\nabla \tilde{X}^{s,X_s}_T - \nabla \tilde{X}^{t,x}_T$ like in Lemma 5.2. For the Γ differences we first apply the inequalities of Hölder and BDG:

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

Since f_y and f_z are bounded we have $\tilde{\mathbb{E}}|\tilde{\Gamma}^{s,X_s}_r|^q + \tilde{\mathbb{E}}|\tilde{\Gamma}^{t,x}_r|^q \leq C(T,f,q)$. Similar to (28), since f_x, f_y, f_z are Lipschitz continuous w.r.t. the space variables,

$$\begin{aligned}
|f_x(\tilde{\Theta}^{s,X_s}_r) - f_x(\tilde{\Theta}^{t,x}_r)| \quad &= \quad \Big| f_x(r, \tilde{X}^{s,X_s}_r, u(r, \tilde{X}^{s,X_s}_r), \sigma(r, \tilde{X}^{s,X_s}_r) u_x(r, \tilde{X}^{s,X_s}_r)) \\
&\qquad - f_x(r, \tilde{X}^{t,x}_r, u(r, \tilde{X}^{t,x}_r), \sigma(r, \tilde{X}^{t,x}_r) u_x(r, \tilde{X}^{t,x}_r)) \Big| \\
&\leq \quad C(T,f,\sigma,c^{2,3}_{4.3})(1 + |\tilde{X}^{s,X_s}_r|^{p_0+1} + |\tilde{X}^{t,x}_r|^{p_0+1}) \frac{|\tilde{X}^{s,X_s}_r - \tilde{X}^{t,x}_r|}{(T-r)^{\frac{1}{2}}},
\end{aligned}$$

so that Lemma 5.2 yields

$$\tilde{\mathbb{E}}|f_x(\tilde{\Theta}^{s,X_s}_r) - f_x(\tilde{\Theta}^{t,x}_r)|^q \leq C(b,\sigma,T,p_0,f,c^{2,3}_{4.3},q)(1 + |X_s|^{p_0+1} + |x|^{p_0+1})^q \frac{|X_s - x|^q + |s-t|^{\frac{q}{2}}}{(T-r)^{\frac{1}{2}}}.$$

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

$$\|\tilde{\mathbb{E}}[\tilde{\Gamma}^{s,X_s}_T - \tilde{\Gamma}^{t,x}_T]\|_p \quad \leq \quad C(b,\sigma,T,p_0,f,g,c^{2,3}_{4.3},p)\Psi(x)|s-t|^{\frac{1}{2}}$$

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 (12). By (4),

$$X_{t_{m+1}}^{n,t_m,x} = x + hb(t_{m+1},x) + \sqrt{h}\sigma(t_{m+1},x)\varepsilon_{m+1} \tag{39}$$

so that

$$T_{m+1,\pm}u^n(t_{m+1},X_{t_{m+1}}^{n,t_m,x}) = u^n(t_{m+1},x+hb(t_{m+1},x)\pm\sqrt{h}\sigma(t_{m+1},x)). \tag{40}$$

Proposition 4.5. *Let Assumption 2.3 hold and assume that u^n is a solution of*

$$\begin{aligned} &u^n(t_m,x) - hf(t_{m+1},x,u^n(t_m,x),\mathcal{D}_{m+1}^n u^n(t_{m+1},X_{t_{m+1}}^{n,t_m,x})) \\ &= \frac{1}{2}[T_{m+1,+}u^n(t_{m+1},X_{t_{m+1}}^{n,t_m,x}) + T_{m+1,-}u^n(t_{m+1},X_{t_{m+1}}^{n,t_m,x})], \quad m=0,\dots,n-1, \end{aligned} \tag{41}$$

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

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

and

$$|u_{xx}^n(t_m,x) - u_{xx}^n(t_m,\bar{x})| \le C_{u^n,3}(1+|x|^{6p_0+7}+|\bar{x}|^{6p_0+7})|x-\bar{x}|^\alpha, \tag{42}$$

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 (41), since g is C^2 and f_y is bounded, for sufficiently small h we conclude by induction (backwards in time) that $u_x^n(t_m,x)$ exists for $m=0,\dots,n-1$, and that it holds

$$\begin{aligned} u_x^n(t_m,x) = \; & hf_x(t_{m+1},x,u^n(t_m,x),\mathcal{D}_{m+1}^n u^n(t_{m+1},X_{t_{m+1}}^{n,t_m,x})) \\ & + hf_y(t_{m+1},x,u^n(t_m,x),\mathcal{D}_{m+1}^n u^n(t_{m+1},X_{t_{m+1}}^{n,t_m,x}))u_x^n(t_m,x) \\ & + hf_z(t_{m+1},x,u^n(t_m,x),\mathcal{D}_{m+1}^n u^n(t_{m+1},X_{t_{m+1}}^{n,t_m,x}))\partial_x\mathcal{D}_{m+1}^n u^n(t_{m+1},X_{t_{m+1}}^{n,t_m,x}) \\ & + \tfrac{1}{2}(\partial_x T_{m+1,+}u^n(t_{m+1},X_{t_{m+1}}^{n,t_m,x}) + \partial_x T_{m+1,-}u^n(t_{m+1},X_{t_{m+1}}^{n,t_m,x})). \end{aligned}$$

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

Step 2. As stated in the proof of Proposition 2.5, the finite difference equation (41) is the associated equation to (8) in the sense that we have the representations (19). We will use that $u^n(t_m,x) = Y_{t_m}^{n,t_m,x}$ and exploit the BSDE

$$Y_{t_m}^{n,t_m,x} = g(X_T^{n,t_m,x}) + \int_{(t_m,T]} f(s,X_{s^-}^{n,t_m,x},Y_{s^-}^{n,t_m,x},Z_{s^-}^{n,t_m,x})d[B^n]_s - \int_{(t_m,T]} Z_{s^-}^{n,t_m,x}dB_s^n, \tag{43}$$

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

$$\begin{aligned} \nabla Y_{t_m}^n = \partial_x Y_{t_m}^n = \; & g'(X_T^n)\partial_x X_T^n + \int_{(t_m,T]} f_x\partial_x X_{s^-}^n + f_y\partial_x Y_{s^-}^n + f_z\partial_x Z_{s^-}^n d[B^n]_s \\ & - \int_{(t_m,T]} \partial_x Z_{s^-}^n dB_s^n. \end{aligned} \tag{44}$$

Similarly as in the proof of [28, Theorem 3.1] the BSDE (44) can be derived from (43) as a limit of difference quotients w.r.t. x. Notice that the generator of (44) is random but has the same Lipschitz constant and linear growth bound as f. Assumption 2.3 allows us to find a $p_0 \geq 0$ and a $K > 0$ such that

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

In order to get estimates simultaneously for (43) and (44) we show the following lemma.

Lemma 4.6. *We fix n and assume a BSDE*

$$\mathsf{Y}_{t_k} = \xi^n + \int_{(t_k,T]} f(s, \mathsf{X}_{s-}, \mathsf{Y}_{s-}, \mathsf{Z}_{s-}) d[B^n]_s - \int_{(t_k,T]} \mathsf{Z}_{s-} dB^n_s, \quad m \leq k \leq n, \tag{45}$$

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

$$\begin{aligned} |\mathsf{f}(\omega,t,x,y,z) - \mathsf{f}(\omega,t,x',y',z')| &\leq L_f(|x-x'| + |y-y'| + |z-z'|), \\ |\mathsf{f}(\omega,t,x,y,z)| &\leq (K_f + L_f)(1 + |x| + |y| + |z|). \end{aligned} \tag{46}$$

Then for any $p \geq 2$,

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

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

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

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

Proof. (i) By Itô's formula (see [22, Theorem 4.57]) we get for $p \geq 2$

$$\begin{aligned} |\mathsf{Y}_{t_k}|^p &= |\xi^n|^p - p\int_{(t_k,T]} \mathsf{Y}_{s-}|\mathsf{Y}_{s-}|^{p-2}\mathsf{Z}_{s-}dB^n_s + p\int_{(t_k,T]} \mathsf{Y}_{s-}|\mathsf{Y}_{s-}|^{p-2}\mathsf{f}(s,\mathsf{X}_{s-},\mathsf{Y}_{s-},\mathsf{Z}_{s-})d[B^n]_s \\ &\quad - \sum_{s\in(t_k,T]} [|\mathsf{Y}_s|^p - |\mathsf{Y}_{s-}|^p - p\mathsf{Y}_{s-}|\mathsf{Y}_{s-}|^{p-2}(\mathsf{Y}_s - \mathsf{Y}_{s-})]. \end{aligned} \tag{47}$$

Following the proof of [25, 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

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

where $\gamma_p > 0$ is computed in [34, 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

$$\begin{aligned} &- \sum_{s\in(t_k,T]} [|\mathsf{Y}_s|^p - |\mathsf{Y}_{s-}|^p - p\mathsf{Y}_{s-}|\mathsf{Y}_{s-}|^{p-2}(\mathsf{Y}_s - \mathsf{Y}_{s-})] \\ &\leq -\gamma_p \sum_{\ell=k}^{n-1} |\mathsf{Y}_{t_\ell}|^{p-2}(\mathsf{f}(t_{\ell+1}, \mathsf{X}_{t_\ell}, \mathsf{Y}_{t_\ell}, \mathsf{Z}_{t_\ell})h - \mathsf{Z}_{t_\ell}\sqrt{h}\varepsilon_{\ell+1})^2 \end{aligned}$$

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

Hence we get from (47)

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

From Young's inequality and (46) 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

$$
\begin{aligned}
|\mathsf{Y}_{t_k}|^p \;\le\; & |\xi^n|^p - p \int_{(t_k,T]} \mathsf{Y}_{s-} |\mathsf{Y}_{s-}|^{p-2} \mathsf{Z}_{s-} dB^n_s + c \int_{(t_k,T]} 1 + |\mathsf{X}_{s-}|^p + |\mathsf{Y}_{s-}|^p d[B^n]_s \\
& - \tfrac{\gamma_p}{2} \int_{(t_k,T]} |\mathsf{Y}_{s-}|^{p-2} |\mathsf{Z}_{s-}|^2 d[B^n]_s. \qquad (48)
\end{aligned}
$$

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

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

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 (48) yields

$$
\mathbb{E}|\mathsf{Y}_{t_k}|^p + \tfrac{\gamma_p}{2} \mathbb{E} \int_{(t_k,T]} |\mathsf{Y}_{s-}|^{p-2} |\mathsf{Z}_{s-}|^2 d[B^n]_s \le \mathbb{E}|\xi^n|^p + c\mathbb{E} \int_{(t_k,T]} 1 + |\mathsf{X}_{s-}|^p + |\mathsf{Y}_{s-}|^p d[B^n]_s. \qquad (49)
$$

By Gronwall's lemma and the polynomial growth of $x \mapsto \mathbb{E}|\xi^n|^p$, and $x \mapsto \mathbb{E} \int_{(t_k,T]} 1 + |\mathsf{X}_{s-}|^p d[B^n]_s$,

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

and inserting this into (49) yields

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

(ii) From (48) we derive by the inequality of BDG and Young's inequality that for $t_m \le t_k \le T$

$$
\mathbb{E} \sup_{t_k < s \le T} |\mathsf{Y}_{s-}|^p
$$

$$
\begin{aligned}
&\leq \mathbb{E}|\xi^n|^p + C(p)\mathbb{E}\left(\int_{(t_k,T]} |\mathsf{Y}_{s-}|^{2p-2}|\mathsf{Z}_{s-}|^2 d[B^n]_s\right)^{\frac{1}{2}} + c\mathbb{E}\int_{(t_k,T]} 1 + |\mathsf{X}_{s-}|^p + |\mathsf{Y}_{s-}|^p d[B^n]_s \\
&\leq \mathbb{E}|\xi^n|^p + c\mathbb{E}\int_{(t_k,T]} 1 + |\mathsf{X}_{s-}|^p d[B^n]_s + C(p)\mathbb{E}\left[\sup_{t_k<s\leq T} |\mathsf{Y}_{s-}|^{\frac{p}{2}}\left(\int_{(t_k,T]} |\mathsf{Y}_{s-}|^{p-2}|\mathsf{Z}_{s-}|^2 d[B^n]_s\right)^{\frac{1}{2}}\right] \\
&\quad + c\mathbb{E}\int_{(t_k,T]} |\mathsf{Y}_{s-}|^p d[B^n]_s \\
&\leq \mathbb{E}|\xi^n|^p + c\mathbb{E}\int_{(t_k,T]} 1 + |\mathsf{X}_{s-}|^p d[B^n]_s + C(p)\mathbb{E}\int_{(t_k,T]} |\mathsf{Y}_{s-}|^{p-2}|\mathsf{Z}_{s-}|^2 d[B^n]_s \\
&\quad + \mathbb{E}\sup_{t_k<s\leq T} |\mathsf{Y}_{s-}|^p(\tfrac{1}{4} + c(T-t_k)).
\end{aligned}
$$

We assume that h is sufficiently small so that we find a t_k with $c(T-t_k) < \frac{1}{4}$. We rearrange the inequality to have $\mathbb{E}\sup_{t_k<s\leq T}|\mathsf{Y}_{s-}|^p$ on the l.h.s., and from (i) we conclude that

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

Now we may repeat the above step for $\mathbb{E}\sup_{t_\ell<s\leq t_k}|\mathsf{Y}_{s-}|^p$ with $c(t_k-t_\ell) < \frac{1}{4}$ 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 (45),

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

so that by (46) and the inequalities of BDG and Hölder we have that

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

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

$$
\mathbb{E}\Big(\int_{(t_k,T]}|\mathsf{Z}_{s-}|^2 d[B^n]_s\Big)^{\frac{p}{2}} \leq C(T,f,g,p,p_0,b,\sigma)(1+|x|^{(p_0+1)p}). \tag{50}
$$

Repeating this procedure eventually yields (iii). □

Step 3. Applying Lemma 4.6 to (43) and (44) we see that for all $m = 0,...,n$ we have

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

and

$$
|u_x^n(t_m,x)| = (\mathbb{E}(\partial_x Y_{t_m}^{n,t_m,x})^2)^{\frac{1}{2}} \leq c(T,f,g,p_0,b,\sigma)(1+|x|^{p_0+1}). \tag{51}
$$

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

$$\begin{aligned}
\partial^2_x Y^n_{t_m} &= g''(X^n_T)(\partial_x X^n_T)^2 + g'(X^n_T)\partial^2_x X^n_T \\
&\quad + \int_{(t_m,T]} \sum_{\substack{a\in\{0,1,2\}^3 \\ i_1+i_2+i_3=2}} (D^a f)(s,X^n_{s-},Y^n_{s-},Z^n_{s-})(\partial_x X^n_{s-})^{i_1}(\partial_x Y^n_{s-})^{i_2}(\partial_x Z^n_{s-})^{i_3} d[B^n]_s \\
&\quad + \int_{(t_m,T]} \sum_{\substack{a\in\{0,1\}^3 \\ i_1+i_2+i_3=1}} (D^a f)(s,X^n_{s-},Y^n_{s-},Z^n_{s-})(\partial^2_x X^n_{s-})^{i_1}(\partial^2_x Y^n_{s-})^{i_2}(\partial^2_x Z^n_{s-})^{i_3} d[B^n]_s \\
&\quad - \int_{(t_m,T]} \partial^2_x Z^n_{s-} dB^n_s.
\end{aligned} \tag{52}$$

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 (52), 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

$$\begin{aligned}
&\mathbb{E}|\hat{\mathsf{Y}}_{t_m}|^2 + \frac{1}{2}\mathbb{E}\int_{(t_m,T]}|\hat{\mathsf{Z}}_{s-}|^2 d[B^n]_s \\
&\quad \le C\Big[\mathbb{E}|\hat\xi^n|^2 + \mathbb{E}\Big(\int_{(t_m,T]}|f_0(s^-)|d[B^n]_s\Big)^2 + \mathbb{E}\int_{(t_m,T]}|\hat{\mathsf{X}}_{s-}|^2 + |\hat{\mathsf{Y}}_{s-}|^2 d[B^n]_s\Big].
\end{aligned} \tag{53}$$

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 $\mathbb{E}\sup_{t_m<s\le T}|\partial^j_x X^n_s|^p$ 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 4.6(ii)-(iii), so that Gronwall's lemma implies

$$|\hat{\mathsf{Y}}^{t_m,x}_{t_m}| = |u_{xx}(t_m,x)| \le C\Psi^2(x). \tag{54}$$

Finally, to show (42), one uses (52) and derives an inequality as in (53) but now for the difference $\partial^2_x Y^{n,t_m,x}_{t_m} - \partial^2_x Y^{n,t_m,\bar x}_{t_m}$.

Before proving it, let us state the following lemma.

Lemma 4.7. *Let Assumption 2.3 hold. We have*

$$\Big(\mathbb{E}\sup_s |Z^{n,t_m,x}_{s^-} - Z^{n,t_m,\bar x}_{s^-}|^p\Big)^{1/p} \le C(\Psi(x)^2 + \Psi(\bar x)^2)|x-\bar x|, \quad p\ge 2, \tag{55}$$

$$\mathbb{E}\Big(\int_{(t_m,T]} |\partial_x Z^{n,t_m,x}_{s^-} - \partial_x Z^{n,t_m,\bar x}_{s^-}|^2 d[B^n]_s\Big)^{\frac p2} \le C(\Psi^{4p}(x) + \Psi^{4p}(\bar x))|x-\bar x|^p, \quad p\ge 2, \tag{56}$$

$$\mathbb{E}\left(\int_{(t_m,T]} |\partial^2_x Z^{n,t_m,x}_{s^-}|^2 d[B^n]_s\right)^{\frac p2} \le C\Psi^{4p}(x), \quad p\ge 2, \tag{57}$$

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

Proof of Lemma 4.7. (55): Introduce $G(t_{k+1},x) := \mathcal{D}^n_{k+1}u^n(t_{k+1}, X^{n,t_k,x}_{t_{k+1}})$. Using relations (39)–(40) and the bounds (51) and (54) for u^n_x and u^n_{xx}, 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 $Z^{n,t_m,x}_{t_k} = \mathcal{D}^n_{k+1}u^n(t_{k+1}, X^{n,t_k,\eta}_{t_{k+1}}) = G(t_{k+1},\eta)$ where $\eta = X^{n,t_m,x}_{t_k}$, the previous bound yields

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

uniformly for each $t_m \le t_k < T$. Inequality (55) then follows by applying the Cauchy-Schwarz inequality and standard L_p-estimates for the process X^n.

(56): This can be shown similarly as Lemma 4.6-(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 (44) itself.

(57): This one gets repeating again the proof of Lemma 4.6-(iii) but now for the BSDE (52). □

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 (50) and (55)

$$\begin{aligned}
&\mathbb{E}\Big(\int_{(t_m,T]} |Z^{n,t_m,x}_{s^-} - Z^{n,t_m,\bar x}_{s^-}|\,|\partial_x Z^{n,t_m,x}_{s^-}|^2 d[B^n]_s\Big)^2 \\
&\le \; (\mathbb{E}\sup_s |Z^{n,t_m,x}_{s^-} - Z^{n,t_m,\bar x}_{s^-}|^4)^{\frac12}\Big(\mathbb{E}\Big(\int_{(t_m,T]}|\partial_x Z^{n,t_m,x}_{s^-}|^2 d[B^n]_s\Big)^4\Big)^{\frac12} \\
&\le \; C(\Psi^4(x)+\Psi^4(\bar x))|x-\bar x|^2\Psi^4(x).
\end{aligned}$$

The second term is bounded by using (50) and (56):

$$\mathbb{E}\Big(\int_{(t_m,T]} |\partial_x Z^{n,t_m,x}_{s^-}|^2 - |\partial_x Z^{n,t_m,\bar x}_{s^-}|^2 d[B^n]_s\Big)^2$$

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

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

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

we need the bound (57).

The result follows then from Gronwall's lemma.

Remark 4.8. *Proposition 4.5 implies that there exists a constant $C = C(T, f, g, p, p_0, b, \sigma) > 0$ such that*

$$
|\partial_x \mathcal{D}^n_{m+1} u^n(t_{m+1}, X^{n,t_m,x}_{t_{m+1}}) - \partial_x \mathcal{D}^n_{m+1} u^n(t_{m+1}, X^{n,t_m,\bar{x}}_{t_{m+1}})| \leq C(1 + \hat{\Psi}(x) + \hat{\Psi}(\bar{x}))|x-\bar{x}|^\alpha, \tag{58}
$$

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

$$
\hat{\Psi}(x) := 1 + |x|^{6p_0+8}. \tag{59}
$$

□

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 (Lemma A.1 [19]). *For all $0 \leq k \leq m \leq n$ and $p > 0$, it holds for $h = \frac{T}{n}$ and τ_k defined in (22) 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 and its discretisation, where we assume that B and $\tilde{B}$ are connected as in (25).

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

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

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

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

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

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

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

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

(iii): This can be easily seen because the process $(\nabla X^{s,y}_r)_{r\in[s,T]}$ solves a linear SDE with bounded coefficients.
(iv): The process solves (61). The estimate follows from the inequality of BDG and Gronwall's lemma.
(v): Recall that from (4) and (24) we have

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

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

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

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

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

We define an 'increasing' map $i(r) := t_{l+1}$ for $(t_l, t_{l+1}]$ 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)

$$\begin{aligned}
&\tilde{\mathbb{E}}|\tilde{X}^{t_k,x}_{t_k+t_m} - \tilde{X}^{n,t_k,y}_{t_k+t_m}|^p \\
&\quad \le \quad C(b,p)\left(|x-y|^p + \tilde{\mathbb{E}}\int_0^{t_m} |r-i(r)|^{\frac{p}{2}} + |\tilde{X}^{t_k,x}_{t_k+r} - \tilde{X}^{t_k,x}_{t_k+d(r)}|^p + |\tilde{X}^{t_k,x}_{t_k+d(r)} - \tilde{X}^{n,t_k,y}_{t_k+d(r)}|^p dr\right) \\
&\qquad + C(p)\tilde{\mathbb{E}}|\int_{t_m\wedge\tilde{\tau}_m}^{t_m} \sigma(t_k+r, \tilde{X}^{t_k,x}_{t_k+r})d\tilde{B}_r|^p \\
&\qquad + C(p)\tilde{\mathbb{E}}|\int_{t_m\wedge\tilde{\tau}_m}^{\tilde{\tau}_m} \sum_{l=0}^{m-1} \sigma(t_{k+l+1}, \tilde{X}^{n,t_k,y}_{t_{k+l}})\mathbf{1}_{(\tilde{\tau}_l,\tilde{\tau}_{l+1}]}(r)d\tilde{B}_r|^p \\
&\qquad + C(p)\tilde{\mathbb{E}}|\int_0^{t_m\wedge\tilde{\tau}_m} \sigma(t_k+r, \tilde{X}^{t_k,x}_{t_k+r}) - \sum_{l=0}^{m-1} \sigma(t_{k+l+1}, \tilde{X}^{n,t_k,y}_{t_{k+l}})\mathbf{1}_{(\tilde{\tau}_l,\tilde{\tau}_{l+1}]}(r)d\tilde{B}_r|^p. \qquad (60)
\end{aligned}$$

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

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

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

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

Finally, by the BDG inequality

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

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)

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

Using Lemma 5.1-(iii) one concludes similarly as in the proof of (ii) that $\tilde{\mathbb{E}}|\tilde{X}^{t_k,x}_{t_k+\tilde{\tau}_l\wedge t_m}-\tilde{X}^{t_k,x}_{t_k+t_l}|^p\leq C(b,\sigma,T,p)h^{\frac{p}{4}}$. Then (60) combined with the above estimates implies that

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

Then Gronwall's lemma yields

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

(vi): We have

$$\begin{aligned}
\nabla\tilde{X}^{n,t_k,y}_{t_k+t_m}\ &=\ 1+\int_{(0,t_m]}b_x(t_k+r,X^{n,t_k,y}_{t_k+r^-})\nabla\tilde{X}^{n,t_k,y}_{t_k+r^-}d[\tilde{B}^n,\tilde{B}^n]_r\\
&\quad+\int_{(0,t_m]}\sigma_x(t_k+r,\tilde{X}^{n,t_k,y}_{t_k+r^-})\nabla\tilde{X}^{n,t_k,y}_{t_k+r^-}d\tilde{B}^n_r
\end{aligned}\tag{61}$$

and

$$\nabla \tilde{X}^{t_k,x}_{t_k+t_m} = 1 + \int_0^{t_m} b_x(t_k+r, \tilde{X}^{t_k,x}_{t_k+r}) \nabla \tilde{X}^{t_k,x}_{t_k+r} dr + \int_0^{t_m} \sigma_x(t_k+r, \tilde{X}^{t_k,x}_{t_k+r}) \nabla \tilde{X}^{t_k,x}_{t_k+r} d\tilde{B}_r. \tag{62}$$

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 (25). Under Assumption 2.1 it holds that*

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

Proof. For $N^{n,\tau_k}_{\tilde{\tau}_m}$ and $N^{t_k}_{t_m}$ given by (10) and (15), 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,

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

The assertion follows then from Lemma 5.1 and from the estimates

$$\tilde{\mathbb{E}} \sup_{s\in[0,t_{m-k}]\cap[\tilde{\tau}_{\ell-1},\tilde{\tau}_\ell]} |a_{t_k+s} - a^n_{\tau_k+\tilde{\tau}_\ell}|^4 \le C(b,\sigma,T,\delta)(|X_{t_k} - X^n_{t_k}|^4 + h) \tag{63}$$

$$\tilde{\mathbb{E}} \sup_{s\in[0,t_{m-k}]} |a_{t_k+s}|^4 + \tilde{\mathbb{E}} \max_{1\le\ell\le m-k} |a^n_{\tau_k+\tilde{\tau}_\ell}|^4 \le 2\|\sigma\|^4_\infty \delta^{-4}. \tag{64}$$

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 (63) let us split $a_{t_k+s} - a^n_{\tau_k+\tilde\tau_\ell}$ in the following way:

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

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 (62) and (61), respectively. Therefore,

$$\tilde{\mathbb E} \sup_{s\in[0,t_{m-k}]} |\nabla \tilde X^{t_k,X_{t_k}}_{t_k+s}|^p \le C(b,\sigma,T,p) \quad \text{and} \quad \tilde{\mathbb E} \max_{0\le\ell\le m-k} |\nabla \tilde{\mathcal X}^{\tau_k,\mathcal X_{\tau_k}}_{\tilde\tau_\ell+\tau_k}|^p \le C(b,\sigma,T,p). \tag{65}$$

For the second term we get

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

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 (64) use the estimates (65). □

Acknowledgement

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

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

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

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

[6] 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.

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

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

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

[10] J.-F. Chassagneux, D. Crisan, and F. Delarue. *Numerical Method for FBSDEs of McKean-Vlasov type.* https://arxiv.org/abs/1703.02007, 2017.

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

[12] J.-F. Chassagneux and A. Richou. *Rate of convergence for discrete-time approximation of reflected BSDEs arising in switching problems.* https://arxiv.org/abs/1602.00015, 2016.

[13] 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.

[14] 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.

[15] 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.

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

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

[18] 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.

[19] C. Geiss, C. Labart, and A. Luoto. *Random walk approximation of BSDEs with Hölder continuous terminal condition.* arXiv:1806.07674, 2018.

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

[21] 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.

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

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

[24] 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.

[25] 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.

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

[27] 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.

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

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

[30] 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.

[31] 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.

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

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

[34] S. Yao. *Lp Solutions of Backward Stochastic Differential Equations with Jumps.* arXiv:1007.2226v1, 2010.

[35] 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.* https://arxiv.org/abs/1708.03223, 2017.

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

135. ZHANG, GUO, Liouville theorems for stationary flows of generalized Newtonian fluids. (14 pp.) 2012
136. RAJALA, TUOMAS, Use of secondary structures in the analysis of spatial point patterns. (27 pp.) 2012
137. LAUKKARINEN, EIJA, On Malliavin calculus and approximation of stochastic integrals for Lévy processes. (21 pp.) 2012
138. GUO, CHANGYU, Generalized quasidisks and the associated John domains. (17 pp.) 2013
139. ÄKKINEN, TUOMO, Mappings of finite distortion: Radial limits and boundary behavior. (14 pp.) 2014
140. ILMAVIRTA, JOONAS, On the broken ray transform. (37 pp.) 2014
141. MIETTINEN, JARI, On statistical properties of blind source separation methods based on joint diagonalization. (37 pp.) 2014
142. TENGVALL, VILLE, Mappings of finite distortion: Mappings in the Sobolev space $W^{1,n-1}$ with integrable inner distortion. (22 pp.) 2014
143. BENEDICT, SITA, Hardy-Orlicz spaces of quasiconformal mappings and conformal densities. (16 pp.) 2014
144. OJALA, TUOMO, Thin and fat sets: Geometry of doubling measures in metric spaces. (19 pp.) 2014
145. KARAK, NIJJWAL, Applications of chaining, Poincaré and pointwise decay of measures. (14 pp.) 2014
146. JYLHÄ, HEIKKI, On generalizations of Evans and Gangbo's approximation method and L^∞ transport. (20 pp.) 2014
147. KAURANEN, AAPO, Space-filling, energy and moduli of continuity. (16 pp.) 2015
148. YLINEN, JUHA, Decoupling on the Wiener space and variational estimates for BSDEs. (45 pp.) 2015
149. KIRSILÄ, VILLE, Mappings of finite distortion on generalized manifolds. (14 pp.) 2015
150. XIANG, CHANG-LIN, Asymptotic behaviors of solutions to quasilinear elliptic equations with Hardy potential. (20 pp.) 2015
151. ROSSI, EINO, Local structure of fractal sets: tangents and dimension. (16 pp.) 2015
152. HELSKE, JOUNI, Prediction and interpolation of time series by state space models. (28 pp.) 2015
153. REINIKAINEN, JAAKKO, Efficient design and modeling strategies for follow-up studies with time-varying covariates. (36 pp.) 2015
154. NUUTINEN, JUHO, Maximal operators and capacities in metric spaces. (22 pp.) 2016
155. BRANDER, TOMMI, Calderón's problem for p-Laplace type equations. (21 pp.) 2016
156. ÄRJE, JOHANNA, Improving statistical classification methods and ecological status assessment for river macroinvertebrates. (30 pp.) 2016
157. HELSKE, SATU, Statistical analysis of life sequence data. (40 pp.) 2016
158. RUOSTEENOJA, EERO, Regularity properties of tug-of-war games and normalized equations. (16 pp.) 2017
159. ZHANG, YI, Planar Sobolev extension domains. (13 pp.) 2017
160. YLITALO, ANNA-KAISA, Statistical inference for eye movement sequences using spatial and spatio-temporal point processes. (45 pp.) 2017
161. LINDQVIST, PETER, Notes on the p-Laplace equation. (106 pp.) 2017
162. LEHTONEN, JERE, Injectivity results for the geodesic ray transform. (15 pp.) 2017
163. GOLO, SEBASTIANO NICOLUSSI, Topics in the geometry of non-Riemannian Lie groups. (14 pp.) 2017
164. KOPRA, JUHO, Statistical modelling of selective non-participation in health examination surveys. (25 pp.) 2018
165. HEINO, JOONAS, On the local and global regularity of tug-of-war games. (20 pp.) 2018
166. MUKHERJEE, SHIRSO, Regularity of quasilinear sub-elliptic equations in the Heisenberg group. (18 pp.) 2018
167. NANDI, DEBANJAN, Applications of the quasihyperbolic metric. (14 pp.) 2018
168. TIKKA, SANTTU, Improving identification algorithms in causal inference. (35 pp.) 2018

ISBN 978-951-39-7575-3 (print)
ISBN 978-951-39-7576-0 (pdf)
ISSN 1457-8905