Backward Euler method for stochastic differential equations with non-Lipschitz coefficients driven by fractional Brownian motion

Hao Zhou1 · Yaozhong Hu2 · Yanghui Liu3

Received: 5 August 2022 / Accepted: 21 June 2023 / Published online: 7 July 2023
© The Author(s), under exclusive licence to Springer Nature B.V. 2023

Abstract

We study the traditional backward Euler method for stochastic differential equation driven by fractional Brownian motion whose drift coefficient satisfies the one-sided Lipschitz condition. The backward Euler scheme is proved to be of order one and this rate is optimal by showing the asymptotic error distribution result. Numerical experiments are performed to validate our claims about the optimality of the rate of convergence.

Keywords Backward Euler method · Stochastic differential equation · Malliavin derivative · Strong convergence · Asymptotic error distribution

Mathematics Subject Classification 65L04 · 60H10 · 60H35

Communicated by Charles-Edouard Bréhier.

Hao Zhou is supported by the China Scholarship Council. Yaozhong Hu is supported by an NSERC discovery grant and a centennial fund from University of Alberta at Edmonton. Yanghui Liu is supported by the PSC-CUNY Award 64353-00 52.

Hao Zhou haozhouhit@163.com
Yaozhong Hu yaozhong@ualberta.ca
Yanghui Liu yanghui.liu@baruch.cuny.edu

1 School of Mathematics, Harbin Institute of Technology, Harbin 150001, China
2 Department of Mathematical and Statistical Sciences, University of Alberta, Edmonton, AB T6G 2G1, Canada
3 Department of Mathematics, Baruch College, CUNY, New York, NY 10010, USA
1 Introduction

In this paper, we consider the following $m$-dimensional stochastic differential equation (SDE for short)
\[ dX_t = b(X_t)dt + dB_t, \quad t \in (0, T], \quad (1.1) \]
where $B_t$ is a fractional Brownian motion (fBm for short) with Hurst parameter $H > 1/2$ and $X_0 = x_0$ is a constant in $\mathbb{R}^m$. We assume that the drift function $b : \mathbb{R}^m \to \mathbb{R}^m$ satisfies the one-sided Lipschitz condition and that $b$ and its derivatives may be unbounded. One typical example that we have in mind is given by $b(x) = x - x^3$, $x \in \mathbb{R}$ or any polynomial of odd degree with a negative leading coefficient. The existence and uniqueness of the solution for this equation are given by Proposition 7.1 in Appendix. The work [1] studied the ergodicity of the solution for this type of equation. To our knowledge the existence and uniqueness of solution for a multiplicative noise under the one-sided Lipschitz condition are unknown. In this paper, we will focus on the additive SDE and leave the multiplicative case for future research.

This paper aims to consider the numerical approximation of the Eq. (1.1). The numerical approximation of SDE driven by Brownian motion is well studied; see e.g. the monographs [2, 3]. The numerical approximation for fractional SDE has also received much attention in recent works; see e.g. [4–10]. The optimal approximation for multiplicative noise SDE is considered in [11]. The multilevel Monte Carlo method for additive SDE is demonstrated in [4].

When the drift function $b$ is one-sided Lipschitz only, it is well-known that the explicit methods for the SDE are conditionally stable and implicit methods are usually preferred (see Table 1 in Example 6.1 for performance comparison among some numerical methods). The numerical approximation for such SDE has been investigated in [6], in which the almost sure convergence rate $H$ of the backward Euler method is obtained.

In this paper, we focus on the backward Euler method. We prove that the strong convergence rate for the backward Euler method is 1.0. The main step in our proof is to derive a decomposition (see (4.42) and (5.26)) of the error process in terms of the quantities of the form $(I - \Delta \partial b)^{-1}$ which is shown to be bounded by $e^{\lambda \Delta}$ under the one-sided Lipschitz condition. Here $\partial b$ is the Jacobian of $b$, $\lambda$ is the largest eigenvalue of $\partial b$, $\Delta > 0$ is a small number, and $I$ is the identity matrix. This decomposition, together with some Malliavin calculus techniques allows us to treat the difficulties

| Value  | $T = 0.08$ | $T = 0.16$ | $T = 0.24$ | $T = 0.32$ | $T = 0.40$ | $T = 0.48$ | $T = 0.56$ | $T = 0.64$ | $T = 0.72$ |
|--------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| EM     | -5.0498   | 4.8234    | -11.586945.4107 | -4.2063e+54.4809e+6 | -8.0082e+464.1562e+127 | -Inf       |
| CN     | -0.7427   | 3.7617    | -4.6843   | 23.6835   | -2.1032e+52.2404e+6 | -4.0041e+462.0781e+127 | -Inf       |
| BEM    | 3.5643    | 2.7000    | 2.2182    | 1.9562    | 1.6969    | 1.5096    | 1.4973    | 1.7401    | 1.5886    |
| Exact  | 2.3221    | 1.7058    | 1.6053    | 1.6447    | 1.3966    | 1.2202    | 1.3487    | 1.6903    | 1.4840    |
due to the unboundedness of the coefficient in the equation. In the second part of our paper, we show that the asymptotic error of the backward Euler method converges in $L^p$ to the solution of a linear SDE for $p \in [1, 2)$. Our main approach in this part is the combination of a variation of parameter method and limit theorems for weighted random sums (see [7, 12]).

The remainder of the paper is structured as follows. Section 2 concerns some basic preparation work on Malliavin calculus and Young integrals. The properties of the exact solution and Malliavin derivatives of the SDE (1.1) are shown in Sect. 3. In Sect. 4, the convergence rate of the backward Euler method is investigated. The convergence rate is proved to be an optimal one in Sect. 5. Then in Sect. 6, we report our numerical experiments and illustrate the accuracy of the method.

2 Preliminaries

2.1 Elements of Malliavin calculus

Let $(B_t = (B^1_t, \ldots, B^m_t), t \in [0, T])$ be an $m$-dimensional fBm of Hurst parameter $H \in (0, 1)$ defined in a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Recall that an fBm $B_t$ is a continuous ($m$-dimensional) mean zero Gaussian process with covariance given by

$$
\mathbb{E}(B^i_t B^j_s) = \frac{\delta_{ij}}{2} \left[t^{2H} + s^{2H} - |t-s|^{2H}\right], \quad \forall \ s, t \in [0, T],
$$

where $i, j = 1, 2, \ldots, m$ and $\delta_{ij}$ is the Kronecker symbol. We shall assume that the Hurst parameter is greater than $1/2$: $H > 1/2$. The case $H < 1/2$ demands different handling, in particular for the central limiting type theorem, and will be considered in our future projects. For a set $A$ we denote by $1_A$ the indicate function such that $1_A(x) = 1$ if $x \in A$ and $1_A(x) = 0$ otherwise. Associated with this covariance, we define

$$(1_{(a,b]}, 1_{(c,d)})_{\mathcal{H}}$$

$$= \int_b^a \int_c^d \phi(u, v) du dv = \frac{1}{2} \left[|d-a|^{2H} + |c-b|^{2H} - |d-b|^{2H} - |c-a|^{2H}\right],$$

where $\phi(u, v) = H(2H - 1)|u - v|^{2H-2}$, and

$$\phi(u, v) = H(2H - 1)|u - v|^{2H-2}.$$

$\langle \cdot, \cdot \rangle_{\mathcal{H}}$ can be extended by (bi-)linearity to a scalar product on the set $\mathcal{S}$ of step functions of the form $f(t) = \sum_{i=1}^n f_i \cdot 1_{(a_i, b_i]}(t)$, $f_i \in \mathbb{R}$. We denote by $\mathcal{H}$ the Hilbert space obtained by completing $\mathcal{S}$ with respect to the above scalar product. Note that $\mathcal{H}$ contains generalized functions. If $f, g : [0, T] \to \mathbb{R}$ are integrable
measurable functions, then
\[
\langle f, g \rangle_H = \int_0^T \int_0^T f(u)g(v)\phi(u, v)dudv.
\]

For the stochastic analysis associated with the fBm, we refer to [5, 13, 14] and we shall use their notations freely. For example, we shall use the Itô type stochastic integral \( \delta (f) = \int_0^T f(s)dB_s \) and \( \delta (f \cdot 1_{(a, b]}) = \int_a^b f(s)dB_s \) freely. But let us recall some notations and results on the Malliavin calculus for the fBm, since some of the results we are going to use cannot be found in the standard literature. We denote by \( P \) the set of all nonlinear functionals (random variables) on \((\Omega, \mathcal{F}, P)\) of the form
\[
F = f(B_{t_1}, \ldots, B_{t_n}),
\]
where \( 0 \leq t_1 < \cdots < t_n \leq T \) and \( f = f((x_{ij})_{1 \leq i \leq m, 1 \leq j \leq n}) \) is a smooth function of \( n \times m \) variables of polynomial growth. For \( f \in P \) of the above form, we define its Malliavin derivative as
\[
D_r^{(i)} F = \sum_{j=1}^n \frac{\partial}{\partial x_{ij}} f(B_{t_1}, \ldots, B_{t_n}) \cdot 1_{[0, t_i]}(r), \quad i = 1, \ldots, m.
\]
We consider \( D^{(i)} F = D^{(i)} F \) and \( DF = (D^{(1)} F, \ldots, D^{(m)} F) \) as \( H \)-valued random variables. We can also define the higher order derivatives \( D^k F \) which can be considered as an \( H^{\otimes k} \)-valued random variable. For \( p \geq 1 \) and \( k \geq 1 \), we denote by
\[
\| F \|_{k, p}^p := \sum_{\ell=0}^k \mathbb{E} \left[ \left\| D^\ell F \right\|_{H^{\otimes \ell}}^p \right],
\]
and we denote by \( \mathbb{D}^{k, p} \) the completion of \( P \) with respect to the above norm. We recall that the following integration by parts formula holds for \( F \in \mathbb{D}^{1, p} \), \( p \geq 1 \) and \( g \in H \)
\[
\mathbb{E} \left[ F \int_0^T g(s)dB_s \right] = \mathbb{E} \left[ \int_0^T \int_0^T D_s Fg(t)\phi(s, t)dtsdt \right] = \mathbb{E} \left[ \langle DF, g \rangle_H \right]. \tag{2.1}
\]
For \( f \in C^1(\mathbb{R}^m) \) with bounded partial derivatives and \( G = (G_1, \ldots, G_m) \in \mathbb{D}^{1, p} \), we have the chain rule: \( Df(G) = \sum_{i=1}^m \frac{\partial}{\partial G_i} f(G)DG_i \).

We end this section by recalling the following differentiation rule.

**Lemma 2.1** Let \( Z \) be an arbitrary random matrix in \((\Omega, \mathcal{F}, P)\). Assume that all entries of \( Z \) are Malliavin differentiable and \( Z^{-1} \) exists almost surely. Then
\[
DZ^{-1} = -Z^{-1}(DZ)Z^{-1}.
\]
If furthermore, all entries of \( Z \) are twice Malliavin differentiable, then
\[
D^2 Z^{-1} = 2Z^{-1} \cdot DZ \cdot Z^{-1} \cdot DZ \cdot Z^{-1} - Z^{-1} \cdot D^2 Z \cdot Z^{-1}.
\]

Proof See p.172 in [15]. \( \square \)

2.2 Elements of Young integral

We define the Hölder continuous functions as follows.

Definition 2.1 Let \( f \) be a continuous function on \([0, T]\) and let \( \alpha \in (0, 1) \). The Hölder semi-norm \( \| \cdot \|_\alpha \) is defined by
\[
\| f \|_\alpha = \sup_{0 \leq s, t \leq T} \frac{|f_t - f_s|}{|t - s|^\alpha}.
\]

When \( \| f \|_\alpha < \infty \), we call \( f \) a Hölder continuous function of order \( \alpha \).

We have the following result on the Young integral (see [16]).

Lemma 2.2 Let \( \alpha, \beta \in (0, 1) \) be such that \( \alpha + \beta > 1 \). Suppose that \( f \) and \( g \) are Hölder continuous of order \( \alpha \) and \( \beta \) on \([0, T]\) respectively. Then the Young integral \( h_t := \int_0^t (f_s - f_0) dg_s \) is well-defined for \( t \in [0, T] \), and there exists a constant \( C \) depending on \( \alpha \) and \( \beta \) such that the following relation holds
\[
\| h \|_\beta \leq C \| f \|_\alpha \cdot \| g \|_\beta.
\]

3 Dissipativity and properties of the solution

In this section, we derive some properties of the solution for Eq. (1.1) under the one-sided Lipschitz condition (see (3.3)) and the polynomial growth condition (see (3.4)). The existence and uniqueness of the solution will be proved in Proposition 7.1.

3.1 Moment bounds for Eq. (1.1)

We make the following assumptions throughout the remaining part of the paper. In the following, we denote by
\[
\partial b \equiv \left( \frac{\partial b_i}{\partial x_j} \right)_{1 \leq i, j \leq m}
\]
the Jacobian of (the vector field) \( b \), which is a continuous function from \( \mathbb{R}^m \) to the set of \( m \times m \) matrices. We also denote by \( \partial^2 b(x) = \left( \frac{\partial^2 b_i}{\partial x_j \partial x_k} (x) \right)_{1 \leq i, j, k \leq m} \) and \( \partial^3 b(x) = \left( \frac{\partial^3 b_i}{\partial x_j \partial x_k \partial x_\ell} (x) \right)_{1 \leq i, j, k, \ell \leq m} \). In the following and throughout the paper for a vector
we denote by $|x|$ the Euclidean norm and for a matrix $A$ we define its norm by

$$|A| = \|A\|_\infty = \sup_{|x|=1} |Ax|.$$  

(3.2)

**Assumption 3.1** We assume the following.

(A1) The coefficient $b$ is one-sided Lipschitz. Namely, there is a constant $\kappa$ such that

$$\langle x - y, b(x) - b(y) \rangle \leq \kappa |x - y|^2, \quad \forall x, y \in \mathbb{R}^m.$$  

(3.3)

(A2) The coefficient $b$ itself and its first and second derivatives are of polynomial growth. Namely, there are constants $\kappa$ and $\mu$ such that

$$|b(x)| + |\partial b(x)| + |\partial^2 b(x)| \leq \kappa (1 + |x|^\mu), \quad \forall x \in \mathbb{R}^m.$$  

(3.4)

where $|\partial b(x)|$ and $|\partial^2 b(x)|$ stand for

$$|\partial b(x)| = \sqrt{\sum_{i,j=1}^m \left| \partial_{x_j} b_i(x) \right|^2} \quad \text{and} \quad |\partial^2 b(x)| = \sqrt{\sum_{i,j,k=1}^m \left| \partial_{x_j} \partial_{x_k} b_i(x) \right|^2}.$$  

(3.5)

Under Assumption 3.1, we obtain the following results.

**Proposition 3.2** Let Assumption 3.1 be satisfied. Then Eq. (1.1) has a unique solution $X_t$. For any integer $p \geq 1$, there exists a constant $C$ depending only on $\kappa$, $\mu$, $p$ and $T$ such that

$$\mathbb{E} \left[ \sup_{0 \leq t \leq T} |X_t|^p \right] \leq C.$$  

(3.6)

**Proof** Take $V_t = X_t - B_t$. Then Eq. (1.1) becomes a deterministic equation with random coefficient

$$dV_t = b(V_t + B_t)dt,$$

and can be solved in a pathwise way, see the result in Proposition 7.1. Let us differentiate $|V_t|^2$ to obtain

$$d \frac{d}{dt} |V_t|^2 = 2\langle V_t, \dot{V}_t \rangle = 2\langle V_t, b(V_t + B_t) \rangle.$$  

(3.7)

Equation (3.7) can be rewritten as

$$\frac{d}{dt} |V_t|^2 = 2\langle (V_t + B_t) - B_t, b(V_t + B_t) - b(B_t) \rangle + 2\langle V_t, b(B_t) \rangle.$$  

(3.8)
Now applying Assumption 3.1 (A1) to the first inner product on the right side of (3.8) and applying the elementary inequality $\langle a, b \rangle \leq \frac{1}{2}(|a|^2 + |b|^2)$ and Assumption 3.1 (A2) to the second inner product, we obtain
\[
\frac{d}{dt}|V_t|^2 \leq (2\kappa + 1)|V_t|^2 + \kappa^2 (1 + |B_t|^{2\mu}).
\]
By Gronwall’s lemma, it turns out that
\[
|V_t|^2 \leq e^{(2\kappa+1)t}|V_0|^2 + \int_0^t \kappa^2 e^{(2\kappa+1)(t-s)}(1 + |B_s|^{2\mu})ds.
\]
Taking the square root on both sides and taking into account that $V_0 = x_0$, we obtain
\[
sup_{0 \leq t \leq T} |V_t| \leq e^{(\kappa+1/2)T}|x_0| + \sqrt{\int_0^T \kappa^2 e^{(2\kappa+1)(t-s)}(1 + |B_s|^{2\mu})ds}.
\]
By the self-similarity property of the fBm and then Fernique’s lemma, we have
\[
\mathbb{E}\left[\sup_{0 \leq t \leq T} |B_t|^\mu\right] = \mathbb{E}\left[\sup_{0 \leq t \leq 1} |B_t|^\mu\right] \cdot T^{\mu H} = CT^{\mu H}.
\]
(3.9)
Now we use the relation $X_t = V_t + B_t$ to obtain
\[
\mathbb{E}\left[\sup_{0 \leq t \leq T} |X_t|^p\right] \leq 2^p \mathbb{E}\left(\sup_{0 \leq t \leq T} |V_t|^p + \sup_{0 \leq t \leq T} |B_t|^p\right) \leq C,
\]
for some finite constants $C$. The proof is completed. □

The next lemma considers the Hölder continuity of $X_t$.

**Lemma 3.1** Let $X_t$ be the solution of Eq. (1.1) and let the assumptions be as in Proposition 3.2. Take $\alpha$ such that $0 < \alpha < H$. Then
\[
\|X_t\|_\alpha \leq \kappa (1 + \sup_{0 \leq t \leq T} |X_t|^{\mu}) + \|B\|_\alpha \quad \text{and} \quad \mathbb{E}\left[\sup_{0 \leq s, s \leq T} \frac{|X_t - X_s|^p}{|t-s|^{\alpha p}}\right] \leq C, \quad \forall p \geq 1.
\]
(3.10)

**Proof** From the Eq. (1.1) satisfied by $X_t$, it follows
\[
X_t - X_s = \int_s^t b(X_u)du + B_t - B_s.
\]
(3.11)
Applying Assumption 3.1 (A2) and then (3.6) to $b(X_u)$ and taking Hölder norm on both sides of (3.11), we obtain the first relation in (3.10). The second relation in (3.10) follows from the first one by Fernique’s theorem. □
3.2 The Malliavin derivatives for Eq. (1.1)

To find the rate of convergence of our numerical scheme in the next section, we will need the Malliavin differentiability of Eq. (1.1). To this end, we make the following additional assumption on the coefficient.

**Assumption 3.3** We assume that \( \partial_i b(x), i = 0, \ldots, 3 \) are of polynomial growth, namely, there are constants \( \kappa, \mu \) such that

\[
\max_{i=0,\ldots,3} |\partial_i^j b(x)| \leq \kappa (1 + |x|^\mu), \quad \forall \ x \in \mathbb{R}^m, \tag{3.12}
\]

where \( |\partial^i b(x)|, i = 0, 1, 2 \) are defined in (3.5) and we set \( |\partial^3 b(x)| = \sqrt{\sum_{i,j,k,\ell=1}^m \left| \frac{\partial^3 b_i}{\partial x_j \partial x_k \partial x_\ell} (x) \right|^2} \).

Let \( r \in [0, T] \). Taking the Malliavin derivative \( D_r \) on both sides of Eq. (1.1) leads to

\[
D_r X_t = \int_r^t \partial b(X_s) D_r X_s ds + \text{Id}, \quad t \in [r, T], \tag{3.13}
\]

and \( D_r X_t = 0 \) for \( t < r \). On the left-hand side of (3.13), we have denoted by \( D_r X_t = \left( D_r^{(j)} X_t^k, \ j, k = 1, \ldots, m \right) \) while on the right-hand side we have used the notation

\[
\partial b(X_s) D_r X_s = \left( \sum_{i=1}^m \frac{\partial b_k}{\partial x_i} (X_s) D_r^{(j)} X_t^i, \ j, k = 1, \ldots, m \right).
\]

It follows from Proposition 7.1 that there exists a unique solution \( D_r X_t \) for the Eq. (3.13). The following lemma is taken from [17].

**Lemma 3.2** The coefficient \( b \) satisfies the one-sided Lipschitz condition (A1) if and only if

\[
\langle x, \partial b(y)x \rangle \leq \kappa |x|^2, \quad \forall \ x, y \in \mathbb{R}^m.
\]

for any \( x \in \mathbb{R}^m \).

Lemma 3.2 can be used to prove the following bounds for the Malliavin derivatives of the solution \( X_t \).

**Proposition 3.4** Let \( X_t \) be the solution of Eq. (1.1) and let Assumptions 3.1 and 3.3 be satisfied. Then for any \( p \geq 1 \), there is a finite constant \( C \) depending only on \( \kappa, \mu, T, p \) such that

\[
\mathbb{E} \left[ \sup_{0 \leq r, t \leq T} |D_r X_t|^p \right] \leq C, \tag{3.14}
\]
and

$$\mathbb{E} \left[ \sup_{0 \leq r, r', t \leq T} |D_{rr}^2 X_t|^p \right] \leq C.$$  

(3.15)

**Proof** Fix $r$ and consider (3.13) as an equation of $U_t$, where $U_t := D_r X_t$, $r \leq t \leq T$. Then

$$dU_t = \partial b(X_t) U_t dt, \quad t \in [r, T],$$

and $U_r = \text{Id}$. Differentiating $|U_t|^2$ yields

$$\frac{d}{dt} |U_t|^2 = 2 \langle U_t, \partial b(X_t) U_t \rangle \leq 2\kappa |U_t|^2.$$  

(3.16)

Applying Gronwall’s lemma yields

$$|D_r X_t|^2 \leq C e^{2\kappa (t - r)}.$$  

(3.17)

This proves (3.14). Now we want to prove (3.15). Differentiating (3.13) again, we have for $t \geq r \vee r'$

$$D_{rr'}^2 X_t = \int_{r \vee r'}^t \partial b(X_s) D_{rr'}^2 X_s ds + \int_{r \vee r'}^t \langle \partial^2 b(X_s), D_{rr'} X_s \otimes D_r X_s \rangle ds,$$  

(3.18)

and $D_{rr'}^2 X_t = 0$ for $t < r \vee r'$, where the inner product on the right-hand side of (3.18) stands for

$$\langle \partial^2 b(X_s), D_{rr'} X_s \otimes D_r X_s \rangle = \left( \sum_{k, \ell, \ell' = 1}^m \frac{\partial^2 b_k}{\partial x_k \partial x_j} (X_s) D_{\ell \ell'}^{\ell'}(X_s) D_{\ell \ell'}(X_s) \right)^{1/2}.$$  

As in (3.16), by differentiating $|D_{rr'}^2 X_t|^2$ in $t$ and then performing the similar argument, we obtain the relation

$$\sup_{0 \leq r, r', t \leq T} |D_{rr'}^2 X_t|^2 \leq C e^{2\kappa T}.$$  

From the Assumption 3.3 and our bound (3.14) for $D_r X_t$, it follows

$$\sup_{0 \leq r, r', t \leq T} |D_{rr'}^2 X_t|^2 \leq C \sup_{0 \leq t \leq T} |X_t|^{2\mu}.$$  

Now combining the above with the moment bound (3.6), we conclude (3.15). \qed
Remark 3.1 It is interesting to point out that we have the almost sure uniform bound (3.17) for the first Malliavin derivative of the true solution $D_t X_t$.

4 Rate of convergence for the backward Euler scheme

In this section, we apply the backward Euler scheme to approximate the solution of (1.1). The convergence rate of this scheme will also be studied. In the next section, we will show that this rate of convergence is optimal.

4.1 Preparations

Let
\[ \pi : 0 = t_0 < t_1 < \cdots < t_{n-1} < t_n = T \] (4.1)

be a partition of the time interval $[0, T]$. We denote by
\[ \Delta_k = t_{k+1} - t_k, \quad \Delta B_k = B_{t_{k+1}} - B_{t_k}, \quad |\pi| = \sup_{0 \leq k \leq n-1} \Delta_k. \] (4.2)

The classical backward Euler scheme applied to (1.1) is
\[ Y_{t_{k+1}}^\pi = Y_{t_k}^\pi + b(Y_{t_{k+1}}^\pi) \Delta_k + \Delta B_k, \quad k = 0, 1, \ldots, n - 1. \] (4.3)

This is an implicit scheme. To find $Y_{t_{k+1}}^\pi$ from $Y_{t_k}^\pi$, we need to solve a function equation
\[ Y_{t_{k+1}}^\pi - b(Y_{t_{k+1}}^\pi) \Delta_k = Y_{t_k}^\pi + \Delta B_k. \] (4.4)

This scheme gives all the values of the approximation at the partition points $t_0, t_1, \ldots, t_n$. To compare the approximation solution with the true solution, we need to know the value of the approximation solution at all time instants. We shall use the following interpolation: let $Y_{t}^\pi$ satisfy
\[ Y_{t_{k+1}}^\pi = Y_{t_k}^\pi + b(Y_{t_{k+1}}^\pi)(t - t_k) + (B_t - B_{t_k}), \quad t \in (t_k, t_{k+1}], \quad k = 0, 1, \ldots, n - 1. \] (4.5)

To simplify the notation in the remaining part of the paper, we omit the explicit dependence of $Y_{t}^\pi$ on the partition $\pi$.

Before we proceed to study the rate of convergence, let us recall the following preparatory results; see [17, Lemma 2.2].

Lemma 4.1 If $J$ is an $m \times m$ square matrix satisfying
\[ \langle x, Jx \rangle \leq \lambda \|x\|^2, \quad \forall x \in \mathbb{R}^m, \] (4.6)
then for all \( t \in \mathbb{R} \) such that \( \lambda t < 1 \), \( \text{Id} - tJ \) is invertible and

\[
\| (\text{Id} - tJ)^{-1} \|_{\infty} \leq (1 - \lambda t)^{-1},
\]

where we recall that \( \| \cdot \|_{\infty} \) means the operator norm for an operator from \( \mathbb{R}^m \) to \( \mathbb{R}^m \) (see (3.2)).

We also have the following result.

**Lemma 4.2** Assume that the coefficient \( b \) satisfies the one-sided Lipschitz condition (A1) for some \( \kappa > 0 \), and that \( b \) is continuously differentiable. Then when \( t \kappa < 1 \), the matrix \( \text{Id} - t \int_0^1 \partial b(x + y \theta) d\theta \) is invertible for any \( x, y \in \mathbb{R}^m \) and we have

\[
\left\| \left( \text{Id} - t \int_0^1 \partial b(x + y \theta) d\theta \right)^{-1} \right\|_{\infty} \leq (1 - \kappa t)^{-1}.
\]

**Proof** Set \( J = \int_0^1 \partial b(x + y \theta) d\theta \). It follows from Lemma 2.1 in [17] that \( J \) satisfies condition (4.6) with \( \lambda = \kappa \). The relation (4.8) is then obtained by applying Lemma 4.1 to \( J \).

\( \square \)

### 4.2 Properties of the approximate solutions

We will need some properties of the approximated solution \( Y_t \) generated by the backward Euler scheme (4.3) for our convergence result. In this subsection, we are concerned with the upper-bound estimates of the approximate solutions.

**Proposition 4.1** Suppose that Assumptions 3.1 and 3.3 hold true with constants \( \kappa, \mu > 0 \). Let \( Y_t \) be the solution of the backward Euler scheme defined in (4.5) with partition \( \pi \) and \( k \)th step size \( \Delta_k > 0 \) defined in (4.1)–(4.2). Assume that \( \kappa \cdot |\pi| < 1 \) for all \( k \).

Then for any \( p \geq 1 \), there is a finite constant \( C \) such that

\[
\mathbb{E} \left[ \sup_{0 \leq t \leq T} |Y_t|^p \right] \leq C.
\]

**Proof** We first argue that (4.4) is well-defined under the condition \( \kappa \cdot |\pi| < 1 \). Define the mapping \( f : x \to x - b(x) \cdot \Delta_k - (Y_{t_k} + \Delta B_k) \). It suffices to show that \( f(x) = 0 \) has a unique solution. We compute the Jacobian of \( f \)

\[
\partial f(x) = \text{Id} - \partial b(x) \cdot \Delta_k.
\]

As in the proof of Lemma 4.2, we can apply Lemma 2.1 in [17] to show that \( \partial b(x) \) satisfies condition (4.6) with \( \lambda = \kappa \) for all \( x \). Therefore, taking into account the relation \( \kappa \cdot |\pi| < 1 \), we can apply Lemma 4.1 to get \( |\partial f(x)| \geq 1 - \kappa \cdot |\pi| > 0 \) for all \( x \), where recall that \( |\cdot| \) stands for the operator norm. Taking into account the continuity
of $\partial f$, it follows from the inverse function theorem that $f(x)$ has a unique solution. Therefore (4.4) is well-defined.

In the following we show that $E\left[\sup_{0 \leq k \leq n} |Y_{t_k}|^p\right] < \infty$. The estimate (4.9) for the supremum over all $t \in [0, T]$ can be proved similarly and is thus omitted for sake of simplicity. The proof can be considered as a discretized version of that of (3.6).

Denote by $\tilde{Y}_{t_k} = Y_{t_k} - B_{t_k}$ for $0 \leq k \leq n$. Then we can write Eq. (4.3) as

$$\tilde{Y}_{t_{k+1}} = \tilde{Y}_{t_k} + \left[b(\tilde{Y}_{t_{k+1}} + B_{t_{k+1}}) - b(B_{t_{k+1}})\right] \Delta_k + b(B_{t_{k+1}}) \Delta_k$$

$$= \tilde{Y}_{t_k} + \Delta_k \cdot \int_0^1 \partial b(\theta \tilde{Y}_{t_{k+1}} + B_{t_{k+1}}) d\theta \cdot \tilde{Y}_{t_{k+1}} + b(B_{t_{k+1}}) \Delta_k . \quad (4.10)$$

Solving Eq. (4.10) for $\tilde{Y}_{t_{k+1}}$, we obtain

$$\tilde{Y}_{t_{k+1}} = \beta_k (\tilde{Y}_{t_k} + b(B_{t_{k+1}}) \Delta_k) , \quad (4.11)$$

where

$$\beta_k = \left(\text{Id} - \Delta_k \cdot \int_0^1 \partial b(\theta \tilde{Y}_{t_{k+1}} + B_{t_{k+1}}) d\theta\right)^{-1} .$$

Iterating (4.11), we get

$$\tilde{Y}_{t_{k+1}} = \sum_{j=0}^k \beta_k \cdots \beta_j b(B_{t_{j+1}}) \Delta_j + \beta_k \cdots \beta_0 Y_0 . \quad (4.12)$$

Let us apply Lemma 4.2 to $\beta_k$ with $t = \Delta_k$, $x = B_{t_{k+1}}$ and $y = \tilde{Y}_{t_{k+1}}$. This yields

$$|\beta_k| \leq (1 - \kappa \Delta_k)^{-1} \leq e^{2\kappa \Delta_k} . \quad (4.13)$$

Applying the estimate (4.13) and Assumption 3.1 (A2) for $b$ to (4.12), we have

$$|\tilde{Y}_{t_{k+1}}| \leq e^{2\kappa T} \sum_{j=0}^k \sup_{0 < j \leq k-1} |b(B_{t_j})| \Delta_j + e^{2\kappa T} |Y_0|$$

$$\leq e^{2\kappa T} T \left(1 + \sup_{0 < j \leq k-1} |B_{t_j}|^{\mu}\right) + e^{2\kappa T} |Y_0| . \quad (4.14)$$

The estimate (4.9) follows by taking supremum for $k = 0, \ldots, n$, then taking expectation on both sides of (4.14) and taking into account the estimate of $B$ in (3.9). \qed
4.3 Malliavin derivatives of the numerical scheme

In this subsection, we consider the Malliavin derivatives for the backward Euler scheme under the one-sided Lipschitz condition.

Proposition 4.2 Let b satisfy Assumptions 3.1 and 3.3 and let Y_t be defined by (4.5). Then for any p \geq 1, there is a finite constant C independent of the partition \pi such that

\[ E \left[ \sup_{0 \leq r, t \leq T} |D_r Y_t|^p \right] \leq C \quad \text{and} \quad E \left[ \sup_{0 \leq r, r', t \leq T} |D^2_{rr} Y_t|^p \right] \leq C. \tag{4.15} \]

Proof Take r \in [0, T]. Note that since Y is adapted to the filtration generated by B, we have D_r Y_t = 0 for all r > t. In the following, we focus on the case when r \leq t. Let k_0 be such that t_{k_0} < r \leq t_{k_0+1}. For a fixed t \in [0, T], we set k_1 = \max\{i : t_i < t, t_i \in \pi\} and define

\[ \eta(t) = t_{k_1}. \tag{4.16} \]

Take the Malliavin derivative \( D = (D^{(1)}, \ldots, D^{(m)}) \) on both sides of (4.5). This yields

\[ D_r Y_t = D_r Y_{\eta(t)} + \partial b(Y_t) D_r Y_t(t - \eta(t)) + \Id \cdot 1_{(\eta(t), t]}(r). \tag{4.17} \]

Solving equation in (4.17) for D_r Y_t, we get

\[ D_r Y_t = (\Id - \partial b(Y_t)(t - \eta(t)))^{-1} \cdot (D_r Y_{\eta(t)} + \Id \cdot 1_{(\eta(t), t]}(r)). \tag{4.18} \]

Iterating (4.18), we obtain

\[ D_r Y_t = \alpha_{t} \alpha_{\eta(t)} \cdots \alpha_{t_{k_0+1}}, \tag{4.19} \]

where

\[ \alpha_{t} = (\Id - (t - \eta(t)) \partial b(Y_t))^{-1}. \]

Now applying Lemma 4.2 to \( \alpha_{t} \) with y = 0 and x = Y_t, we get

\[ |D_r Y_t| \leq |\alpha_{t}| \cdot |\alpha_{\eta(t)}| \cdots |\alpha_{t_{k_0+1}}| \leq (1 - \Delta_{k_1} \kappa)^{-1} (1 - \Delta_{k_1-1} \kappa)^{-1} \cdots (1 - \Delta_{k_0} \kappa)^{-1} \leq e^{2 \kappa (t - r)} \cdot (\Delta_{k_1} + \Delta_{k_1-1} + \cdots + \Delta_{k_0}) \leq e^{2 \kappa (t - r)}. \tag{4.20} \]

Recall that |·| = \|·\|_{\infty} stands for the operator norm defined in (3.2). This proves the first bound in (4.15).
Finally, applying (4.9) we have for any $|\alpha_{t_0+1}|$
\[
D_{r'}^{(j)} (D_r Y_t) = D_{r'}^{(j)} [\alpha_t \alpha_{t_0+1}] \quad \sum_{k=k_0+1}^{k_1} \alpha_t \alpha_{t_0} \cdots \alpha_{t_k+1} (D_{r'}^{(j)} \alpha_{t_k}) \alpha_{t_{k-1}} \cdots \alpha_{t_{t_{k_0+1}}} \quad + (D_{r'}^{(j)} \alpha_t) \alpha_{t_0} \cdots \alpha_{t_{t_{k_0+1}}}.
\]

To compute $D_{r'}^{(j)} \alpha_{t_k}$ or $D_{r'}^{(j)} \alpha_t$ we use Lemma 2.1, it satisfies
\[
D_{r'}^{(j)} \alpha_{t_k} = - (\text{Id} - \Delta_{t_k} - \partial b(Y_{t_k})) (\text{Id} - \Delta_{t_k} - \partial b(Y_{t_k}))^{-1}
\]
\[
= - (\text{Id} - \Delta_{t_k} - \partial b(Y_{t_k})) (\Delta_{t_k} - \partial^2 b(Y_{t_k}) D_{r'}^{(j)} Y_{t_k}) (\text{Id} - \Delta_{t_k} - \partial b(Y_{t_k}))^{-1}.
\]

Hence, we further have
\[
|D_{r'}^{(j)} (D_r Y_t)| \leq \sum_{k=k_0+1}^{k_1} |\alpha_t| |\alpha_{t_0}| \cdots |\alpha_{t_{k+1}}| |D_{r'}^{(j)} \alpha_{t_k}| |\alpha_{t_{k-1}}| \cdots |\alpha_{t_{t_{k+1}}}| + |D_{r'}^{(j)} \alpha_t| |\alpha_{t_0}| \cdots |\alpha_{t_{t_{k_0+1}}}| \leq e^{2\kappa (t-r)} \sum_{k=k_0+1}^{k_1} |D_{r'}^{(j)} \alpha_{t_k}| + e^{2\kappa (t-r)} |D_{r'}^{(j)} \alpha_t|,
\]
(4.21)
where in the last inequality we used bounds analogous to (4.20). By Assumption 3.3 and then (4.20), we get
\[
|D_{r'}^{(j)} \alpha_{t_k}| \leq C \Delta_{t_k} (1 + |Y_{t_k}|^\mu) |D_{r'}^{(j)} Y_{t_k}| \leq C \Delta_{t_k} (1 + |Y_{t_k}|^\mu).
\]
(4.22)

The same bound holds for $|D_{r'}^{(j)} \alpha_t|$. The proof is similar to $|D_{r'}^{(j)} \alpha_{t_k}|$ and is omitted. Substituting (4.22) into (4.21), we obtain
\[
|D_{r'}^{(j)} (D_r Y_t)| \leq C e^{2\kappa (t-r)} \sum_{k=k_0+1}^{k_1} \Delta_{t_k} (1 + |Y_{t_k}|^\mu) + C e^{2\kappa (t-r)} \Delta_{t_k} (1 + |Y_{t_k}|^\mu) \leq C (1 + |Y_{t_k}|^\mu).
\]

Finally, applying (4.9) we have for any $p \geq 1$
\[
\mathbb{E} \left[ \sup_{0 \leq t \leq T} |D_{r'}^{(j)} (D_r Y_t)|^p \right] \leq C \mathbb{E} \left[ \sup_{0 \leq t \leq T} (1 + |Y_t|^p) \right] \leq \infty.
\]

This proves the second inequality in (4.15).


4.4 Estimate for $E(\Delta B^i_tuv)$ and $E(\Delta B^i_tuv \Delta B^j_{st})$

In this subsection, we derive an upper-bound estimate result for the increments of the fBm $B_t$, which will be useful in the proof of convergence rate in Sect. 4.5. For convenience let us denote by

$$\Delta B^i_{u,v} = B^i_v - B^i_u,$$

(4.23)

for $0 \leq u \leq v \leq T$ and $i = 1, \ldots, m$.

**Lemma 4.3** Let $F : \Omega \rightarrow \mathbb{R}$ be a random variable that possesses the second Malliavin derivative. If $E|F| < \infty$ and if $\sup_{0 \leq s \leq T} E|D^{(i)}_s F| < \infty$, then

$$\left| E\left(F \Delta B^i_{u,v}\right) \right| \leq C|u - v|.$$  (4.24)

Assume further that $\sup_{0 \leq s, t \leq T} E|D^{ij}_{st} F| < \infty$. Then for any $0 \leq u < v \leq T$ and $0 \leq s < t \leq T$

$$\left| E\left(F \Delta B^i_{u,v} \Delta B^j_{s,t}\right) \right| \leq C \langle I(s,t), I(u,v) \rangle_{\mathcal{H}} + C|v - u| \cdot |t - s|.$$  (4.25)

**Proof** Applying (2.1) with $g(t) = 1_{(u,v)}(t)$ yields

$$\left| E\left(F \Delta B^i_{u,v}\right) \right| = \left| E\left(\int_0^T \int_0^T D^{(i)}_s F 1_{(u,v)}(t) \phi(s,t) ds dr \right) \right|$$

$$\leq \sup_{0 \leq s \leq T} E|D^{(i)}_s F| \cdot \int_u^v \int_0^T \phi(s,t) ds dr.$$  

Now we have

$$\int_0^T \phi(s,t) ds = \frac{H}{2H - 1} \left[ (T - t)^{2H-1} + t^{2H-1} \right] \leq C T^{2H-1}.$$  (4.26)

Thus (4.24) is proved.

Next, we turn to prove (4.25). Applying (2.1) twice, we obtain

$$\left| E\left(F \Delta B^i_{u,v} \Delta B^j_{s,t}\right) \right|$$

$$= \left| E\left(\langle D^{ij} F, 1_{(u,v) \times (s,t)} \rangle_{\mathcal{H} \otimes \mathcal{H}} \right) + E\left[ F 1_{(u,v)}, I(s,t) \right] 1_{[i=j]} \right|$$

$$= \left| E\left(\int_0^T \int_0^T \int_u^v \int_s^t D^{(j)} \phi(x,y) \phi(\xi,\zeta) dy d\xi dx d\zeta \right) \right|$$

$$+ E\left[ F 1_{(u,v)}, I(s,t) \right] 1_{[i=j]}$$

$$\leq E\left[ \sup_{0 \leq x, \xi \leq T} |D^{ij}_{\xi x} F| \int_0^T \int_0^T \int_u^v \int_s^t \phi(\xi,\zeta) dy d\xi dx d\zeta \right]$$

$$+ E\left| F \right| \langle I(s,t), I(u,v) \rangle_{\mathcal{H}}.$$  (4.27)
As in (4.26) we derive
\[
\int_0^T \int_0^T \int_u^v \int_s^t \phi(x, y)\phi(\xi, \zeta) dy\zeta dx\xi \leq |v - u| \cdot |t - s|. \tag{4.28}
\]
Substituting (4.28) into (4.27), we obtain the estimate (4.25). 

\[\Box\]

### 4.5 Rate of convergence

After these preparations, we return to explore the convergence of the backward Euler scheme, namely the convergence of \(Y_t\) defined by (4.5) to the solution \(X_t\) defined by the stochastic differential Eq. (1.1). We shall also obtain the rate of this convergence, which will be shown to be optimal in the next section. To complete the proof, we introduce the following notations and essential lemmas. Recall that \(t_k, k = 0, \ldots, n\) and \(\pi\) are defined in (4.1)–(4.2) and \(\eta(t)\) is defined in (4.16). We define processes \(\Gamma_t\) and \(A_t\) for all \(t\) in \([0, T]\)

\[
\Gamma_t = \text{Id} - (t - \eta(t)) \int_0^1 \partial b(uX_t + (1 - u)Y_t) du, \tag{4.29}
\]
and

\[
A_t = \Gamma_t^{-1} = \left(\text{Id} - (t - \eta(t)) \int_0^1 \partial b(uX_t + (1 - u)Y_t) du\right)^{-1}. \tag{4.30}
\]

For convenience we will also denote by \(A_{k+1} = A_{t_{k+1}}\) for \(k = 0, \ldots, n\). Resorting to Lemma 3.2 it is easily seen the existence of \(A_t\) when \(\kappa \cdot |\pi| < 1\). Furthermore, using Lemma 4.2 leads to the following lemma for \(A_t\).

**Lemma 4.4** Let \(A_t\) be defined by (4.30). Suppose that the partition \(\pi\) satisfies \(\kappa \cdot |\pi| < 1\). Then

\[
|A_t| \leq (1 - \kappa(t - \eta(t)))^{-1} \leq e^{2\kappa(t - \eta(t))}. \tag{4.31}
\]

In the following, we consider the Malliavin derivatives of \(\Gamma_t\).

**Lemma 4.5** Let \(\Gamma_t\) be defined by (4.29). Then for \(p \geq 1\), there exists a constant \(C\) independent of \(|\pi|\) such that

\[
\mathbb{E}\left[\sup_{0 \leq t, \xi, \zeta \leq T} |D_\xi \Gamma_t|^p\right] \leq C|\pi|^p \quad \text{and} \quad \mathbb{E}\left[\sup_{0 \leq t, \xi, \zeta \leq T} |D_{\xi\xi}^2 \Gamma_t|^p\right] \leq C|\pi|^p. \tag{4.32}
\]

**Proof** By the definition of \(\Gamma_t\) we have

\[
D_\xi^{(j)} \Gamma_t = -(t - \eta(t)) \int_0^1 D_\xi^{(j)} \partial b(uX_t + (1 - u)Y_t) du
\]

\[\square\] Springer
\[= -(t - \eta(t)) \int_0^1 \partial^2 b(uX_t + (1 - u)Y_t) \cdot (uD^{(j)}_\xi X_t + (1 - u)D^{(j)}_\xi Y_t)du.\]

Applying Assumption 3.1 (A2) and then using the elementary inequality \(ab \leq \frac{1}{2}(a^2 + b^2)\) and taking into account that \(t - \eta(t) \leq |\pi|\), we obtain
\[
|D^{(j)}_\xi \Gamma_t| \leq C(t - \eta(t))(1 + |X_t|^{2\mu} + |Y_t|^{2\mu} + |D^{(j)}_\xi X_t|^2 + |D^{(j)}_\xi Y_t|^2). 
\]

Apply Propositions 3.2, 3.4, 4.1 and 4.2 to \(X_t\), \(D^{(j)}_\xi X_t\), \(Y_t\) and \(D^{(j)}_\xi Y_t\) respectively. We obtain
\[
\mathbb{E} \left[ \sup_{0 \leq t, \xi \leq T} |D_\xi \Gamma_t|^p \right] \leq C|\pi|^p, \quad \forall p \geq 1. 
\]

This is the first inequality in (4.32).

Differentiating (4.29) twice gives
\[
D^{ij}_{\xi \zeta} \Gamma_t = -(t - \eta(t)) \int_0^1 \partial^3 b(uX_t + (1 - u)Y_t) 
\cdot (uD^{(j)}_\xi X_t + (1 - u)D^{(i)}_\xi Y_t) \otimes (uD^{(j)}_\xi X_t + (1 - u)D^{(j)}_\xi Y_t)du 
- (t - \eta(t)) \int_0^1 \partial^2 b(uX_t + (1 - u)Y_t) \cdot (uD^{ij}_{\xi \zeta} X_t + (1 - u)D^{ij}_{\xi \zeta} Y_t)du. 
\]

Thus applying Assumption 3.3 on the coefficient \(b\), we have
\[
|D^{2}_{\xi \zeta} \Gamma_t| \leq C|\pi| \left( 1 + |X_t|^{2\mu} + |Y_t|^{2\mu} \right) 
\cdot \left( (|D_\xi X_t| + |D_\xi Y_t|)(|D_\zeta X_t| + |D_\zeta Y_t|) + |D^{2}_{\xi \xi} X_t| + |D^{2}_{\xi \zeta} Y_t| \right) 
\leq C|\pi| \left( 1 + |X_t|^{2\mu} + |Y_t|^{2\mu} + |D_\xi X_t|^4 + |D_\xi Y_t|^4 \right. 
+ |D_\xi X_t|^4 + |D_\zeta Y_t|^4 + |D^{2}_{\xi \xi} X_t|^2 + |D^{2}_{\xi \zeta} Y_t|^2 \bigg). 
\]

Applying Propositions 3.2, 3.4, 4.1 and 4.2 again yields the second inequality of (4.32). This proves the lemma. \(\Box\)

Let us now turn to the estimate of the Malliavin derivatives of \(A_t\).

**Lemma 4.6** Let \(A_t\) be defined by (4.30). Then for \(p \geq 1\), there exists a constant \(C\) independent of \(|\pi|\) such that
\[
\mathbb{E} \left[ \sup_{0 \leq t, \xi, \zeta \leq T} |D_\xi A_t|^p \right] \leq C|\pi|^p \quad \text{and} \quad \mathbb{E} \left[ \sup_{0 \leq t, \xi, \zeta \leq T} |D^{2}_{\xi \zeta} A_t|^p \right] \leq C|\pi|^p. 
\]
**Proof** Differentiating $A_t$ and taking into account the relation $A_t = \Gamma_t^{-1}$ and Lemma 2.1, we have

$$D_\xi A_t = -A_t \cdot D_\xi \Gamma_t \cdot A_t.$$ 

Applying Lemmas 4.4 and 4.5 for $A_t$ and $D_\xi \Gamma_t$, we thus obtain the first inequality in (4.33).

Applying Lemma 2.1 again we obtain

$$D^2_\xi A_t = A_t \cdot D_\xi \Gamma_t \cdot A_t \cdot D_\xi \Gamma_t \cdot A_t + A_t \cdot D_\xi \Gamma_t \cdot A_t \cdot D_\xi \Gamma_t \cdot A_t - A_t \cdot D^2_\xi \Gamma_t \cdot A_t.$$ 

Then applying Lemmas 4.4 and 4.5 again, we obtain the second inequality in (4.33). $\square$

For $s, t \in [0, T]$ and $j = 0, \ldots, n$, we define

$$b_1(s, j) = \int_0^1 \partial b(uX_{t_{j+1}/2} + (1 - u)X_s)du.$$ (4.34)

We will need the following estimate for $b_1$. The proof follows from a similar procedure as for $\Gamma_t$ in Lemma 4.5 and is omitted.

**Lemma 4.7** Let $b_1$ be defined by (4.34). Then for any $p \geq 1$, we have

$$E \left[ \sup_{0 \leq \xi \leq T} |b_1(s, j)|^p \right] < \infty, \quad E \left[ \sup_{0 \leq \xi \leq T} |D_\xi b_1(s, j)|^p \right] < \infty$$

and

$$E \left[ \sup_{0 \leq \xi, \xi \leq T} |D^2_\xi b_1(s, j)|^p \right] < \infty.$$ 

Now we can state one of the main results in this paper.

**Theorem 4.3** Assume that Assumptions 3.1 and 3.3 hold. Let $X_t$ satisfy (1.1), let $Y_t$ satisfy (4.5) and let the partition $\pi$ satisfy $\kappa \cdot |\pi| < 1$. Then there is a constant $C$ independent of the partition $\pi$ but dependent on $\kappa, \mu, T$ such that

$$\sup_{0 \leq t \leq T} E \left[ |Y_t - X_t|^2 \right] \leq C |\pi|^2$$ (4.35)

and

$$\int_0^T E |Y_t - X_t|^2 dt \leq C |\pi|^2.$$ (4.36)
Remark 4.1 Our proof does not extend to the estimate of \( \sup_{0 \leq t \leq T} \mathbb{E} \left[ |Y_t - X_t|^p \right] \) for \( p \geq 1 \) since the use of an expansion of the second moment for the error process (see (4.45)). But we believe that the result can be improved to all \( p \geq 1 \) by a more careful analysis of the error process.

Proof of Theorem 4.3 We will first prove (4.35). The second statement (4.36) follows from (4.35) easily. The proof is divided into seven steps.

Step 1: A representation for the error process. We can write (1.1) as

\[
X_t = X_{t_k} + \int_{t_k}^{t} b(X_s) ds + (B_t - B_{t_k} - B_{t_k}^\tau) - R_k(t),
\]

where \( t_k \leq t \leq t_{k+1} \) and

\[
R_k(t) = \int_{t_k}^{t} b(X_s) - b(X_t) ds.
\]

For ease of notations, we denote by \( R_k := R_k(t_{k+1}) \) and \( R_k^\tau := R_k(t_{k+1} \wedge t) \).

Denote by the error process \( Z_t = X_t - Y_t \). For convenience we will also denote by \( Z_k = Z_{tk} \). Subtracting (4.5) from (4.37), we get for \( t_k \leq t \leq t_{k+1} \)

\[
Z_t = Z_k + (t - t_k)[b(X_t) - b(Y_t)] + R_k(t).
\]

We can rewrite the above formula as

\[
Z_t = Z_k + (t - t_k)[b(X_t) - b(Y_t)] + R_k(t).
\]

That is

\[
Z_t - (t - t_k) \left[ \sum_{j=0}^{k} A_j Z_j + A_k R_k(t) \right] = Z_k + R_k(t).
\]

Recall that \( A_t \) is defined in (4.30) and we denote by \( A_k = A_{t_k} \). Therefore, multiplying \( A_t \) on both sides of (4.39), we get

\[
Z_t = A_t Z_k + A_t R_k(t).
\]

Taking \( t = t_{k+1} \) particularly arrives at

\[
Z_{k+1} = A_{k+1} Z_k + A_{k+1} R_k.
\]

Denote by

\[
A_j^\tau = A_t \left( \prod_{\ell=j+1}^{k} A_{\ell} \right).
\]
where \( k \) is such that \( t_k < t \leq t_{k+1} \) and here we use the convention that if \( \{ \ell : j + 1 \leq \ell \leq k \} = \emptyset \) then \( \prod_{\ell = j + 1}^{k} A_{\ell} = 1 \). By iterating (4.41) and then using (4.40) and \( R_k^t = R_k(t_{k+1} \wedge t) \), we obtain

\[
Z_t = \sum_{j=0}^{k} A_j^t R_j^t. \tag{4.42}
\]

**Step 2: Estimate of \( A_j^t \).** Due to (4.31), we have

\[
|A_j^t| \leq (1 - \kappa(t - t_k))^{-1} \cdots (1 - \kappa \Delta_{j+1})^{-1} (1 - \kappa \Delta_j)^{-1} \leq e^{2\kappa(t-t_j)} \leq C. \tag{4.43}
\]

This can be used to obtain rate of convergence as in [17]. However, this rate will not be optimal. To obtain the optimal rate estimate, we need to bound the Malliavin derivative of \( A_j^t \) as in [7]. A straightforward computation for \( D_\xi A_j^t \) and \( D^2_{\xi, \zeta} A_j^t \) yields

\[
D_\xi A_j^t = D_\xi A_t \left( \prod_{\ell = j+1}^{k} A_\ell \right) + \sum_{\ell = j+1}^{k} A_t \left( A_{j+1} \cdots A_{\ell-1} \cdot D_\xi A_{\ell} \cdot A_{\ell+1} \cdots A_k \right),
\]

and

\[
D^2_{\xi, \zeta} A_j^t = D^2_{\xi, \zeta} A_t \left( \prod_{\ell = j+1}^{k} A_\ell \right) + \sum_{\ell = j+1}^{k} A_t \left( A_{j+1} \cdots A_{\ell-1} \cdot D^2_{\xi, \zeta} A_{\ell} \cdot A_{\ell+1} \cdots A_k \right) + \sum_{\ell = j+1}^{k} D_\xi A_t \left( A_{j+1} \cdots A_{\ell-1} \cdot D_\xi A_{\ell} \cdot A_{\ell+1} \cdots A_k \right) + \sum_{\ell = j+1}^{k} D_\xi A_t \left( A_{j+1} \cdots A_{\ell-1} \cdot D_\xi A_{\ell} \cdot A_{\ell+1} \cdots A_k \right) + \sum_{\ell \neq \ell', \ell, \ell' = j+1} A_t \left( A_{j+1} \cdots D_\xi A_{\ell} \cdots D_\xi A_{\ell'} \cdots A_k \right).
\]

These two formulas combined with (4.31), (4.33) and (4.43) yield

\[
\mathbb{E} \left[ \sup_{0 \leq \xi, t \leq T} |D_\xi A_j^t|^p \right] \leq C \quad \text{and} \quad \mathbb{E} \left[ \sup_{0 \leq \xi, \zeta, t \leq T} |D^2_{\xi, \zeta} A_j^t|^p \right] \leq C, \tag{4.44}
\]

for any \( p \geq 1. \)
Step 3: A decomposition for the error process. Now we come back to the main proof. Recall the Eq. (4.42). Thus we have

\[ |Z_t|^2 = \sum_{j=0}^{k} A_j R_j|^2 = \sum_{j,j'=0}^{k} A_j R_j \cdot A_j' R_j' =: \sum_{j,j'=0}^{k} I_{j,j'} \quad \text{(4.45)} \]

It is worthy to notice from (4.38) that

\[
R_j^r = \int_{t_j}^{t_j+1} \left( \int_0^1 \partial b(u X_{t_j+1} + (1-u) X_s) \, du \right) \Delta X_{s,t_j+1} \, ds
\]

\[
= \int_{t_j}^{t_j+1} b_1(s, j) \Delta X_{s,t_j+1} \, ds \quad \text{(4.46)}
\]

where \( \Delta X_{s,t} = X_t - X_s \) and recall that \( b_1 \) is defined by (4.34). So we can write

\[
I_{j,j'} = \int_{t_j}^{t_j+1} \int_{t_{j'}}^{t_{j'}+1} \left( A_j b_1(s, j) \Delta X_{s,t_j+1} \cdot A_j' b_1(s', j') \Delta X_{s',t_{j'}+1} \right) \, ds' \, ds \quad \text{(4.47)}
\]

Since

\[
\Delta X_{s,t_j+1} = X_{t_j+1} - X_s = \int_s^{t_j+1} b(X_u) \, du + \Delta B_{s,t_j+1} \quad \text{(4.47)}
\]

we can rewrite (4.47) as follows

\[
I_{j,j'} = \int_{t_j}^{t_j+1} \int_{t_{j'}}^{t_{j'}+1} \left[ A_j b_1(s, j) \left( \int_s^{t_j+1} b(X_u) \, du + \Delta B_{s,t_j+1} \right) \right.
\]

\[
\left. \cdot A_j' b_1(s', j') \left( \int_{s'}^{t_{j'+1}} b(X_v) \, dv + \Delta B_{s',t_{j'+1}} \right) \right] \, ds' \, ds \quad \text{(4.48)}
\]

where

\[
I_1 := \int_{t_j}^{t_j+1} \int_{t_{j'}}^{t_{j'+1}} \int_s^{t_j+1} \int_{s'}^{t_{j'+1}} \left( A_j b_1(s, j) b(X_u) \cdot A_j' b_1(s', j') b(X_v) \right) \, du \, dv \, ds' \, ds \;
\]

\[
I_2 := \int_{t_j}^{t_j+1} \int_{t_{j'}}^{t_{j'+1}} \int_s^{t_j+1} \int_{s'}^{t_{j'+1}} \left[ A_j b_1(s, j) b(X_u) \cdot A_j b_1(s', j') \Delta B_{s,t_{j'+1}} \right] \, du \, ds' \, ds \;
\]

\[
I_3 := \int_{t_j}^{t_j+1} \int_{t_{j'}}^{t_{j'+1}} \int_s^{t_j+1} \int_{s'}^{t_{j'+1}} \left[ A_j b_1(s, j) \Delta B_{s,t_{j'+1}} \cdot A_j' b_1(s', j') b(X_v) \right] \, dv \, ds' \, ds \;
\]

\[
I_4 := \int_{t_j}^{t_j+1} \int_{t_{j'}}^{t_{j'+1}} \left( A_j b_1(s, j) \Delta B_{s,t_{j'+1}} \cdot A_j' b_1(s', j') \Delta B_{s',t_{j'+1}} \right) \, ds' \, ds \;
\]

In the above, we have omitted the dependence of \( I_1, I_2, I_3, I_4 \) on \( j, j' \) for the sake of simplicity.
Step 4: Estimates of $I_1$. In the following we derive the upper-bounds for $I_1, \ldots, I_4$. First, let us consider $I_1$. From the polynomial growth assumption on $b$, Proposition 3.2, (4.43) and Lemma 4.7, we obtain that

$$
\mathbb{E} \left| A_j^i b_1(s, j) b(X_u) \cdot A_j^i b_1(s', j') b(X_v) \right| \leq C.
$$

Thus

$$
\mathbb{E} |I_1| \leq C \int_{t_j}^{t_{j+1} \wedge T} \int_{t_{j'}}^{t_{j'+1} \wedge T} \int_s^{t_{j+1} \wedge T} \int_{s'}^{t_{j'+1} \wedge T} \mathrm{d}u \mathrm{d}s' \mathrm{d}s \leq C \Delta_j^2 \Delta_{j'}^2. \tag{4.49}
$$

Step 5: Estimate of $I_2$ and $I_3$. We turn to $I_2$. From (4.43), (4.44), Lemma 4.7, the polynomial growth assumption on $b$, Proposition 3.2 and Proposition 3.4, we see easily (by using the product rule of Malliavin derivative) that

$$
\mathbb{E} \left[ \sup_{s,s',t,u,j,j',\xi} \left| D_\xi \left( A_j^i b_1(s, j) b(X_u) \cdot A_j^i b_1(s', j') \right) \right| \right] < C.
$$

Thus by Lemma 4.3, we get

$$
\mathbb{E} \left| A_j^i b_1(s, j) b(X_u) \cdot A_j^i b_1(s', j') \Delta B_{s',t_{j'+1} \wedge T} \right| \leq C \Delta_{j'}.
$$

Consequently, we have

$$
\mathbb{E} |I_2| \leq C \Delta_{j'} \int_{t_j}^{t_{j+1} \wedge T} \int_{t_{j'}}^{t_{j'+1} \wedge T} \int_s^{t_{j+1} \wedge T} \int_{s'}^{t_{j'+1} \wedge T} \mathrm{d}u \mathrm{d}s' \mathrm{d}s \leq C \Delta_j^2 \Delta_{j'}^2. \tag{4.50}
$$

Exactly in the same way, we obtain

$$
\mathbb{E} |I_3| \leq C \Delta_j^2 \Delta_{j'}^2. \tag{4.51}
$$

Step 6: Estimate of $I_4$. Denote by

$$
F(s, s', t, j, j') := A_j^i b_1(s, j) \cdot A_j^i b_1(s', j').
$$

Similar to the above argument we can show that

$$
\mathbb{E} \left[ \sup_{s,s',t,j,j',\xi,\zeta} \left( |F(s, s', t, j, j')| + |D_\xi F(s, s', t, j, j')| + |D_\xi^2 F(s, s', t, j, j')| \right) \right] \leq C.
$$
From Lemma 4.3 it follows then

\[
\mathbb{E}|I_4| \leq C \int_{t_j}^{t_{j+1}} \int_{t_{j'}}^{t_{j'+1}} \left| \mathbb{E} \left( F(s, s', t, j, j') \Delta B_{s, t_{j+1}} \Delta B_{s', t_{j'+1}} \right) \right| ds' ds
\]

\[
\leq C \int_{t_j}^{t_{j+1}} \int_{t_{j'}}^{t_{j'+1}} \left[ \langle 1(s, t, j), 1(s', t, j) \rangle \mathcal{H} + |t_{j+1} - s| \cdot |t_{j'+1} - s'| \right] ds' ds
\]

\[
\leq C \Delta_j \Delta_j' \langle 1(t, t_{j+1}), 1(t', t_{j'+1}) \rangle \mathcal{H} + C \Delta_j^2 \Delta_j'^2.
\]

(4.52)

**Step 7: Conclusion.** Plugging (4.48)–(4.52) into (4.45), we conclude that

\[
\mathbb{E} \left[ |Z_t|^2 \right] \leq C \sum_{j, j'=0}^{n-1} \Delta_j^2 \Delta_j'^2 + C \Delta_j \Delta_j' \sum_{j, j'=0}^{n-1} \langle 1(t, t_{j+1}), 1(t', t_{j'+1}) \rangle \mathcal{H} \leq C |\pi|^2.
\]

This proves the theorem. \(\square\)

## 5 Limiting distribution

In this section, we consider the asymptotic error distribution of the backward Euler scheme. For convenience, we consider the equidistant discretization only, namely we consider the partition \(\pi\) defined in (4.1) such that \(|\pi| = T/n\) and \(t_k = k|\pi|, k = 0, 1, \ldots, n\). Our result shows that the asymptotic error \(nZ_t\) converges in \(L_p\) to a nonzero limit for \(p \in [1, 2)\). In particular, the exact rate of the backward Euler scheme is \(1/n\).

### 5.1 A linear ordinary differential equation

Let \(\phi_t : [0, T] \rightarrow \mathbb{R}^{m \times m}\) be the solution of the linear differential equation

\[
\phi_t = \text{Id} + \int_0^t \partial b(X_s)\phi_s ds.
\]

(5.1)

We have the following estimate for (5.1).

**Lemma 5.1** Let \(\phi_t\) be defined by (5.1). For all \(s, t : 0 \leq s < t \leq T\), we have relations

\[
|\phi_t| \leq e^{\lambda(t-s)} |\phi_s| \quad \text{and} \quad \sup_{0 \leq t \leq T} |\phi_t'| \leq \kappa \sup_{0 \leq t \leq T} (1 + |X_t|^\mu) \cdot \sup_{0 \leq t \leq T} |\phi_t|.
\]

(5.2)

**Proof** Take \(s, t : 0 \leq s < t \leq T\). Let \(\phi_{i,t}\) be the \(i\)th column of \(\phi_t\), \(i = 1, 2, \ldots, m\). Using (5.1), we can write that

\[
\partial f(x) = \text{Id} - \partial b(x) \cdot \Delta_k.
\]
As in the proof of Lemma 4.2, we can show that $\partial b(X_u)$ satisfies condition (4.6) with $\lambda = \kappa$. Therefore, we get
\[
|\phi_{i,t}|^2 \leq |\phi_{i,s}|^2 + 2\lambda \int_{s}^{t} |\phi_{i,u}|^2 du.
\]
Gronwall’s inequality then yields
\[
|\phi_{i,t}|^2 \leq |\phi_{i,s}|^2 + 2\lambda \int_{s}^{t} |\phi_{i,u}|^2 du.
\]
This gives the first inequality in (5.2). Differentiating the $i$th component of $\phi_t$ in (5.1) and using Assumption 3.1 (A2), it is readily checked that the second inequality in (5.2) holds. This completes the proof of the lemma.

Now we want to approximate equation (5.1) by the following “forward-backward” Euler scheme
\[
\phi_\pi^\tau = \phi_{tk} + \partial b(X_{tk})\phi_{tk} \cdot (t - tk), \quad \phi_\pi^0 = \text{Id}.
\]
In particular, when $t = tk+1$, we have
\[
\phi_{tk+1}^\tau = \phi_{tk} + \partial b(X_{tk})\phi_{tk+1} \Delta_k, \quad \phi_0^\tau = \text{Id}.
\]
Note that solving (5.3) for $\phi_t^\tau$, we have
\[
\phi_t^\tau = (I - \partial b(X_{t\ell})(t - t\ell))^{-1} \phi_{tk} = \tilde{A}_t \cdot \phi_{tk}^\tau,
\]
where $t_\ell < t \leq t_{\ell+1}$, $\ell = 0, \ldots, n - 1$, and we define
\[
\tilde{A}_t = (I - \partial b(X_{t\ell})(t - t\ell))^{-1}.
\]

**Proposition 5.1** Let $\phi_t$ and $\phi_t^\tau$ be defined by (5.1) and (5.3), for any $0 < \alpha < H$ and $p > 1$, we have the following $L^p$-convergence
\[
\mathbb{E} \left[ \sup_{0 \leq t \leq T} |\phi_t^\tau - \phi_t|^p \right] \leq C |\tau|^\alpha p.
\]

**Proof** On the subinterval $t \in (tk, tk+1]$, we write
\[
\phi_t = \phi_{tk} + \int_{tk}^{t} \partial b(X_s)\phi_s ds
\]
\[
= \phi_{tk} + \partial b(X_{tk})\phi_t \cdot (t - tk) + \tilde{R}_t,
\]
where \( t_k < t \leq t_{k+1} \), and we denote by

\[
\tilde{R}_t = \int_{t_k}^t \partial b(X_s) \phi_s \, ds - \int_{t_k}^t \partial b(X_{t_k}) \phi_t \, ds.
\]

Thus, we have

\[
\phi_t = (\text{Id} - \partial b(X_{t_k})(t - t_k))^{-1} \left( \phi_{t_k} + \tilde{R}_t \right) = \tilde{A}_t \phi_{t_k} + \tilde{A}_t \tilde{R}_t. \tag{5.6}
\]

Denote by \( z_t = \phi_t - \phi^\pi_t \) and \( z_k = \phi_{t_k} - \phi^\pi_{t_k} \). Then, similar to (4.40), by taking difference between (5.4) and (5.6) we obtain

\[
z_t = \tilde{A}_t z_k + \tilde{A}_t \tilde{R}_t. \tag{5.7}
\]

Computing (5.7) recursively leads to

\[
z_t = \tilde{A}_t \tilde{R}_t + \tilde{A}_t \tilde{A}_t \tilde{R}_t + \cdots + \tilde{A}_t \tilde{A}_t \cdots \tilde{A}_t \tilde{R}_t. \tag{5.8}
\]

Rewrite \( \tilde{R}_t \) as

\[
\tilde{R}_t = \int_{t_k}^t \left( \partial b(X_s) - \partial b(X_{t_k}) \right) \phi_s \, ds + \int_{t_k}^t \partial b(X_{t_k}) (\phi_s - \phi_t) \, ds. \tag{5.9}
\]

Now apply Assumption 3.1 (A2), relation (3.10) and relations in (5.2) to the first integral in (5.9), we derive

\[
\left| \int_{t_k}^t \left( \partial b(X_s) - \partial b(X_{t_k}) \right) \phi_s \, ds \right| \leq C|\pi|^{1+\alpha} \left( 1 + \sup_{0 \leq s \leq T} |X_s|^\mu \right) \cdot \left( 1 + \sup_{0 \leq s \leq T} |X_s|^\mu + \|B\|_{\alpha} \right).
\]

On the other hand, we apply Assumption 3.1 (A2) and relation (5.2) again to the second integral to get

\[
\left| \int_{t_k}^t \partial b(X_{t_k}) (\phi_s - \phi_t) \, ds \right| \leq C|\pi|^2 \left( 1 + \sup_{0 \leq s \leq T} |X_s|^\mu \right)^2.
\]

Combining the above two estimates, we have

\[
|\tilde{R}_t| \leq C|\pi|^{1+\alpha} \left( 1 + \sup_{0 \leq s \leq T} |X_s|^\mu + \|B\|_{\alpha} \right)^2. \tag{5.10}
\]
Let us recall the Lemma 4.2. We take \( y = 0 \) and assume that \( \kappa \cdot |\pi| < 1 \). Then we have the following estimate for \( \tilde{A}_t \)

\[
|\tilde{A}_t| \leq (1 - \kappa (t - t_\ell))^{-1} \leq e^{2\kappa (t - t_\ell)}. \tag{5.11}
\]

Substituting the estimates (5.10) and (5.11) into (5.8), it has

\[
|z_t| \leq C|\pi|^\alpha \left( 1 + \sup_{0 \leq s \leq T} |X_s|^\mu + \|B\|_\alpha \right)^2,
\]

where \( C \) is a constant with respect to \( \kappa, \mu, H, T \) but independent of the partition \( \pi \). According to the boundness of \( X_t \), see (3.6), we obtain earlier proof of the proposition.

\[ \square \]

Take \( t_k < t \leq t_{k+1} \) and \( t_j \leq t_k \) and denote by \( \phi_{t_j}(t) = \phi_t \phi_{t_j}^{-1} \). According to the definition (5.1) of \( \phi_t \), we have

\[
\phi_{t_j}(t) = \text{Id} + \int_{t_j}^t \partial b(X_s)\phi_{t_j}(s)ds, \quad t_j \leq t \leq T. \tag{5.12}
\]

On the other hand, using the expression (5.3), we get for \( \phi_{t_j}^\pi(t) := \phi_t^\pi (\phi_{t_j}^\pi)^{-1} \)

\[
\phi_{t_j}^\pi(t) = \text{Id} + \int_{t_j}^t \partial b(X_{tk})\phi_{t_j}^\pi(s)ds, \quad t_j \leq t_k < t \leq t_{k+1} \leq T. \tag{5.13}
\]

Therefore, in a similar way as for Proposition 5.1, we can show the following.

**Proposition 5.2** Let \( \phi_{t_j}(t) \) and \( \phi_{t_j}^\pi(t) \) be defined by (5.12) and (5.13) separately. Then for any \( 0 < \alpha < H \) and \( p > 1 \), we have the following \( L^p \)-convergence

\[
\mathbb{E} \left[ \sup_{0 \leq t_j < t \leq T} |\phi_{t_j}^\pi(t) - \phi_{t_j}(t)|^p \right] \leq C|\pi|^\alpha p.
\]

**5.2 The asymptotic error of the backward Euler scheme**

In this subsection, we prove our main result on the asymptotic error of the backward Euler scheme. The proof is based on the representation (4.42) of the error process \( Z_t \) and some approximation results for \( R_{t_j} \) and \( A_{t_j} \).

We first prove the following lemma.

**Lemma 5.2** Recall that \( R_k = R_k(t_{k+1}) = \int_{t_k}^{t_{k+1}} b(X_s) - b(X_{t_{k+1}})ds \) is defined by (4.38). Let

\[
\hat{R}_k = R_k + R_{1k} + R_{2k}, \tag{5.14}
\]

\[ \square \] Springer
where we define

\[ R_{1k} := \partial bb(X_{t_k}) \int_{t_k}^{t_{k+1}} (t_{k+1} - s) \, ds \quad \text{and} \quad R_{2k} := \partial b(X_{t_k}) \int_{t_k}^{t_{k+1}} (B_{t_{k+1}} - B_s) \, ds, \]

and we denote by

\[ \partial bb(x) = \sum_{i=1}^{m} \frac{\partial b}{\partial x_i}(x) b_i(x). \]

Then for any \( 1/2 < \alpha < H \), we have

\[ \left| \hat{R}_k \right| \leq G_\alpha \Delta_k^{2\alpha + 1}, \tag{5.15} \]

where \( G_\alpha \) is some random variable independent of \( n \) and \( k \) such that \( \| G_\alpha \|_{L^p(\Omega)} < \infty \) for any \( p > 0 \).

**Proof** Using Itô formula and (1.1), we can write

\[ b(X_{t_{k+1}}) - b(X_s) = \int_s^{t_{k+1}} \partial b(X_u) dB_u - \int_s^{t_{k+1}} \partial(bb)(X_u) \, du + \int_s^{t_{k+1}} \partial b(X_u) du. \tag{5.16} \]

Let us denote by

\[ r_k(s) = \int_s^{t_{k+1}} \partial bb(X_u) - \partial(bb)(X_{t_k}) du + \int_s^{t_{k+1}} \partial b(X_u) - \partial b(X_{t_k}) dB_u. \tag{5.17} \]

Then using (5.16), it is clear that

\[ b(X_{t_{k+1}}) - b(X_s) = \partial bb(X_{t_k})(t_{k+1} - s) + \partial b(X_{t_k})(B_{t_{k+1}} - B_s) + r_k(s). \tag{5.18} \]

Integrating both sides of Eq. (5.18) and recalling the definition of \( \hat{R}_k \) in (5.14), we get

\[ \hat{R}_k = - \int_{t_k}^{t_{k+1}} r_k(s) \, ds. \tag{5.19} \]

To show (5.15) it suffices to prove a bound for \( r_k(s) \). We apply Lemma 2.2 to the second integral of the right side of (5.17) to get for any \( \alpha \in (1/2, H) \)

\[ |r_k(s)| \leq \| \partial bb(X) \|_\alpha \Delta_k^{1+\alpha} + \| \partial b(X) \|_\alpha \cdot \| B \|_\alpha \Delta_k^{2\alpha}. \]

Thanks to Assumption 3.1 (A2) and Lemma 3.1 for \( X \), we obtain that \( \| \partial bb(X) \|_\alpha \) and \( \| \partial b(X) \|_\alpha \cdot \| B \|_\alpha \) are bounded by some random variable \( G_\alpha \) which has any finite
moment. That is, we have the estimate
\[
\sup_{t_k \leq s \leq t_{k+1}} |r_k(s)| \leq G u \Delta_k^{2\alpha}.
\]

Applying the above estimate to (5.19), we obtain (5.15).

Let us turn to the estimate of \(A_t^j\).

**Lemma 5.3** For any \(p \in [1, 2)\) and any \(0 < \alpha < H\), we have
\[
\sup_{0 \leq t_j < t \leq T} \mathbb{E} \left[ |A_t^j - \phi_t^\pi (\phi_t^j)^{-1}|^p \right] \leq C|\pi|^{\alpha p}.
\] (5.20)

**Proof** By iterating (5.4), we observe that
\[
\phi_t^\pi = \tilde{A}_t \prod_{\ell=j+1}^k \tilde{A}_\ell \cdot \phi_t^j.
\] (5.21)

So multiplying \((\phi_t^j)^{-1}\) on both sides of (5.21) leads to
\[
\phi_t^\pi (\phi_t^j)^{-1} = \tilde{A}_t \prod_{\ell=j+1}^k \tilde{A}_\ell.
\]

On the other hand, recall the expression of \(A_t^j\) (see (4.30) and (4.41))
\[
A_t^j = A_t \left( \prod_{\ell=j+1}^k A_\ell \right) = A_t \prod_{\ell=j+1}^k \left( \text{Id} - \Delta_{\ell-1} \int_0^1 \partial b(uX_{\ell t} + (1-u)Y_{\ell t})du \right)^{-1}.
\]

Thus, applying the estimates (4.31) and (5.11) of \(A_t\) and \(\tilde{A}_t\), we get
\[
|A_t^j - \phi_t^\pi (\phi_t^j)^{-1}| \leq |A_t - \tilde{A}_t| \prod_{\ell=j+1}^k |A_\ell| + |\tilde{A}_t| \sum_{i=j+1}^k |A_{j+1}| \cdots |A_{i-1}| \cdot |A_i - \tilde{A}_i| \cdot |\tilde{A}_{i+1}| \cdots |\tilde{A}_k| \leq C|A_t - \tilde{A}_t| + C \sum_{i=j+1}^k |A_i - \tilde{A}_i|.
\] (5.22)

It follows from the definitions of \(A_t\) in (4.30) and \(\tilde{A}_t\) in (5.5) that for all \(t \in [0, T]\)
\[
\tilde{A}_t^{-1} - A_t^{-1} = (t - \eta(t)) \left( \partial b(X_{\eta(t)}) - \int_0^1 \partial b(uX_t + (1-u)Y_t)du \right).
\]
Therefore, we have

\[ |A_t - \tilde{A}_t| = |\tilde{A}_t (\tilde{A}_t^{-1} - A_t^{-1}) A_t| \]
\[ \leq \Delta_i^{-1} |\tilde{A}_t| \cdot \left| \partial b(X_{\eta(t)}) - \int_0^1 \partial b(uX_t + (1 - u)Y_t) du \right| \cdot |A_t|. \]

Applying Assumption 3.1 (A2), Lemma 4.4, and (5.11) we thus obtain

\[ |A_t - \tilde{A}_t| \leq C \Delta_i^{-1} \left( 1 + |X_t|^{\mu} + |Y_t|^{\mu} \right) \left( |X_t - Y_t| + |X_t - X_{\eta(t)}| \right). \]

Applying (3.6), (4.9), the convergence result (4.35) and the first relation in (3.10), we get for all \( p \in [1, 2) \)

\[
\mathbb{E} \left[ |A_t - \tilde{A}_t|^p \right] \leq C \Delta_i^{-1} \mathbb{E} \left[ \left( 1 + |X_t|^{\mu p} + |Y_t|^{\mu p} \right) \left( |X_t - Y_t|^p + |X_t - X_{\eta(t)}|^p \right) \right] 
\leq C \Delta_i^{-1} \left[ \mathbb{E} \left( 1 + |X_t|^{\frac{2\mu p}{2-p}} + |Y_t|^{\frac{2\mu p}{2-p}} \right)^{\frac{2-p}{2}} \cdot \left( \mathbb{E} (|X_t - Y_t|^2 + |X_t - X_{\eta(t)}|^2) \right)^{\frac{p}{2}} \right] 
\leq C \Delta_i^{-1} |\pi|^{\alpha p}.
\]

for any \( 0 < \alpha < H \), where in the second inequality in (5.23) we applied the Hölder’s inequality with the conjugate pair \((2/(2 - p), 2/p)\). Finally, substituting (5.23) into (5.22), we conclude the estimate (5.20). The proof is now complete.

Combining Lemma 5.3 with Proposition 5.2, we obtain the following lemma.

**Lemma 5.4** For any \( p \in [1, 2) \) and any \( 0 < \alpha < H \), we have

\[ \sup_{0 \leq t_j < t \leq T} \mathbb{E} \left[ \left| A_{t_j} - \phi_{t_j}^{-1} \phi_t \right|^p \right] \leq C |\pi|^{\alpha p}. \]

Now we state the main theorem of this section.

**Theorem 5.3** For any \( p \in [1, 2) \) and any \( 1/2 < \alpha < H \), we have the following convergence in \( L^p \) as \( |\pi| \to 0 \)

\[ nZ_t \to \hat{U}_t := \frac{1}{2} \int_0^t \phi_t \phi_s^{-1} \partial bb(X_s) ds + \frac{1}{2} \int_0^t \phi_t \phi_s^{-1} \partial b(X_s) dB_s. \quad (5.24) \]

Moreover, \( \hat{U}_t \) satisfies

\[ d\hat{U}_t = \partial b(X_t) \hat{U}_t dt + \frac{1}{2} \partial bb(X_t) dt + \frac{1}{2} \partial b(X_t) dB_t. \quad (5.25) \]
Proof Let \( k \) be such that \( t_k < t \leq t_{k+1} \). Recalling relation (4.42), we have

\[
nZ_t = \sum_{j=0}^{k-1} A'_j (nR_j) + A'_k (nR'_k).
\]

Recall the representation (5.14) for \( R_k \). We consider the following decomposition of \( nZ_t \):

\[
nZ_t = \sum_{j=0}^{k-1} \left( A'_j - \phi_t \phi_{t_j}^{-1} \right) (nR_j) + \sum_{j=0}^{k-1} \phi_t \phi_{t_j}^{-1} (n \hat{R}_j) - \sum_{j=0}^{k-1} \phi_t \phi_{t_j}^{-1} (nR_{1j})

- \sum_{j=0}^{k-1} \phi_t \phi_{t_j}^{-1} (nR_{2j}) + A'_k (nR'_k)

:= \hat{I}_1 + \hat{I}_2 + \hat{I}_3 + \hat{I}_4 + \hat{I}_5. \tag{5.26}
\]

In the following, we consider the convergence of \( \hat{I}_i \), \( i = 1, \cdots, 5 \), individually.

We first consider the convergence of \( \hat{I}_5 \). Note that by applying Assumption (A2) and the Hölder continuity of \( X_t \) in (3.10) to (4.38), we get

\[
E |R'_k|^p \leq C |\pi|^{(1+\alpha)p}. \tag{5.27}
\]

So, taking into account the estimate of \( A'_j \) in (4.43), we have

\[
E |\hat{I}_5|^p = E \left| A'_k (nR'_k) \right|^p < C |\pi|^{\alpha p}.
\]

This implies the convergence \( \hat{I}_5 \to 0 \) as \( |\pi| \to 0 \).

For the convergence of \( \hat{I}_1 \), we apply Lemma 5.4 and relation (5.27). This yields

\[
E |\hat{I}_1|^p \leq C \sum_{j=0}^{k-1} E \left| \left( A'_j - \phi_t \phi_{t_j}^{-1} \right) (nR_j) \right|^p \leq C \sum_{j=0}^{k-1} |\pi|^{2\alpha p} \leq C |\pi|^{2\alpha p-1}.
\]

Since \( 2\alpha p - 1 > 0 \) for \( 1/2 < \alpha < H \), we conclude that \( \hat{I}_1 \to 0 \) as \( |\pi| \to 0 \).

For \( \hat{I}_2 \), we apply Lemmas 5.1 and 5.2 to get

\[
E |\hat{I}_2|^p \leq C \sum_{j=0}^{k-1} E \left| \phi_t \phi_{t_j}^{-1} (n \hat{R}_j) \right|^p \leq C \sum_{j=0}^{k-1} |\pi|^{2\alpha p} \leq C |\pi|^{2\alpha p-1}.
\]

Because \( 2\alpha p - 1 > 0 \) when \( 1/2 < \alpha < H \), we get \( \hat{I}_2 \to 0 \) as \( |\pi| \to 0 \).
Let us consider the convergence of $\hat{I}_3$ and $\hat{I}_4$. According to the definition of $R_{1j}$, we have

$$\hat{I}_3 = \sum_{j=0}^{k-1} \phi_t \phi_{t_j}^{-1} \cdot \partial b(X_{t_j}) \cdot \frac{n}{2} (t_{j+1} - t_j)^2 = \frac{1}{2n} \sum_{j=0}^{k-1} \phi_t \phi_{t_j}^{-1} \partial b(X_{t_j}).$$

Therefore by the continuity of $\phi_t \phi_{s}^{-1} \partial b(X_s)$ in $L_p$, we obtain

$$\hat{I}_3 \to \frac{1}{2} \int_0^t \phi_t \phi_{s}^{-1} \partial b(X_s) \, ds$$

in $L^p$ as $|\pi| \to 0$.

For $\hat{I}_4$, we write it as follows

$$\hat{I}_4 = \sum_{j=0}^{k-1} n \phi_t \phi_{t_j}^{-1} \partial b(X_{t_j}) \int_{t_j}^{t_{j+1}} \int_s^{t_{j+1}} dB_u ds = n \sum_{j=0}^{k-1} \phi_t \phi_{t_j}^{-1} \partial b(X_{t_j}) \int_{t_j}^{t_{j+1}} (u - t_j) dB_u.$$

It follows from Lemma 3.1 and Assumption 3.1 (A2) for $\partial b$ that the process $\phi_s^{-1} \partial b(X_s), s \in [0, T]$ is Hölder continuous of order $H$ in $L^p$ for $p \geq 1$. We can thus apply [7, Corollary 7.2] to get the convergence in $L^p$ as $|\pi| \to 0$

$$\hat{I}_4 \to \frac{1}{2} \int_0^t \phi_t \phi_{s}^{-1} \partial b(X_s) dB_s.$$

In summary of the convergence of $\hat{I}_i, i = 1, \cdots, 5$, we conclude the convergence (5.24). It is straightforward to check that $\hat{U}_t$ satisfies (5.25).

6 Numerical experiments

In this section, we validate our theoretical results by performing two numerical experiments. In the following $B_t, t \geq 0$ denotes by a standard fBm with Hurst parameter $H \in (0, 1)$.

Example 6.1 We consider the following one-sided Lipschitz stochastic differential equation

$$dX_t = -X_t^3 dt + dB_t,$$

with $X_0 = 5$ and Hurst parameter $H = 0.6$. In this example, we compare the performance of three numerical methods, namely the backward Euler method (BEM), the forward Euler-Maruyama method (EM) and the Crank-Nicolson method (CN).

We use the backward Euler method with step size 0.0001 to compute a numerical substitute for the sample of the exact solution. For the simulation of the three numerical
methods we consider two step sizes: 0.02 and 0.08. Tables 1 and 2 show the results for the sampling in two different step sizes. The exact solution and the three numerical approximations of Eq. (6.1) are evaluated at \( T = 0.1, 0.2, \ldots, 0.8 \). Clearly, the step size 0.08 is too large for the Euler-Maruyama method and the Crank-Nicolson method on this one-sided Lipschitz problem. On the other hand, when the step size is cut to 0.02 all three methods perform well.

**Example 6.2** Let \( B_t = (B^1_t, B^2_t) \) be a two-dimensional fBm with Hurst parameter \( H \in (0, 1) \) and assume that \( B^1_t \) and \( B^2_t \) are mutually independent. In this example we consider the two-dimensional stochastic differential equation studied in \([18]\):

\[
\begin{align*}
\quad & dX_t = \left( X_t - Y_t - X^3_t - X_t Y^2_t \right) dt + dB^1_t, \\
\quad & dY_t = \left( X_t + Y_t - X^2_t Y_t - Y^3_t \right) dt + dB^2_t,
\end{align*}
\]  

(6.2)

with the initial value \( X_0 = 1, Y_0 = 1 \) for \( t \in [0, 1] \). It is easy to verify that Eq. (6.2) satisfies the one-sided Lipschitz condition.

To find the convergence rate of the backward Euler method, in our numerical computation we take the time step sizes \( 2^{-k}, k = 5, 6, \ldots, 9 \). As in the previous Example 6.1 the numerical solution with a much smaller step size (in this example we take the step size \( 2^{-11} \)) is used to represent the reference solution. We perform \( M = 1000 \) simulations, and we compute the mean square error by

\[
\epsilon_n = \sqrt{\frac{1}{M} \sum_{\ell=1}^{M} (|X^\ell_n - X^\ell|^2 + |Y^\ell_n - Y^\ell|^2)},
\]

where \( X^\ell \) and \( Y^\ell \) denote by the \( \ell \)-th exact solution, \( X^\ell_n \) and \( Y^\ell_n \) denote by the \( \ell \)-th numerical solution with step size \( |\pi| = 1/n \). By calculating the convergence order \( \mathcal{O} \) as defined by

\[
\mathcal{O} = \log_2 \left( \frac{\epsilon_n}{\epsilon_{2n}} \right),
\]

accurate numerical values of error and convergence orders under different step sizes are listed in Table 3. For the four subfigures in Fig. 1, the red dotted reference line has
Table 3  Mean squared of errors and convergence rates of backward Euler method for SDE (6.2)

| $|\pi|$ | $H = 0.6$ | $H = 0.7$ | $H = 0.8$ | $H = 0.9$ |
|------|--------|--------|--------|--------|
|      | Error  | $O$    | Error  | $O$    | Error  | $O$    | Error  | $O$    |
| $1/2^5$ | 3.3335e$-2$  | –      | 2.2177e$-2$  | –      | 1.7053e$-2$  | –      | 1.4462e$-2$  | –      |
| $1/2^6$ | 1.5464e$-2$  | 1.1081 | 1.0176e$-2$  | 1.1239 | 8.0469e$-3$  | 1.0835 | 6.9912e$-3$  | 1.0487 |
| $1/2^7$ | 7.1006e$-3$  | 1.1229 | 4.7004e$-3$  | 1.1143 | 3.8213e$-3$  | 1.0744 | 3.3680e$-3$  | 1.0536 |
| $1/2^8$ | 3.1505e$-3$  | 1.1724 | 2.1154e$-3$  | 1.1519 | 1.7612e$-3$  | 1.1175 | 1.5667e$-3$  | 1.1042 |
| $1/2^9$ | 1.3181e$-3$  | 1.2571 | 8.9192e$-4$  | 1.2459 | 7.5221e$-4$  | 1.2274 | 6.7140e$-4$  | 1.2225 |

Fig. 1 The strong mean square convergence error of approximation of SDE (6.2)

(a) $H = 0.6$  (b) $H = 0.7$

(c) $H = 0.8$  (d) $H = 0.9$

Thus we numerically verify that our theoretical results are in accordance with numerical results.

Acknowledgements  The authors wish to thank the referees for many useful comments and suggestions.
Declarations

Conflict of interest  The authors declare that they have no known competing financial interests that are directly related to the work submitted for publication.

7 Appendix

In this section, we consider the differential equation

$$x_t = x_{t_0} + \int_{t_0}^t b(s, x_s)ds, \quad t \geq t_0,$$

(7.1)

where we assume that

(i) The function $b : [t_0, \infty) \times \mathbb{R}^m \to \mathbb{R}^m$ is one-sided Lipschitz: there is a constant $\kappa > 0$ such that

$$\langle x - y, b(s, x) - b(s, y) \rangle \leq \kappa |x - y|^2, \quad \forall x, y \in \mathbb{R}^m \text{ and } s \geq t_0.$$

(ii) Both $b$ and $\partial b$ are continuous functions in $[t_0, \infty) \times \mathbb{R}^m$. Here $\partial b$ denotes by the Jacobian $\frac{\partial b(s, x)}{\partial x}$ of $b$ as we defined in Eq. (3.1).

Example 7.1  It is easy to show that the following two equations are examples of (7.1) and the conditions (i)–(ii) are satisfied.

(1) The equation

$$y'_t = b(g_t + y_t),$$

where $b$ is one-sided Lipschitz and $g$ is a continuous function. In fact, in this case $b(s, x) = b(g_s + x)$ and we have

$$\langle x - y, b(s, x) - b(s, y) \rangle = \langle (g_s + x) - (g_s + y), b(g_s + x) - b(g_s + y) \rangle$$

$$\leq \kappa |x - y|^2, \quad \forall x, y \in \mathbb{R}^m \text{ and } s \geq t_0.$$

(2) The linear equation

$$x'_t = U_t x_t,$$

where the eigenvalues of $U_t$ are uniformly bounded for all $t \geq 0$.

The following existence and uniqueness results about Eq. (7.1) are mostly well-known. We did not find the exact explicit result for (7.1) in the literature and so we include a proof for the sake of completeness.
Proposition 7.1 There exists a unique solution $x$ to Eq. (7.1), and the solution satisfies the relation

$$|x_t|^2 \leq \left( |x_{t_0}|^2 + \int_{t_0}^{t} |b(s, 0)|^2 \, ds \right) \cdot e^{(2\kappa + 1)(t-t_0)}, \quad t \in [t_0, \infty).$$

Proof We start by defining

$$\tau = \sup \{ t : \text{there exists a solution to equation (7.1) on } [t_0, t] \}.$$

We first note that by Peano’s theorem it is easy to see that $\tau > t_0$. On the other hand, the one-sided Lipschitz condition implies that for any $t < \tau$ there exists a unique solution on the interval $[t_0, t]$ (see e.g. [19, Lemma 12.1]).

In the following we show that $\tau = \infty$ by contradiction, which then concludes the existence of the solution on $[t_0, \infty)$.

We calculate

$$|x_t|^2 = |x_{t_0}|^2 + 2 \int_{t_0}^{t} (x_s, b(s, x_s)) \, ds$$

$$= |x_{t_0}|^2 + 2 \int_{t_0}^{t} (x_s, b(s, x_s) - b(s, 0)) \, ds + 2 \int_{t_0}^{t} (x_s, b(s, 0)) \, ds$$

$$\leq |x_{t_0}|^2 + 2\kappa \int_{t_0}^{t} |x_s|^2 \, ds + \int_{t_0}^{t} |x_s|^2 \, ds + \int_{t_0}^{t} |b(s, 0)|^2 \, ds.$$

Applying Gronwall’s inequality yields

$$|x_t|^2 \leq \left( |x_{t_0}|^2 + \int_{t_0}^{t} |b(s, 0)|^2 \, ds \right) e^{(2\kappa + 1)(t-t_0)}$$

$$\leq \left( |x_{t_0}|^2 + \int_{t_0}^{\tau} |b(s, 0)|^2 \, ds \right) e^{(2\kappa + 1)(\tau-t_0)}.$$

This implies that $x_t$ is bounded on $[t_0, \tau)$. Since $x_t' = b(t, x_t)$ we obtain that $x$ is uniformly continuous on the interval $[t_0, \tau)$. In particular, the limit $\lim_{t \to \tau} = x_{\tau}$ exists. This shows that $x$ is a solution to Eq. (7.1) on $[t_0, \tau]$.

Now consider the equation $x_t' = b(t, x_t)$ with initial value $x_\tau$. Peano’s theorem implies that the equation has a solution on the interval $[\tau, \tau + \varepsilon]$ for some $\varepsilon > 0$. Combining the two functions $(x_t, t \in [t_0, \tau])$ and $(x_t, t \in [\tau, \tau + \varepsilon])$, we obtain a solution to Eq. (7.1) on the interval $[t_0, \tau + \varepsilon]$. This contradicts the definition of $\tau$. We conclude that $\tau = \infty$. The proof is complete. \qed

Our next result addresses the Malliavin differentiability of Eq. (1.1). The proof is an application of [20, Lemma 1.2.3]. Recall that the Malliavin derivative is defined in Sect. 2.1.

Lemma 7.1 Let $X_t$ be the solution of Eq. (1.1). For any $t \geq 0$, the Malliavin derivative $D X_t$ of $X_t$ exists in $\mathbb{D}^{1,2}$. 

\vspace{1cm}
Proof Let $Y_t^\pi$ be defined by (4.5). Then for any fixed $t \in [0, T]$, we can show that $Y_t^\pi$ converges to $X_t$ in $L^2$. This can be done by following the method of Sect. 4 without the use of Malliavin calculus. Indeed, bounding the right-hand side of (4.46) we get that $E|R_t^j|^2 \leq C|\pi|^{2\alpha+2}$ for $0 < \alpha < H$. Applying this bound and also the bound $|A_t^j| \leq C$ in (4.43) to the expression of $Z_t = X_t - Y_t^\pi$ in (4.42), we obtain that

$$Y_t^\pi \to X_t \quad \text{in} L^2 \quad \text{as} |\pi| \to 0. \quad (7.2)$$

From the first inequality of (4.15) it follows that

$$\sup_{|\pi|} E\left[\|DY_t^\pi\|_{\mathcal{H}}^2\right] < \infty. \quad (7.3)$$

Applying [20, Lemma 1.2.3] with $F = X_t$ and $F_n = Y_t^\pi$ and taking into account (7.2) and (7.3), we conclude that $DX_t$ exists in $\mathbb{D}^{1,2}$. □

References

1. Hairer, M.: Ergodicity of stochastic differential equations driven by fractional Brownian motion. Ann. Probab. 33(2), 703–758 (2005)
2. Kloeden, P.E., Platen, E.: Numerical Solution of Stochastic Differential Equations, Applications of Mathematics (New York), vol. 23. Springer, Berlin (1992)
3. Milstein, G.N.: Numerical Integration of Stochastic Differential Equations, Mathematics and its Applications, vol. 313. Kluwer Academic, Dordrecht (1995)
4. Kloeden, P.E., Neuenkirch, A., Pavani, R.: Multilevel Monte Carlo for stochastic differential equations with additive fractional noise. Ann. Oper. Res. 189, 255–276 (2011)
5. Mishura, Y.S.: Stochastic Calculus for Fractional Brownian Motion and Related Processes. Lecture Notes in Mathematics, vol. 1929. Springer-Verlag, Berlin (2008)
6. Riedel, S., Wu, Y.: Semi-implicit Taylor schemes for stiff rough differential equations. arXiv:2006.13689 (2020)
7. Hu, Y., Liu, Y., Nualart, D.: Rate of convergence and asymptotic error distribution of Euler approximation schemes for fractional diffusions. Ann. Appl. Probab. 26(2), 1147–1207 (2016)
8. Hong, J., Huang, C., Kamrani, M., Wang, X.: Optimal strong convergence rate of a backward Euler type scheme for the Cox-Ingersoll-Ross model driven by fractional Brownian motion. Stoch. Process. Appl. 130(5), 2675–2692 (2020)
9. Boufoussi, B., Tudor, C.A.: Kramers-Smoluchowski approximation for stochastic evolution equations with FBM. Rev. Roumaine Math. Pures Appl. 50(2), 125–136 (2005)
10. Hu, Y., Nualart, D., Song, X.: A singular stochastic differential equation driven by fractional Brownian motion. Statist. Probab. Lett. 78(14), 2075–2085 (2008)
11. Neuenkirch, A.: Optimal pointwise approximation of stochastic differential equations driven by fractional Brownian motion. Stoch. Process. Appl. 118(12), 2294–2333 (2008)
12. Liu, Y., Tindel, S.: Discrete rough paths and limit theorems. Ann. Inst. Henri Poincaré Probab. Stat. 56(3), 1730–1774 (2020)
13. Biagini, F., Hu, Y., Øksendal, B., Zhang, T.: Stochastic Calculus for Fractional Brownian Motion and Applications, Probability and its Applications (New York). Springer, London Ltd, London (2008)
14. Hu, Y.: Integral transformations and anticipative calculus for fractional Brownian motions. Mem. Amer. Math. Soc. 175(825), viii+127 (2005)
15. Hu, Y.: Analysis on Gaussian Spaces. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ (2017)
16. Young, L.C.: An inequality of the Hölder type, connected with Stieltjes integration. Acta Math. 67(1), 251–282 (1936)
17. Hu, Y.: Semi-implicit Euler-Maruyama scheme for stiff stochastic equations. In: Koerezlioglu, H. (ed.) Stochastic Analysis and Related Topics V: The Silivri Workshop. Progr. Probab. 38, pp. 183–202. Birkhäuser, Boston (1996)

18. Abbaszadeh, M., Marquez, H.J.: Design of nonlinear state observers for one-sided Lipschitz systems. arXiv:1302.5867 (2013)

19. Hairer, E., Wanner, G.: Solving Ordinary Differential Equations. II, Springer Series in Computational Mathematics, vol. 14. Springer-Verlag, Berlin (1996)

20. Nualart, D.: The Malliavin Calculus and Related Topics, second ed., in: Probability and its Applications (New York). Springer-Verlag, Berlin (2006)

**Publisher’s Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.