IMA Journal of Numerical Analysis (2021) **41**, 1608–1638
doi:10.1093/imanum/draa019
Advance Access publication on 24 July 2020

Optimal rate of convergence for two classes of schemes to stochastic differential equations driven by fractional Brownian motions

Jialin Hong, Chuying Huang* and Xu Wang
Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China;
School of Mathematical Sciences, University of Chinese Academy of Sciences, Beijing 100049, China
*Corresponding author: huangchuying@lsec.cc.ac.cn

[Received on 3 November 2019; revised on 5 February 2020]

This paper investigates numerical schemes for stochastic differential equations driven by multi-dimensional fractional Brownian motions (fBms) with Hurst parameter $H \in (\frac{1}{2}, 1)$. Based on the continuous dependence of numerical solutions on the driving noises, we propose the order conditions of Runge–Kutta methods for the strong convergence rate $2H - \frac{1}{2}$, which is the optimal strong convergence rate for approximating the Lévy area of fBms. We provide an alternative way to analyse the convergence rate of explicit schemes by adding 'stage values' such that the schemes are interpreted as Runge–Kutta methods. Taking advantage of this technique the strong convergence rate of simplified step-N Euler schemes is obtained, which gives an answer to a conjecture in Deya *et al.* (2012) when $H \in (\frac{1}{2}, 1)$. Numerical experiments verify the theoretical convergence rate.

Keywords: fractional Brownian motion; strong convergence rate; Runge–Kutta method; simplified step-N Euler scheme.

1. Introduction

In this paper we consider the strong convergence rate of numerical schemes for the following stochastic differential equation (SDE):

$$
\begin{aligned}
dY_t &= V(Y_t)\mathrm{d}X_t = \sum_{l=1}^{d} V_l(Y_t)\mathrm{d}X_t^l, \quad t \in (0, T], \\
Y_0 &= y \in \mathbb{R}^m,
\end{aligned}
\tag{1.1}
$$

where $X_t = (X_t^1, \cdots, X_t^d)^\top \in \mathbb{R}^d$ with $X_t^1 = t$ and $X_t^2, \cdots, X_t^d$ being independent fractional Brownian motions (fBms) with Hurst parameter $H \in (\frac{1}{2}, 1)$. The well posedness is interpreted through Young's integral or fractional calculus pathwisely; see Zähle (1998), Lyons *et al.* (2007), Friz & Victoir (2010) and references therein.

The fBm X^l on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is a centred Gaussian process with continuous sample paths. Its covariance satisfies

$$
\mathbb{E}\Big[X_s^l X_t^l\Big] = \frac{1}{2}\Big(t^{2H} + s^{2H} - |t-s|^{2H}\Big), \quad \forall\, s, t \in [0, T], \tag{1.2}
$$

where $H \in (0,1)$ is called the Hurst parameter. The fBm is a semi-martingale and Markovian process only if $H = \frac{1}{2}$, that is, the standard Brownian motion. Otherwise, the process exhibits long-range or short-range dependence when $H \in (\frac{1}{2}, 1)$ or $H \in (0, \frac{1}{2})$, respectively. This brings in wide applications of SDEs driven by fBms, such as the flows in porous media (Cao *et al.*, 2017, 2018), the stochastic volatility model (Chronopoulou & Viens, 2012; Garnier & Sølna, 2017), the simulation of transient noise in circuit (Denk & Winkler, 2007), the rough Hamiltonian systems (Hong *et al.*, 2018), the stochastic modelling in nanoscale biophysics (Kou, 2008) and the range of cumulated water flows in hydrology (Mandelbrot & Van Ness, 1968). However, there are many obstacles in both the simulation for noises and analysis of strong convergence rate for numerical schemes.

On one hand the covariance of fBms causes difficulties in simulating Nth-level ($N \geqslant 2$) iterated integrals in multi-dimensional case. For the standard Brownian case $H = \frac{1}{2}$ Milstein & Tretyakov (2004) give an approximation for iterated integrals, which can be simulated by specific independent and identically distributed Gaussian random variables. For the case $H \neq \frac{1}{2}$ other techniques need to be explored. An implementable choice is to substitute the Nth-level iterated integral of X by $\frac{1}{N!}(\Delta X_k)^N$ directly. The corresponding numerical schemes are called simplified step-N Euler schemes; see Gradinaru & Nourdin (2009), Deya *et al.* (2012), Friz & Riedel (2014) and Bayer *et al.* (2016). Another way is taking advantage of stages values to design Runge–Kutta methods, which are derivative free and can be particularly chosen as structure-preserving methods; see Milstein & Tretyakov (2004), Hong *et al.* (2018) and references therein.

On the other hand, since fBms with $H \neq \frac{1}{2}$ have no independent increments or martingale property, approaches for the analysis of the convergence rate for schemes in the fractional setting are different from those via the fundamental convergence theorem, which reveals the relationship between the local error and the global error in Milstein & Tretyakov (2004) for the standard Brownian case. In Deya *et al.* (2012) authors analyse the simplified step-2 Euler scheme by the Wong–Zakai approximations and obtain the pathwise convergence rate $(H - \frac{1}{3})^-$ in Hölder norm for any $H \in (\frac{1}{3}, 1)$. They conjecture that the optimal convergence rate in supremum norm is $2H - \frac{1}{2}$, based on the best probable rate for implementable approximations of the Lévy area of fBms proposed in Neuenkirch *et al.* (2010) and Neuenkirch & Shalaiko (2016). In Hu *et al.* (2016a, 2017) authors prove the optimal strong convergence rate of the Crank–Nicolson scheme and the modified Euler scheme for $H \in (\frac{1}{2}, 1)$, respectively, by combining the techniques of Malliavin calculus and fractional calculus.

The goal of this paper is to demonstrate that both simplified step-N Euler schemes and Runge–Kutta methods satisfying order condition (3.10) achieve the same convergence rate as the best probable rate $2H - \frac{1}{2}$ for approximations of the Lévy area of fBms. Namely, denote by Y^n the numerical solution under study with time step size $h = \frac{T}{n}$, then

$$\left\| \sup_{t\in[0,T]} |Y_t - Y_t^n| \right\|_{L^p(\Omega)} \leqslant Ch^{2H-\frac{1}{2}}, \quad \forall\, p \geqslant 1. \tag{1.3}$$

For general Runge–Kutta methods we construct continuous versions of numerical solutions and stage values, which are continuously dependent on the driving noises. This robustness coincides with the property of the exact solution. Combining estimates for iterated integrals of fBms in Hu *et al.* (2016b, 2017) we utilize the property that numerical solutions are determined implicitly through stage values to derive the order conditions for Runge–Kutta methods. Moreover, since the martingale inequality is not valid for fBms, the Besov–Hölder embedding is used to prove the $L^p(\Omega)$-estimate for the supremum norm of the error.

Further, to obtain the strong convergence rate for the simplified step-N Euler schemes, we introduce 'stage values' to interpret these explicit schemes as Runge–Kutta methods satisfying order condition (3.10) with negligible high-order terms. This approach leads us to the same strong convergence rate $2H-\frac{1}{2}$ and avoids the estimation of the Wong–Zakai approximations. Our result gives an answer to the conjecture in Deya *et al.* (2012) for $H \in (\frac{1}{2}, 1)$. Numerical experiments are performed to verify the optimality.

For $H \in (\frac{1}{4}, \frac{1}{2})$ the optimal strong convergence rate for the simplified step-2 Euler scheme is still an open problem, in which case equation (1.1) is understood via the rough path framework in Davie (2007), Lyons *et al.* (2007) and Friz & Victoir (2010). We refer to Bayer *et al.* (2016), Hong *et al.* (2018) and Liu & Tindel (2019) for related results.

The paper is organized as follows. In Section 2 we recall some definitions and results about fractional calculus and fBms. In Section 3 we prove the solvability of implicit Runge–Kutta methods, derive the order conditions of the strong convergence rate $2H-\frac{1}{2}$ for Runge–Kutta methods and illustrate the approach to view the simplified step-N Euler schemes as Runge–Kutta methods. In Section 4 we show the continuous dependence of numerical solutions under study on the driving noises and then prove the main theorems, Theorems 3.2 and 3.3. Numerical experiments are presented in Section 5.

2. Preliminaries

In this section we introduce some notations, definitions and results about fractional calculus and fBms. We use C as a generic constant that could be different from line to line.

2.1 *Fractional calculus*

Denote by $\mathscr{C}([0,T];\mathbb{R}^d)$ the space of continuous functions from $[0,T]$ to $\mathbb{R}^d$. For any $f \in \mathscr{C}([0,T];\mathbb{R}^d)$, $0 \leqslant s < t \leqslant T$ and $0 < \beta \leqslant 1$ the β-Hölder semi-norm of f on $[s,t]$ is defined by

$$\|f\|_{s,t,\beta} := \sup\left\{\frac{|f_v - f_u|}{(v-u)^\beta},\ s \leqslant u < v \leqslant t\right\},$$

where $|\cdot|$ is the Euclidean norm on $\mathbb{R}^d$. Especially, we use $\|f\|_\beta := \|f\|_{0,T,\beta}$ for short. The Hölder semi-norm is estimated by the Besov–Hölder embedding, which is a corollary from Garsia–Rodemich–Rumsey inequality.

LEMMA 2.1 (see Friz & Victoir, 2010, Corollary A.2) Let $q > 1$, $\alpha \in (\frac{1}{q}, 1)$ and $f \in \mathscr{C}([0,T];\mathbb{R}^d)$. Then there exists a constant $C = C(\alpha, q)$ such that for all $0 \leqslant s < t \leqslant T$,

$$\|f\|^q_{s,t,\alpha-\frac{1}{q}} \leqslant C \int_s^t \int_s^t \frac{|f_u - f_v|^q}{|u-v|^{1+q\alpha}}\,\mathrm{d}u\,\mathrm{d}v.$$

Let $f \in \mathscr{C}([s,t];\mathbb{R})$ be β-Hölder continuous on $[s,t] \subseteq [0,T]$ with $1/2 < \beta < 1$ and $g : [s,t] \to \mathbb{R}$ be a step function defined by $g_t = g_s \mathbf{1}_{\{s\}} + \sum_{k=0}^{n-1} g^k \mathbf{1}_{(t_k,t_{k+1}]}$ with $s = t_0 < t_1 < \cdots < t_n = t$. The integral of g with respect to f can be defined as the Riemann sum

$$\int_s^t g_r\,\mathrm{d}f_r := \sum_{k=0}^{n-1} g^k \left(f_{t_{k+1}} - f_{t_k}\right).$$

For $1-\beta<\alpha<1$, according to fractional calculus (see e.g., Zähle, 1998, Section 2), the integral has the characterization

$$\int_s^t g_r\,\mathrm{d}f_r = (-1)^\alpha \int_s^t D^\alpha_{s+}g_r D^{1-\alpha}_{t-}F_r\,\mathrm{d}r. \tag{2.1}$$

Here $(-1)^\alpha = e^{-i\pi\alpha}$, $F_r := f_r - f_t$, $D^\alpha_{s+}g_r$ and $D^{1-\alpha}_{t-}F_r$ are fractional Weyl derivatives of the order α and $1-\alpha$, respectively:

$$(D^\alpha_{s+}g)_r := \frac{1}{\Gamma(1-\alpha)}\left(\frac{g_r}{(r-s)^\alpha} + \alpha\int_s^r \frac{g_r-g_u}{(r-u)^{\alpha+1}}\,\mathrm{d}u\right),$$

$$(D^{1-\alpha}_{t-}F)_r := \frac{(-1)^{1-\alpha}}{\Gamma(\alpha)}\left(\frac{F_r}{(t-r)^{1-\alpha}} + (1-\alpha)\int_r^t \frac{F_r-F_u}{(u-r)^{2-\alpha}}\,\mathrm{d}u\right).$$

2.2 A priori *estimate for the solution and iterated integrals*

In the sequel, we denote by $\mathscr{C}^N_b(\mathbb{R}^m;\mathbb{R}^M)$ the space of bounded and N-times continuously differentiable functions $V:\mathbb{R}^m\to\mathbb{R}^M$ with bounded derivatives.

The following lemma shows the well posedness of (1.1). Since almost all sample paths of X are β-Hölder continuous for any $\beta\in(0,H)$, it implies that the solution $Y_\cdot$ is continuously dependent on the driving noises in Hölder semi-norm. The numerical solutions proposed in this paper will be shown to possess a similar property in Section 4.

Lemma 2.2 (see e.g., Friz & Victoir, 2010, Theorem 10.14) Let $1/2<\beta<H$. If $V\in\mathscr{C}^1_b(\mathbb{R}^m;\mathbb{R}^{m\times d})$ then there exists a unique solution of (1.1) satisfying almost surely that

$$\|Y\|_\beta \leqslant C(V,\beta,T)\max\left\{\|X\|_\beta, \|X\|_\beta^{1/\beta}\right\},$$

$$\|Y\|_\infty \leqslant |y| + C(V,\beta,T)\max\left\{\|X\|_\beta, \|X\|_\beta^{1/\beta}\right\},$$

where $\|Y\|_\infty := \sup\left\{|Y_u|,\ 0\leqslant u\leqslant T\right\}$. Moreover, for some $C_0>0$ and $0\leqslant s<t\leqslant T$ such that $\|X\|_\beta|t-s|^\beta\leqslant C_0$, the estimate can be improved to

$$\|Y\|_{s,t,\beta} \leqslant C(V,\beta,T,C_0)\|X\|_\beta.$$

Several estimates for the Lévy area type processes and iterated integrals of X are introduced in the following lemma, which plays an important role in the error analysis for numerical schemes.

LEMMA 2.3 (see Hu *et al.*, 2016b, 2017) Denote $t_k = kh$, $k = 0, \cdots, n$ with $h = \frac{T}{n}$ and $n \in \mathbb{N}_+$. Let $X_t^1 = t$ and $X_t^2, \cdots, X_t^d$ be independent fBms with $H > \frac{1}{2}$. Then it holds that for $0 \leqslant t_i < t_j \leqslant T$ and $p \geqslant 1$,

$$\left\| \sum_{k=i}^{j-1} \left[\int_{t_k}^{t_{k+1}} \int_{t_k}^{s} \mathrm{d}X_u^3 \, \mathrm{d}X_s^2 - \int_{t_k}^{t_{k+1}} \int_{s}^{t_{k+1}} \mathrm{d}X_u^3 \, \mathrm{d}X_s^2 \right] \right\|_{L^p(\Omega)} \leqslant C|t_j - t_i|^{\frac{1}{2}} h^{2H-\frac{1}{2}}, \tag{2.2}$$

$$\left\| \sum_{k=i}^{j-1} \left[\int_{t_k}^{t_{k+1}} \int_{t_k}^{s} \mathrm{d}X_u^1 \, \mathrm{d}X_s^2 - \int_{t_k}^{t_{k+1}} \int_{s}^{t_{k+1}} \mathrm{d}X_u^1 \, \mathrm{d}X_s^2 \right] \right\|_{L^p(\Omega)} \leqslant C|t_j - t_i|^{\frac{1}{2}} h^{H+\frac{1}{2}}, \tag{2.3}$$

where $C = C(p)$ above is independent of n. Moreover, for any $l_1, \cdots, l_{N'} \in \{1, \cdots, d\}$, it holds that

$$\left\| \sum_{k=i}^{j-1} \int_{t_k}^{t_{k+1}} \int_{t_k}^{u_1} \cdots \int_{t_k}^{u_{N'-1}} \mathrm{d}X_{u_{N'}}^{l_{N'}} \cdots \mathrm{d}X_{u_2}^{l_2} \, \mathrm{d}X_{u_1}^{l_1} \right\|_{L^p(\Omega)} \leqslant C|t_j - t_i|^{\frac{1}{2}} h^r, \tag{2.4}$$

where $r = N''H + N' - N'' - 1$ when $N'' = \sharp\{l_i : l_i \neq 1\}$ is even, $r = N''H + N' - N'' - H$ when N'' is odd and $C = C(p)$ is independent of n.

In particular, if $N' = N'' = 2$ then $r = 2H - 1 < 2H - \frac{1}{2}$. This implies that the convergence rate of the first-level iterated integrals of X in the form of (2.2) is higher than that in the form of (2.4). We will use the next lemma to connect the convergence rates above for iterated integrals of X with the numerical errors.

LEMMA 2.4 (Hu *et al.*, 2016b, Proposition 8) Let f be a β-Hölder continuous stochastic process in $L^{2p}(\Omega)$ with $\frac{1}{2} < \beta < H$ and $p \geqslant 1$, i.e.,

$$\sup \left\{ \frac{\|f_v - f_u\|_{L^{2p}(\Omega)}}{(v-u)^\beta}, \; 0 \leqslant u < v \leqslant T \right\} < \infty.$$

If a sequence of stochastic processes $\{g_n\}_{n \in \mathbb{N}_+}$ satisfies $g_n(t_i) = \sum_{k=0}^{i-1} \xi_{n,k}$ and

$$\|g_n(t_j) - g_n(t_i)\|_{L^{2p}(\Omega)} \leqslant C|t_j - t_i|^{\frac{1}{2}}, \quad \forall\, 0 \leqslant t_i < t_j \leqslant T,$$

then

$$\left\| \sum_{k=i}^{j-1} f_{t_k} \xi_{n,k} \right\|_{L^p(\Omega)} \leqslant C|t_j - t_i|^{\frac{1}{2}}, \quad \forall\, 0 \leqslant t_i < t_j \leqslant T.$$

All constants $C = C(p,f)$ above are independent of n.

3. Strong convergence rate

For a numerical scheme we apply the uniform partition of the interval $[0, T]$ with step size $h = \frac{T}{n}$, $n \in \mathbb{N}_+$ and denote $t_k = kh$, $k = 0, \cdots, n$.

3.1 *Runge–Kutta methods*

Consider an $\mathbf{s}$-stage Runge–Kutta method applied to (1.1):

$$Y^n_{t_{k+1},i} = Y^n_{t_k} + \sum_{j=1}^{\mathbf{s}} a_{ij} V(Y^n_{t_{k+1},j}) \Delta X_k, \tag{3.1}$$

$$Y^n_{t_{k+1}} = Y^n_{t_k} + \sum_{i=1}^{\mathbf{s}} b_i V(Y^n_{t_{k+1},i}) \Delta X_k \tag{3.2}$$

with $i,j = 1,\cdots,\mathbf{s}$, $k = 0,\cdots,n-1$, $\Delta X_k = X_{t_{k+1}} - X_{t_k} \in \mathbb{R}^d$ and $Y^n_{t_0} = y \in \mathbb{R}^m$. Here $Y^n_{t_{k+1},i}$, $i = 1,\cdots,\mathbf{s}$, are called stage values.

If the method is an implicit one, such as the midpoint scheme, the solvability of (3.1) and (3.2) should be taken into consideration. For the classical case with $H = 1/2$ the increment $X_{t_{k+1}} - X_{t_k}$ is usually replaced by $\zeta\sqrt{h}$, with ζ being a bounded truncation of an $\mathcal{N}(0,1)$-distributed random variable. However, for the case $H \in (1/2,1)$ considered in this paper, the covariance function turns out to be more complicated, which makes it more difficult to get the error of the truncation. To avoid this we show that the solvability of implicit Runge–Kutta methods can be obtained without using the truncation technique if the coefficients are assumed to be bounded.

PROPOSITION 3.1 If $V \in \mathscr{C}^0_b(\mathbb{R}^m;\mathbb{R}^{m\times d})$, then for arbitrary time step size $h > 0$, initial value y and coefficients $\{a_{ij}, b_i : i,j = 1,\cdots,\mathbf{s}\}$, the $\mathbf{s}$-stage Runge–Kutta methods (3.1) and (3.2) has at least one solution for every ω.

Proof. Fix $\omega \in \Omega$, $h > 0$ and $Y^n_{t_k} \in \mathbb{R}^m$.

Let $Z_1,\cdots,Z_{\mathbf{s}} \in \mathbb{R}^m$ and $Z = (Z_1^\top,\cdots,Z_{\mathbf{s}}^\top)^\top \in \mathbb{R}^{m\mathbf{s}}$. We define a map $\phi : \mathbb{R}^{m\mathbf{s}} \to \mathbb{R}^{m\mathbf{s}}$ with

$$\phi(Z) := (\phi(Z)_1^\top,\cdots,\phi(Z)_{\mathbf{s}}^\top)^\top,$$

$$\phi(Z)_i := Z_i - Y^n_{t_k} - \sum_{j=1}^{s} a_{ij} V(Z_j) \Delta X_k(\omega),\ i = 1,\cdots,\mathbf{s}.$$

It suffices to prove that $\phi(Z) = 0$ has at least one solution, which implies the solvability of (3.1) and thus the solvability of the Runge–Kutta method. Let $c := \max\{|a_{ij}| : i,j = 1,\cdots,\mathbf{s}\}$, $\nu := \sup_{y\in\mathbb{R}^m} |V(y)|$ and

$$R = \sqrt{\mathbf{s}}|Y^n_{t_k}| + \mathbf{s}\sqrt{\mathbf{s}}c\nu|\Delta X_k(\omega)| + 1.$$

If $|Z| = R$ then

$$\begin{aligned} Z^\top\phi(Z) &= \sum_{i=1}^{\mathbf{s}} Z_i^\top \left(Z_i - Y^n_{t_k} - \sum_{j=1}^{\mathbf{s}} a_{ij} V(Z_j)\Delta X_k(\omega) \right) \\ &\geqslant |Z| \left(|Z| - \sqrt{\mathbf{s}}|Y^n_{t_k}| - \mathbf{s}\sqrt{\mathbf{s}}c\nu|\Delta X_k(\omega)| \right) > 0. \end{aligned}$$

We aim to show that $\phi(Z) = 0$ has a solution in the ball $B_R := \{Z : |Z| \leqslant R\}$. Assume by contradiction that $\phi(Z) \neq 0$ if $|Z| \leqslant R$. We define a continuous map ψ by $\psi(Z) := -\frac{R\phi(Z)}{|\phi(Z)|}$. Since $\psi : B_R \to B_R$, ψ has at least one fixed point Z^* such that $Z^* = \psi(Z^*)$ and $|Z^*| = R$. This leads to a contradiction since $|Z^*|^2 = \psi(Z^*)^\top Z^* = -\frac{R\phi(Z^*)^\top Z^*}{|\phi(Z^*)|} < 0$. Therefore, ϕ has at least one solution. □

To derive order conditions on coefficients of Runge–Kutta methods with the strong convergence rate $2H - \frac{1}{2}$, we first construct the continuous versions (3.3) and (3.4) for the Runge–Kutta methods (3.1) and (3.2), taking advantages of the stage values $Y^n_{t_k,i}$. Denote $\lceil t \rceil^n := t_{k+1}$ for $t \in (t_k, t_{k+1}]$. In particular, $t = t_k$ if and only if $t = \lceil t \rceil^n$ for some $k = 0, \cdots, n$. The continuous version reads

$$Y^n_{t,i} := Y^n_{(t-h)\vee 0} + \sum_{j=1}^{\mathbf{s}} \int_{(t-h)\vee 0}^{t} a_{ij} V(Y^n_{\lceil s \rceil^n, j})\, \mathrm{d}X_s, \quad i = 1, \cdots, \mathbf{s}, \tag{3.3}$$

$$Y^n_t := y + \sum_{i=1}^{\mathbf{s}} \int_0^t b_i V(Y^n_{\lceil s \rceil^n, i})\, \mathrm{d}X_s, \tag{3.4}$$

where $s \vee t := \max\{s, t\}$. Then the error is decomposed into

$$\begin{aligned} Y_t - Y^n_t =& \left[\int_0^t V(Y_s)\, \mathrm{d}X_s - \int_0^t V(Y^n_s)\, \mathrm{d}X_s \right] \\ &+ \left[\int_0^t V(Y^n_s)\, \mathrm{d}X_s - \int_0^t \sum_{i=1}^{\mathbf{s}} b_i V(Y^n_{\lceil s \rceil^n, i})\, \mathrm{d}X_s \right] \\ =&: L_t + R_t. \end{aligned}$$

For the first term the Taylor expansion yields

$$\begin{aligned} L_t &= \int_0^t V(Y_s)\, \mathrm{d}X_s - \int_0^t V(Y^n_s)\, \mathrm{d}X_s \\ &= \sum_{l=1}^{d} \int_0^t \int_0^1 \nabla V_l(\theta Y_s + (1-\theta) Y^n_s)(Y_s - Y^n_s)\, \mathrm{d}\theta\, \mathrm{d}X^l_s. \end{aligned}$$

For the second term, fix any $t = \lceil t \rceil^n$, we have

$$\begin{aligned} R_t &= \int_0^t V(Y^n_s)\, \mathrm{d}X_s - \int_0^t \sum_{i=1}^{\mathbf{s}} b_i V\left(Y^n_{\lceil s \rceil^n, i}\right)\, \mathrm{d}X_s \\ &= \sum_{k=0}^{nt/T-1} \int_{t_k}^{t_{k+1}} \left[V(Y^n_s) - \sum_{i=1}^{\mathbf{s}} b_i V\left(Y^n_{t_{k+1}, i}\right) \right] \mathrm{d}X_s. \end{aligned}$$

For $i = 1, \cdots, \mathbf{s}$, $q = 1, \cdots, m$, denote by $Y_t^{n,q}$ and $Y_{t,i}^{n,q}$ the qth component of Y_t^n and $Y_{t,i}^n$. We apply the Taylor expansion to $V(Y_{t_{k+1},i}^n)$ at $Y_{t_k}^n$ and at $Y_{t_{k+1}}^n$, respectively, then

$$V\left(Y_{t_{k+1},i}^n\right) = V\left(Y_{t_k}^n\right) + \int_0^1 \sum_{q=1}^m \partial_q V\left(\theta Y_{t_{k+1},i}^n + (1-\theta) Y_{t_k}^n\right)\left(Y_{t_{k+1},i}^{n,q} - Y_{t_k}^{n,q}\right) \, \mathrm{d}\theta$$

and

$$V\left(Y_{t_{k+1},i}^n\right) = V\left(Y_{t_{k+1}}^n\right) + \int_0^1 \sum_{q=1}^m \partial_q V\left(\theta Y_{t_{k+1},i}^n + (1-\theta) Y_{t_{k+1}}^n\right)\left(Y_{t_{k+1},i}^{n,q} - Y_{t_{k+1}}^{n,q}\right) \, \mathrm{d}\theta,$$

where ∂_q denotes the partial differential operator with respect to the qth variable. For simplicity we omit the range of indices in summations in the sequel. Let $\eta \in [0,1]$, then for any $s \in (t_k, t_{k+1}]$,

$$\begin{aligned}
V(Y_s^n) - \sum_i b_i V\left(Y_{t_{k+1},i}^n\right)
&= \eta\left[V(Y_s^n) - \sum_i b_i V\left(Y_{t_k}^n\right)\right] \\
&\quad - \eta \sum_{i,q} b_i \int_0^1 \partial_q V\left(\theta Y_{t_{k+1},i}^n + (1-\theta) Y_{t_k}^n\right)\left(Y_{t_{k+1},i}^{n,q} - Y_{t_k}^{n,q}\right) \, \mathrm{d}\theta \\
&\quad + (1-\eta)\left[V\left(Y_s^n\right) - \sum_i b_i V\left(Y_{t_{k+1}}^n\right)\right] \\
&\quad - (1-\eta) \sum_{i,q} b_i \int_0^1 \partial_q V\left(\theta Y_{t_{k+1},i}^n + (1-\theta) Y_{t_{k+1}}^n\right)\left(Y_{t_{k+1},i}^{n,q} - Y_{t_{k+1}}^{n,q}\right) \, \mathrm{d}\theta \\
&=: \eta R_s^1 + \eta R_s^2 + (1-\eta) R_s^3 + (1-\eta) R_s^4.
\end{aligned}$$

Applying the chain rule and (3.4) to $V(Y_s^n)$, we get

$$\begin{aligned}
R_s^1 &= V(Y_s^n) - \sum_i b_i V(Y_{t_k}^n) \\
&= \left[V(Y_{t_k}^n) + \sum_{q,i} \int_{t_k}^s \partial_q V\left(Y_u^n\right) b_i V^q\left(Y_{t_{k+1},i}^n\right) \, \mathrm{d}X_u\right] - \sum_i b_i V(Y_{t_k}^n) \\
&= \left[V(Y_{t_k}^n) + \sum_{q,i} \partial_q V\left(Y_{t_k}^n\right) \int_{t_k}^s b_i V^q\left(Y_{t_{k+1},i}^n\right) \, \mathrm{d}X_u + E_s^1\right] - \sum_i b_i V\left(Y_{t_k}^n\right),
\end{aligned}$$

where $E_s^1 := \sum_{q,q',i,i'} \int_{t_k}^{s} \left[\int_{t_k}^{u} \partial_{q'}\partial_q V(Y_{u'}^n) b_{i'} V^{q'}(Y_{t_{k+1},i'}^n)\,\mathrm{d}X_{u'}\right] b_i V^q(Y_{t_{k+1},i}^n)\,\mathrm{d}X_u$ represents the remainder term of R_s^1. Therefore, we propose the first condition

$$\sum_{i=1}^{\mathbf{s}} b_i = 1$$

to obtain

$$\int_{t_k}^{t_{k+1}} R_s^1\,\mathrm{d}X_s = \int_{t_k}^{t_{k+1}} \sum_{q,i} \partial_q V\left(Y_{t_k}^n\right) \int_{t_k}^{s} b_i V^q\left(Y_{t_{k+1},i}^n\right)\mathrm{d}X_u\,\mathrm{d}X_s + \int_{t_k}^{t_{k+1}} E_s^1\,\mathrm{d}X_s. \tag{3.5}$$

Similarly,

$$\int_{t_k}^{t_{k+1}} R_s^3\,\mathrm{d}X_s = -\int_{t_k}^{t_{k+1}} \sum_{q,i} \partial_q V\left(Y_{t_{k+1}}^n\right) \int_{s}^{t_{k+1}} b_i V^q\left(Y_{t_{k+1},i}^n\right)\mathrm{d}X_u\,\mathrm{d}X_s + \int_{t_k}^{t_{k+1}} E_s^3\,\mathrm{d}X_s \tag{3.6}$$

with $E_s^3 := \sum_{q,q',i,i'} \int_{s}^{t_{k+1}} \left[\int_{u}^{t_{k+1}} \partial_{q'}\partial_q V(Y_{u'}^n) b_{i'} V^{q'}(Y_{t_{k+1},i'}^n)\,\mathrm{d}X_{u'}\right] b_i V^q(Y_{t_{k+1},i}^n)\,\mathrm{d}X_u$. For R_s^2 and R_s^4 using the definitions from (3.1) and (3.2) that

$$\begin{aligned} Y_{t_{k+1},i}^n - Y_{t_k}^n &= \sum_j a_{ij} V\left(Y_{t_{k+1},j}^n\right) \Delta X_k, \\ Y_{t_{k+1},i}^n - Y_{t_{k+1}}^n &= -\sum_j b_j V\left(Y_{t_{k+1},j}^n\right) \Delta X_k + \sum_j a_{ij} V\left(Y_{t_{k+1},j}^n\right) \Delta X_k, \end{aligned}$$

we have

$$\int_{t_k}^{t_{k+1}} R_s^2\,\mathrm{d}X_s = -\int_{t_k}^{t_{k+1}} \sum_{i,q,j} b_i \partial_q V\left(Y_{t_{k+1},i}^n\right)\left[\int_{t_k}^{t_{k+1}} a_{ij} V^q\left(Y_{t_{k+1},j}^n\right)\mathrm{d}X_u\right]\mathrm{d}X_s + \int_{t_k}^{t_{k+1}} E_s^2\,\mathrm{d}X_s, \tag{3.7}$$

$$\int_{t_k}^{t_{k+1}} R_s^4\,\mathrm{d}X_s = -\int_{t_k}^{t_{k+1}} \sum_{i,q,j} b_i \partial_q V\left(Y_{t_{k+1},i}^n\right)\left[\int_{t_k}^{t_{k+1}} (a_{ij} - b_j) V^q\left(Y_{t_{k+1},j}^n\right)\mathrm{d}X_u\right]\mathrm{d}X_s + \int_{t_k}^{t_{k+1}} E_s^4\,\mathrm{d}X_s. \tag{3.8}$$

The remainder terms E^2_s and E^4_s are defined by

$$E^2_s = \sum_{i,q,q',j} b_i \int_0^1 \left[\int_0^1 \partial_{q'}\partial_q V\left((\theta'\theta + (1-\theta'))Y^n_{t_{k+1},i} + \theta'(1-\theta)Y^n_{t_k}\right)(\theta-1)\left(Y^n_{t_{k+1},i} - Y^n_{t_k}\right) \mathrm{d}\theta'\right] \left[\int_{t_k}^{t_{k+1}} a_{ij} V^q\left(Y^n_{t_{k+1},j}\right) \mathrm{d}X_u\right] \mathrm{d}\theta,$$

$$E^4_s = \sum_{i,q,q',j} b_i \int_0^1 \left[\int_0^1 \partial_{q'}\partial_q V\left((\theta'\theta + (1-\theta'))Y^n_{t_{k+1},i} + \theta'(1-\theta)Y^n_{t_{k+1}}\right)(1-\theta)\left(Y^n_{t_{k+1},i} - Y^n_{t_{k+1}}\right) \mathrm{d}\theta'\right] \left[\int_{t_k}^{t_{k+1}} (a_{ij}-b_j) V^q\left(Y^n_{t_{k+1},j}\right) \mathrm{d}X_u\right] \mathrm{d}\theta.$$

Applying the Taylor expansion to both $V(Y^n_{t_{k+1},i})$ and $V(Y^n_{t_{k+1},j})$ at $Y^n_{t_k}$ in (3.7) and (3.8), and choosing $\eta = \frac{1}{2}$, we propose another condition

$$\sum_{i=1}^{\mathrm{s}} b_i \left(\sum_{j=1}^{\mathrm{s}} b_j - 2a_{ij}\right) = 0$$

such that

$$\sum_{q=1}^{m}\sum_{i=1}^{\mathrm{s}} b_i \sum_{j=1}^{\mathrm{s}} \int_{t_k}^{t_{k+1}} \partial_q V(Y^n_{t_k}) \int_{t_k}^{t_k} a_{ij} V^q(Y^n_{t_k})\,\mathrm{d}X_u\,\mathrm{d}X_s + \int_{t_k}^{t_{k+1}} \partial_q V(Y^n_{t_k}) \int_{t_k}^{t_{k+1}} (a_{ij}-b_j) V^q(Y^n_{t_k})\,\mathrm{d}X_u\,\mathrm{d}X_s = 0.$$

This implies that the second-level iterated integrals of X in (3.7) and (3.8) vanish. Therefore, the leading order term of R_t is shown to appear only in (3.5) and (3.6), which is in the forms of (2.2) and (2.3). More precisely, recalling $\eta = \frac{1}{2}$ and applying the Tayor expansion again to both $V(Y^n_{t_k})$ and $V(Y^n_{t_{k+1}})$ at $Y^n_{t_{k+1},i}$ in (3.5)–(3.6), we derive that the leading order term of R_t is

$$R^{lead}_t = \frac{1}{2}\sum_{k=0}^{nt/T-1}\sum_{q=1}^{m}\sum_{i=1}^{\mathrm{s}} b_i \Bigg[\int_{t_k}^{t_{k+1}}\int_{t_k}^{s} \partial_q V\left(Y^n_{t_{k+1},i}\right) V^q\left(Y^n_{t_{k+1},i}\right) \mathrm{d}X_u\,\mathrm{d}X_s - \int_{t_k}^{t_{k+1}}\int_{s}^{t_{k+1}} \partial_q V\left(Y^n_{t_{k+1},i}\right) V^q\left(Y^n_{t_{k+1},i}\right) \mathrm{d}X_u\,\mathrm{d}X_s\Bigg]. \tag{3.9}$$

Note that the remainder $R_t - R^{lead}_t$ contains the third-level iterated integrals of X in each interval $(t_k, t_{k+1}]$ in the form of (2.4).

Combining the two conditions above we obtain the strong convergence rate for the Runge–Kutta methods, whose proof is postponed to the next section.

THEOREM 3.2 Suppose $V \in \mathscr{C}_b^3(\mathbb{R}^m;\mathbb{R}^{m\times d})$ and $H > 1/2$. Denote $c_i = \sum_{j=1}^{\mathbf{s}} a_{ij}$. If it holds that

$$\sum_{i=1}^{\mathbf{s}} b_i = 1 \quad \text{and} \quad \sum_{i=1}^{\mathbf{s}} b_i c_i = 1/2 \tag{3.10}$$

then the strong convergence order of Runge–Kutta method for (1.1) is $2H - \frac{1}{2}$. More precisely, there exists a constant C independent of n such that

$$\left\| \sup_{t\in[0,T]} |Y_t - Y_t^n| \right\|_{L^p(\Omega)} \leqslant Ch^{2H-\frac{1}{2}}, \quad p \geqslant 1, \tag{3.11}$$

where $h = \frac{T}{n}$ and Y_t^n is defined by (3.4).

REMARK 3.1 Given $X_{t_k}, k = 1, \cdots, n$ as the information of fBms the best approximation for the solution Y_t is the conditional expectation $\mathbb{E}\left[Y_t \,|\, X_{t_1}, X_{t_2}, \cdots, X_{t_n}\right]$. Neuenkirch & Shalaiko (2016) derive both the upper and lower bounds of

$$n^{2H-\frac{1}{2}} \left\| Y_T - \mathbb{E}\left[Y_T \,|\, X_{t_1}, X_{t_2}, \cdots, X_{t_n}\right] \right\|_{L^2(\Omega)}$$

for a two-dimensional linear equation, whose solution is the Lévy area of fBms. This implies that the optimal strong convergence rate of implementable schemes for equations driven by multi-dimensional fBms is $2H - \frac{1}{2}$ in general.

REMARK 3.2 If the diffusion term satisfies the following commutative condition

$$\sum_q \partial_q V_l V_{l'}^q = \sum_q \partial_q V_{l'} V_l^q, \quad 1 < l \leqslant l' \leqslant d,$$

then Fubini's theorem shows

$$R_t^{lead} = \frac{1}{2} \sum_{k=0}^{nt/T-1} \sum_{i,q} \sum_{l\neq 1} b_i \left[\partial_q V_l \left(Y_{t_{k+1},i}^n\right) V_1^q \left(Y_{t_{k+1},i}^n\right) - \partial_q V_1 \left(Y_{t_{k+1},i}^n\right) V_l^q \left(Y_{t_{k+1},i}^n\right) \right]$$
$$\left[\int_{t_k}^{t_{k+1}} \int_{t_k}^{s} \mathrm{d}u \, \mathrm{d}X_s^l - \int_{t_k}^{t_{k+1}} \int_{s}^{t_{k+1}} \mathrm{d}u \, \mathrm{d}X_s^l \right].$$

As a result the strong convergence rate in (3.11) is $H + \frac{1}{2}$ according to (2.3). This indicates that a large class of numerical schemes can achieve the best probable rate of convergence $H + \frac{1}{2}$ for the scalar noise case proved in Neuenkirch (2008).

REMARK 3.3 If the drift and diffusion terms satisfy the commutative condition

$$\sum_q \partial_q V_l V_{l'}^q = \sum_q \partial_q V_{l'} V_l^q, \quad 1 \leqslant l \leqslant l' \leqslant d,$$

then $R_t^{lead} = 0$ and the strong convergence rate is $2H$ based on (2.4). The order is optimal and reasonable in the view that the order condition above coincides with the second-order condition for Runge–Kutta methods in the deterministic case by taking $H = 1$ formally for $X_t = t$.

REMARK 3.4 As H goes to $\frac{1}{2}$ the convergence rate in (3.11) tends to $\frac{1}{2}$ in general cases and to 1 in commutative cases, respectively. This is consistent with classical results for Stratonovich SDEs driven by standard Brownian motions.

REMARK 3.5 If the Hurst parameter of X^l is H_l, $l = 1, \ldots, d$, satisfying $H_1 \geqslant \cdots \geqslant H_d > \frac{1}{2}$, a similar procedure as in Lemma 2.3 and Theorem 3.2 leads to the strong convergence rate $H_{d-1} + H_d - \frac{1}{2}$.

3.2 *Simplified step-N Euler schemes*

Fix an integer $N \geqslant 2$. The simplified step-N Euler scheme applied to (1.1) is

$$Y_{t_{k+1}}^n = Y_{t_k}^n + \sum_{w=1}^{N} \sum_{l_w,\cdots,l_1=1}^{d} \mathscr{V}_{l_w} \cdots \mathscr{V}_{l_1} I(Y_{t_k}^n) \frac{\Delta X_k^{l_w} \cdots \Delta X_k^{l_1}}{w!}, \tag{3.12}$$

where $\mathscr{V}_l$ is identified with the first-order differential operator $\sum_{q=1}^m V_l^q(y) \frac{\partial}{\partial y^q}$, $l = 1, \cdots, d$. Similar to the Runge–Kutta method, we construct a continuous version for the simplified step-N Euler scheme. For $t \in (t_k, t_{k+1}]$ define $\lfloor t \rfloor_n := t_k$ and for $t = 0$ define $\lfloor t \rfloor_n := 0$. Then a continuous version of (3.12) is

$$Y_t^n := y + \int_0^t \sum_{w,l_w,\cdots,l_1} \mathscr{V}_{l_w} \cdots \mathscr{V}_{l_1} I(Y_{\lfloor s \rfloor_n}^n) \frac{\left(X_{\lceil s \rceil^n}^{l_w} - X_{\lfloor s \rfloor_n}^{l_w}\right) \cdots \left(X_{\lceil s \rceil^n}^{l_2} - X_{\lfloor s \rfloor_n}^{l_2}\right)}{w!} \, \mathrm{d}X_s^{l_1}. \tag{3.13}$$

For simplicity we take $N = 2$ in the following. Indeed, our approach gives the same strong convergence rate $2H - \frac{1}{2}$ of simplified step-N Euler schemes for $N \geqslant 2$. We decompose the error by

$$\begin{aligned} Y_t - Y_t^n =& \left[\int_0^t V(Y_s) \, \mathrm{d}X_s - \int_0^t V(Y_s^n) \, \mathrm{d}X_s \right] + \left[\int_0^t V(Y_s^n) \, \mathrm{d}X_s - \int_0^t V(Y_{\lfloor s \rfloor_n}^n) \, \mathrm{d}X_s \right] \\ =&: L_t + R_t. \end{aligned}$$

If we apply the Taylor expansion $V(Y^n_s)$ at $V(Y^n_{\lfloor s\rfloor_n})$ directly, the corresponding leading order term of R_t contains second-level iterated integrals of X in the form of (2.4), i.e.,

$$\begin{aligned} R_t^{lead} &= \sum_{k=0}^{nt/T-1} \sum_{q,l,l'} \int_{t_k}^{t_{k+1}} \int_{t_k}^{s} \partial_q V_l(Y^n_{t_k,i}) V^q_{l'}(Y^n_{t_k,i})\, \mathrm{d}X^{l'}_u\, \mathrm{d}X^l_s \\ &= \frac{1}{2} \sum_{k=0}^{nt/T-1} \sum_{q,l,l'} \left[\int_{t_k}^{t_{k+1}} \int_{t_k}^{s} \partial_q V_l(Y^n_{t_k,i}) V^q_{l'}(Y^n_{t_k,i})\, \mathrm{d}X^{l'}_u\, \mathrm{d}X^l_s + \int_{t_k}^{t_{k+1}} \int_{u}^{t_{k+1}} \partial_q V_l(Y^n_{t_k,i}) V^q_{l'}(Y^n_{t_k,i})\, \mathrm{d}X^l_s\, \mathrm{d}X^{l'}_u \right] \\ &= \frac{1}{2} \sum_{k=0}^{nt/T-1} \sum_{q,l,l'} \left[\int_{t_k}^{t_{k+1}} \int_{t_k}^{s} \partial_q V_l(Y^n_{t_k,i}) V^q_{l'}(Y^n_{t_k,i})\, \mathrm{d}X^{l'}_u\, \mathrm{d}X^l_s + \int_{t_k}^{t_{k+1}} \int_{s}^{t_{k+1}} \partial_q V_l(Y^n_{t_k,i}) V^q_{l'}(Y^n_{t_k,i})\, \mathrm{d}X^l_u\, \mathrm{d}X^{l'}_s \right] \\ &\neq \frac{1}{2} \sum_{k=0}^{nt/T-1} \sum_{q,l,l'} \left[\int_{t_k}^{t_{k+1}} \int_{t_k}^{s} \partial_q V_l(Y^n_{t_k,i}) V^q_{l'}(Y^n_{t_k,i})\, \mathrm{d}X^{l'}_u\, \mathrm{d}X^l_s - \int_{t_k}^{t_{k+1}} \int_{s}^{t_{k+1}} \partial_q V_l(Y^n_{t_k,i}) V^q_{l'}(Y^n_{t_k,i})\, \mathrm{d}X^{l'}_u\, \mathrm{d}X^l_s \right], \\ & \qquad t = \lceil t \rceil^n, \end{aligned}$$

where the Fubini theorem is used in the second step. Roughly speaking it yields the convergence rate $2H - 1$.

To gain a sharp convergence rate, our goal is to show the corresponding leading order term of the error is in the forms of (2.2) and (2.3). The key idea is to compare it with the following 2-stage Runge–Kutta method (the Heun's method)

$$Z^n_{t_{k+1},1} = Z^n_{t_k}, \tag{3.14}$$

$$Z^n_{t_{k+1},2} = Z^n_{t_k} + V\left(Z^n_{t_{k+1},1}\right) \Delta X_k, \tag{3.15}$$

$$Z^n_{t_{k+1}} = Z^n_{t_k} + \frac{1}{2} V\left(Z^n_{t_{k+1},1}\right) \Delta X_k + \frac{1}{2} V(Z^n_{t_{k+1},2}) \Delta X_k, \tag{3.16}$$

which satisfies condition (3.10). We introduce two similar stage values for $Y^n_{t_{k+1}}$:

$$\begin{aligned} Y^n_{t_{k+1},1} &= Y^n_{t_k}, \\ Y^n_{t_{k+1},2} &= Y^n_{t_k} + V\left(Y^n_{t_{k+1},1}\right) \Delta X_k, \end{aligned}$$

and define the continuous versions:

$$Y^n_{t,1} = Y^n_{(t-h)\vee 0}, \tag{3.17}$$

$$Y^n_{t,2} = Y^n_{(t-h)\vee 0} + \int_{(t-h)\vee 0}^{t} V\left(Y^n_{\lceil s\rceil^n,1}\right) \mathrm{d}X_s. \tag{3.18}$$

Notice that

$$V_l(Y^n_{t_{k+1},2}) = V_l(Y^n_{t_k}) + \left[\int_0^1 \sum_{q,l'} \partial_q V_l\left(\theta Y^n_{t_{k+1},2} + (1-\theta) Y^n_{t_k}\right) \mathrm{d}\theta \right] V^q_{l'}(Y^n_{t_k}) \Delta X^{l'}_k$$

and

$$
\begin{aligned}
&\partial_q V_l\left(\theta Y^n_{t_{k+1},2}+(1-\theta)Y^n_{t_k}\right)\\
&\quad=\partial_q V_l(Y^n_{t_k})+\sum_{q',l''}\left[\int_0^1 \partial_{q'}\partial_q V_l\left(\theta'\left[\theta Y^n_{t_{k+1},2}+(1-\theta)Y^n_{t_k}\right]+(1-\theta')Y^n_{t_k}\right)\,\mathrm{d}\theta'\right]\left(\theta V^{q'}_{l''}(Y^n_{t_k})\Delta X^{l''}_k\right)\\
&\quad=:\partial_q V_l(Y^n_{t_k})+G^n_{q,l,k}(\theta).
\end{aligned}
$$

We have

$$
\begin{aligned}
V_l\left(Y^n_{t_{k+1},2}\right)&=V_l\left(Y^n_{t_k}\right)+\sum_{q,l'}\left[\partial_q V_l(Y^n_{t_k})+\int_0^1 G^n_{q,l,k}(\theta)\,\mathrm{d}\theta\right]V^q_{l'}(Y^n_{t_k})\Delta X^{l'}_k\\
&=V_l\left(Y^n_{t_k}\right)+\sum_{q,l'}\partial_q V_l\left(Y^n_{t_k}\right)V^q_{l'}\left(Y^n_{t_k}\right)\Delta X^{l'}_k+\sum_{q,l'}\left[\int_0^1 G^n_{q,l,k}(\theta)\,\mathrm{d}\theta\right]V^q_{l'}\left(Y^n_{t_k}\right)\Delta X^{l'}_k\\
&=:V_l\left(Y^n_{t_k}\right)+\sum_{q,l'}\partial_q V_l\left(Y^n_{t_k}\right)V^q_{l'}\left(Y^n_{t_k}\right)\Delta X^{l'}_k-G^n_{l,t_{k+1}}.
\end{aligned}
$$

Therefore, the simplified step-2 Euler scheme is rewritten in an implicit way:

$$
Y^n_{t_{k+1}}=Y^n_{t_k}+\frac{1}{2}V\left(Y^n_{t_{k+1},1}\right)\Delta X_k+\frac{1}{2}V\left(Y^n_{t_{k+1},2}\right)\Delta X_k+\frac{1}{2}\sum_l G^n_{l,t_{k+1}}\Delta X^l_k, \tag{3.19}
$$

$$
Y^n_t=y+\frac{1}{2}\int_0^t V\left(Y^n_{\lceil s\rceil^n,1}\right)\,\mathrm{d}X^l_s+\frac{1}{2}\int_0^t V\left(Y^n_{\lceil s\rceil^n,2}\right)\,\mathrm{d}X^l_s+\frac{1}{2}\sum_l\int_0^t G^n_{l,\lceil s\rceil^n}\,\mathrm{d}X^l_s, \tag{3.20}
$$

where $\frac{1}{2}\sum_l\int_0^t G^n_{l,\lceil s\rceil^n}\,\mathrm{d}X^l_s$ contains third-level iterated integrals of X.

Based on (3.19) and (3.20), and arguments in Subsection 3.1, we obtain that the simplified step-2 Euler scheme has the same leading order term as scheme (3.14)–(3.16), i.e.,

$$
\begin{aligned}
R^{lead}_t=\frac{1}{4}\sum_{k=0}^{nt/T-1}\sum_{i,q,l,l'}\Bigg[&\int_{t_k}^{t_{k+1}}\int_{t_k}^{s}\partial_q V_l\left(Y^n_{t_{k+1},i}\right)V^q_{l'}\left(Y^n_{t_{k+1},i}\right)\,\mathrm{d}X^{l'}_u\,\mathrm{d}X^l_s\\
&-\int_{t_k}^{t_{k+1}}\int_{s}^{t_{k+1}}\partial_q V_l\left(Y^n_{t_{k+1},i}\right)V^q_{l'}\left(Y^n_{t_{k+1},i}\right)\,\mathrm{d}X^{l'}_u\,\mathrm{d}X^l_s\Bigg].
\end{aligned} \tag{3.21}
$$

Thus, we get the same strong convergence rate for the simplified step-2 Euler scheme as in Theorem 3.2.

THEOREM 3.3 If $N\geqslant 2$, $V\in\mathscr{C}^{N+1}_b(\mathbb{R}^m;\mathbb{R}^{m\times d})$ and $H>1/2$ then the strong convergence order of simplified step-N Euler scheme for (1.1) is $2H-\frac{1}{2}$. More precisely, there exists a constant C independent of n such that

$$\left\| \sup_{t\in[0,T]} |Y_t - Y_t^n| \right\|_{L^p(\Omega)} \leqslant Ch^{2H-\frac{1}{2}}, \quad p \geqslant 1,$$

where $h = \frac{T}{n}$ and Y_t^n is defined by (3.13).

Based on Theorem 3.3, we consider linear interpolation of $Y_{t_k}^n$, i.e.,

$$Y_t^{n,linear} := Y_{t_k}^n + \frac{t - t_k}{h}\left(Y_{t_{k+1}}^n - Y_{t_k}^n\right), \quad \forall\, t \in (t_k, t_{k+1}]. \tag{3.22}$$

Then

$$\begin{aligned}\left\| \sup_{t\in[0,T]} |Y_t - Y_t^{n,linear}| \right\|_{L^p(\Omega)} &\leqslant \left\| \sup_{t\in[0,T]} |Y_t - Y_t^n| \right\|_{L^p(\Omega)} + \left\| \sup_{t\in[0,T]} |Y_t^n - Y_t^{n,linear}| \right\|_{L^p(\Omega)} \\ &\leqslant Ch^{2H-\frac{1}{2}} + C\left\| \sup_{t\in[0,T]} \left| X_t - X_{\lfloor t\rfloor_n} - \frac{t-\lfloor t\rfloor_n}{h}\left(X_{\lceil t\rceil^n} - X_{\lfloor t\rfloor_n}\right) \right| \right\|_{L^{2p}(\Omega)} \\ &\leqslant C\left(h^{2H-\frac{1}{2}} + h^H\sqrt{\log\frac{T}{h}}\right), \quad p \geqslant 1,\end{aligned}$$

where the last inequality follows from Hüsler *et al.* (2003, Theorem 6). This result gives an answer to a conjecture in Deya *et al.* (2012) for the simplified step-2 Euler scheme when $H > \frac{1}{2}$. Indeed, our approach to determining the leading order term of the error of the simplified step-2 Euler scheme is available for general schemes constructed by second-order Taylor expansion. As a consequence, these schemes share the same strong convergence rate $2H - \frac{1}{2}$.

REMARK 3.6 Our result indicates that the simplified step-2 Euler scheme is superior to the classical Euler method (Mishura & Shevchenko, 2008), whose convergence rate is $2H-1$, $H \in (\frac{1}{2}, 1)$. Compared with the optimal convergence rate γ of the modified Euler scheme Hu *et al.*, (2016a) where $\gamma = 2H - \frac{1}{2}$ when $H \in (\frac{1}{2}, \frac{3}{4})$, $\gamma = 1^-$ when $H = \frac{3}{4}$ and $\gamma = 1$ when $H \in (\frac{3}{4}, 1)$, the simplified step-2 Euler scheme has higher convergence rate for $H \in [\frac{3}{4}, 1)$.

4. Proof of Theorems 3.2 and 3.3

In the following we establish the continuous dependence of the numerical solutions on the driving noises for Runge–Kutta methods in Section 4.1 and for simplified step-N Euler schemes in Section 4.2, respectively. In Section 4.3 we combine the continuous dependence with the leading order term of the error to prove the main theorems, Theorem 3.2 and 3.3.

4.1 A priori *estimates for Runge–Kutta methods*

For $f \in \mathscr{C}([0,T],\mathbb{R}^d)$ we introduce the discrete Hölder semi-norm

$$\|f\|_{s,t,\beta,n} := \sup\left\{\frac{|f_v - f_u|}{(v-u)^\beta},\ s \leqslant u < v \leqslant t,\ u = \lceil u \rceil^n,\ v = \lceil v \rceil^n\right\}$$

and use $\|f\|_{\beta,n} := \|f\|_{0,T,\beta,n}$ for short.

LEMMA 4.1 If $V \in \mathscr{C}_b^0(\mathbb{R}^m;\mathbb{R}^{m\times d})$ then for any $n \in \mathbb{N}_+$ and $1/2 < \beta < H$, $\|Y_\cdot^n\|_{\beta,n}$ and $\|Y_{\cdot,i}^n\|_{\beta,n}$ are finite almost surely, $i = 1,\cdots,\mathbf{s}$. More precisely

$$\|Y_\cdot^n\|_{\beta,n} \leqslant C(d,m,n,c,\nu,\mathbf{s})\|X\|_\beta < \infty, \quad a.s.,$$
$$\|Y_{\cdot,i}^n\|_{\beta,n} \leqslant C(d,m,n,c,\nu,\mathbf{s})\|X\|_\beta < \infty, \quad a.s.,$$

with $c := \max\{|a_{ij}|,|b_i| : i,j = 1,\cdots,\mathbf{s}\}$ and $\nu := \sup_{y\in\mathbb{R}^m}|V(y)|$.

Proof. Recall that

$$Y_t^n = y + \sum_{i=1}^{\mathbf{s}} \int_0^t b_i V(Y^n_{\lceil s\rceil^n,i})\,\mathrm{d}X_s,$$
$$Y_{t,i}^n = Y^n_{(t-h)\vee 0} + \sum_{j=1}^{\mathbf{s}} \int_{(t-h)\vee 0}^t a_{ij} V(Y^n_{\lceil s\rceil^n,j})\,\mathrm{d}X_s, \quad i = 1,\cdots,\mathbf{s}.$$

Due to the boundedness of V we have that for $u = \lceil u\rceil^n = t_k$ and $v = \lceil v\rceil^n = t_l$,

$$\begin{aligned}|Y_u^n - Y_v^n| &\leqslant \sum_{i=1}^{\mathbf{s}} \left|\int_{t_l}^{t_k} b_i V(Y^n_{\lceil s\rceil^n,i})\,\mathrm{d}X_s\right| \\ &\leqslant \sum_{i=1}^{\mathbf{s}}\sum_{p=l}^{k-1} |b_i| \left|V\left(Y^n_{t_{p+1},i}\right)\right| |X_{t_{p+1}} - X_{t_p}| \\ &\leqslant \sum_{i=1}^{\mathbf{s}}\sum_{p=l}^{k-1} |b_i| \left|V\left(Y^n_{t_{p+1},i}\right)\right| \|X\|_\beta h^\beta,\end{aligned}$$

which implies

$$\frac{|Y_u^n - Y_v^n|}{|u-v|^\beta} \leqslant \frac{C(d,m,n,c,\nu,\mathbf{s})\|X\|_\beta h^\beta}{|u-v|^\beta} \leqslant C(d,m,n,c,\nu,\mathbf{s})\|X\|_\beta.$$

The estimate for $Y_{t,i}^n$ can also be obtained by a similar procedure. □

Inspired by Hu *et al.* (2016a) we give the following two lemmas as *a priori* estimates for the continuous versions (3.3) and (3.4). Further, Proposition 4.1 shows the bounds for the stage values $Y^n_{\cdot,i}$ and the numerical solution $Y^n_{\cdot}$ in terms of the driving noises in Hölder semi-norm.

LEMMA 4.2 Let α, β and β' satisfy $1-\beta<\alpha<\beta'$. Then for any $s,t\in[0,T]$ satisfying $s<t$ and $s=\lceil s\rceil^n$ there exists a constant $C=C(\alpha,\beta,\beta',T)$ such that

$$\int_s^t (t-r)^{\alpha+\beta-1}\int_s^r \frac{(\lceil r\rceil^n-\lceil u\rceil^n)^{\beta'}}{(r-u)^{\alpha+1}}\,\mathrm{d}u\,\mathrm{d}r \leqslant C(t-s)^{\beta+\beta'}.$$

Proof. Without loss of generality suppose $T=1$. By the definition of $\lceil\cdot\rceil^n$ in Subsection 3.1, we have

$$\begin{aligned}
&\int_s^t (t-r)^{\alpha+\beta-1}\int_s^r \frac{\left(\lceil r\rceil^n-\lceil u\rceil^n\right)^{\beta'}}{(r-u)^{\alpha+1}}\,\mathrm{d}u\,\mathrm{d}r\\
&=\int_{\lceil s\rceil^n+1/n}^t (t-r)^{\alpha+\beta-1}\int_{\lceil s\rceil^n}^{\lceil r\rceil^n-\frac{1}{n}} \frac{\left(\lceil r\rceil^n-\lceil u\rceil^n\right)^{\beta'}}{(r-u)^{\alpha+1}}\,\mathrm{d}u\,\mathrm{d}r\\
&=\int_{\lceil s\rceil^n+1/n}^t (t-r)^{\alpha+\beta-1}\left(\int_{\lceil s\rceil^n}^{\lceil r\rceil^n-\frac{2}{n}}+\int_{\lceil r\rceil^n-\frac{2}{n}}^{\lceil r\rceil^n-\frac{1}{n}}\right)\frac{\left(\lceil r\rceil^n-\lceil u\rceil^n\right)^{\beta'}}{(r-u)^{\alpha+1}}\,\mathrm{d}u\,\mathrm{d}r\\
&=: I_1+I_2.
\end{aligned}$$

For the first term, since $r-u>\frac{1}{n}$ and $\lceil r\rceil^n-\lceil u\rceil^n<r-u+\frac{1}{n}$, we have

$$\begin{aligned}
I_1&=\int_{\lceil s\rceil^n+1/n}^t (t-r)^{\alpha+\beta-1}\int_{\lceil s\rceil^n}^{\lceil r\rceil^n-\frac{2}{n}} \frac{\left(\lceil r\rceil^n-\lceil u\rceil^n\right)^{\beta'}}{(r-u)^{\alpha+1}}\,\mathrm{d}u\,\mathrm{d}r\\
&\leqslant \int_{\lceil s\rceil^n+1/n}^t (t-r)^{\alpha+\beta-1}\int_{\lceil s\rceil^n}^{\lceil r\rceil^n-\frac{2}{n}} \frac{2^{\beta'}(r-u)^{\beta'}}{(r-u)^{\alpha+1}}\,\mathrm{d}u\,\mathrm{d}r\\
&\leqslant C\int_{\lceil s\rceil^n+1/n}^t (t-r)^{\alpha+\beta-1}(r-s)^{\beta'-\alpha}\,\mathrm{d}r\\
&\leqslant C\int_{\lceil s\rceil^n+1/n}^t (t-r)^{\alpha+\beta-1}(t-s)^{\beta'-\alpha}\,\mathrm{d}r\\
&\leqslant C(t-s)^{\beta+\beta'}.
\end{aligned}$$

For the second term

$$
\begin{aligned}
I_2 &= \int_{\lceil s\rceil^n+1/n}^{t} (t-r)^{\alpha+\beta-1} \int_{\lceil r\rceil^n-\frac{2}{n}}^{\lceil r\rceil^n-\frac{1}{n}} \frac{\left(\lceil r\rceil^n - \lceil u\rceil^n\right)^{\beta'}}{(r-u)^{\alpha+1}}\,\mathrm{d}u\,\mathrm{d}r \\
&\leqslant C(t-s)^{\alpha+\beta-1}\left(\frac{2}{n}\right)^{\beta'} \int_{\lceil s\rceil^n+1/n}^{t} \left(r-\lceil r\rceil^n+\frac{1}{n}\right)^{-\alpha}\mathrm{d}r \\
&\leqslant C(t-s)^{\alpha+\beta-1}\left(\frac{2}{n}\right)^{\beta'}\left(\frac{t-s}{\frac{1}{n}}\right)\left(\frac{1}{n}\right)^{-\alpha+1} \\
&\leqslant C(t-s)^{\beta+\beta'}.
\end{aligned}
$$

□

LEMMA 4.3 Let β and β' satisfy $\beta+\beta' > 1$. Let $g \in \mathscr{C}_b^1(\mathbb{R}^m;\mathbb{R})$, $x \in \mathscr{C}([s,t];\mathbb{R})$ and $z \in \mathscr{C}([s,t];\mathbb{R}^m)$. If $\|x\|_\beta$ and $\|z\|_{\beta',n}$ are all finite for any $n \in \mathbb{N}_+$ then for $s = \lceil s\rceil^n$ and $t = \lceil t\rceil^n$,

$$
\left|\int_s^t g(z_{\lceil r\rceil^n})\,\mathrm{d}X_r\right| \leqslant C(g,\beta,\beta',T)(1+\|z\|_{s,t,\beta',n}(t-s)^{\beta'})\|x\|_\beta(t-s)^\beta.
$$

Proof. Considering the equivalence of norms in $\mathbb{R}^m$ we suppose $m=1$ without loss of generality. Let α satisfy $\alpha < \beta'$ and $\beta+\alpha > 1$. According to the characterization of the integral (2.1)

$$
\int_s^t g(z_{\lceil r\rceil^n})\,\mathrm{d}X_r = (-1)^\alpha \int_s^t D_{s+}^\alpha g(z_{\lceil r\rceil^n}) D_{t-}^{1-\alpha}(x_r - x_t)\,\mathrm{d}r.
$$

Combining the fractional Weyl derivatives, we have that for $s < r < t$,

$$
\begin{aligned}
\left|D_{s+}^\alpha g(z_{\lceil r\rceil^n})\right| &\leqslant C\left(\left|\frac{g(z_{\lceil r\rceil^n})}{(r-s)^\alpha}\right| + \int_s^r \frac{\left|g(z_{\lceil r\rceil^n}) - g(z_{\lceil u\rceil^n})\right|}{(r-u)^{\alpha+1}}\,\mathrm{d}u\right) \\
&\leqslant C\left(\frac{1}{(r-s)^\alpha} + \int_s^r \frac{\left|z_{\lceil r\rceil^n} - z_{\lceil u\rceil^n}\right|}{(r-u)^{\alpha+1}}\,\mathrm{d}u\right) \\
&\leqslant C\left(\frac{1}{(r-s)^\alpha} + \|z\|_{s,t,\beta',n}\int_s^r \frac{\left(\lceil r\rceil^n - \lceil u\rceil^n\right)^{\beta'}}{(r-u)^{\alpha+1}}\,\mathrm{d}u\right)
\end{aligned}
$$

and

$$\left|D_{t-}^{1-\alpha}(x_r - x_t)\right| \leqslant C\left(\left|\frac{x_r - x_t}{(t-r)^{1-\alpha}}\right| + \int_r^t \frac{\left|x_r - x_u\right|}{(u-r)^{2-\alpha}}\,\mathrm{d}u\right)$$
$$\leqslant C\|x\|_\beta (t-r)^{\alpha+\beta-1}.$$

Using Lemma 4.2 we obtain

$$\left|\int_s^t g(z_{\lceil r\rceil^n})\,\mathrm{d}X_r\right| \leqslant \int_s^t \left|D_{s+}^{\alpha} g(z_{\lceil r\rceil^n}) D_{t-}^{1-\alpha}(x_r - x_t)\right| \mathrm{d}r$$
$$\leqslant C\int_s^t \left(\frac{1}{(r-s)^\alpha} + \|z\|_{s,t,\beta',n}\int_s^r \frac{\left(\lceil r\rceil^n - \lceil u\rceil^n\right)^{\beta'}}{(r-u)^{\alpha+1}}\,\mathrm{d}u\right)\|x\|_\beta (t-r)^{\alpha+\beta-1}\,\mathrm{d}r$$
$$\leqslant C(1 + \|z\|_{s,t,\beta',n}(t-s)^{\beta'})\|x\|_\beta (t-s)^\beta.$$

□

Proposition 4.1 Let $1/2 < \beta < H$. If $V \in \mathscr{C}_b^1(\mathbb{R}^m;\mathbb{R}^{m\times d})$, then for any $n \in \mathbb{N}_+$,

$$\sum_{i=1}^{\mathbf{s}} \|Y_{\cdot,i}^n\|_\beta \leqslant C(c,\mathbf{s},V,\beta,T)\max\left\{\|X\|_\beta, \|X\|_\beta^{1/\beta}\right\},$$
$$\sum_{i=1}^{\mathbf{s}} \|Y_{\cdot,i}^n\|_\infty \leqslant \mathbf{s}|y| + C(c,\mathbf{s},V,\beta,T)\max\left\{\|X\|_\beta, \|X\|_\beta^{1/\beta}\right\},$$
$$\|Y_\cdot^n\|_\beta \leqslant C(c,\mathbf{s},V,\beta,T)\max\left\{\|X\|_\beta, \|X\|_\beta^{1/\beta+1}\right\},$$
$$\|Y_\cdot^n\|_\infty \leqslant |y| + C(c,\mathbf{s},V,\beta,T)\max\left\{\|X\|_\beta, \|X\|_\beta^{1/\beta+1}\right\},$$

where $c = \max\{|a_{ij}|, |b_i| : i,j = 1,\cdots,\mathbf{s}\}$.

Moreover, there exists some $C_0 > 0$ such that $0 \leqslant s < t \leqslant T$ and $\|X\|_\beta |t-s|^\beta \leqslant C_0$ imply

$$\sum_{i=1}^{\mathbf{s}} \|Y_{\cdot,i}^n\|_{s,t,\beta} \leqslant C(c,\mathbf{s},V,\beta,T,C_0)\|X\|_\beta,$$

$$\|Y_\cdot^n\|_{s,t,\beta} \leqslant C(c,\mathbf{s},V,\beta,T,C_0)\|X\|_\beta.$$

Proof. We first take $s, t \in [0, T]$ satisfying $s = \lceil s \rceil^n$ and $t = \lceil t \rceil^n$. Lemma 4.3 yields

$$\begin{aligned}
\left|Y^n_{t,i} - Y^n_{s,i}\right| \leqslant \sum_{j=1}^{\mathbf{s}} \Bigg| & \int_0^{(t-h)\vee 0} b_j V\left(Y^n_{\lceil r\rceil^n,j}\right) \,\mathrm{d}X_r + \int_{(t-h)\vee 0}^{t} a_{ij} V\left(Y^n_{\lceil r\rceil^n,j}\right) \,\mathrm{d}X_r \\
& - \int_0^{(s-h)\vee 0} b_j V\left(Y^n_{\lceil r\rceil^n,j}\right) \,\mathrm{d}X_r - \int_{(s-h)\vee 0}^{s} a_{ij} V\left(Y^n_{\lceil r\rceil^n,j}\right) \,\mathrm{d}X_r \Bigg| \\
\leqslant \sum_{j=1}^{\mathbf{s}} & \left| \int_s^{(t-h)\vee 0} b_j V\left(Y^n_{\lceil r\rceil^n,j}\right) \,\mathrm{d}X_r \right| + \left| \int_{(t-h)\vee 0}^{t} a_{ij} V\left(Y^n_{\lceil r\rceil^n,j}\right) \,\mathrm{d}X_r \right| \\
& + \left| \int_{(s-h)\vee 0}^{s} b_j V\left(Y^n_{\lceil r\rceil^n,j}\right) \,\mathrm{d}X_r \right| + \left| \int_{(s-h)\vee 0}^{s} a_{ij} V\left(Y^n_{\lceil r\rceil^n,j}\right) \,\mathrm{d}X_r \right| \\
\leqslant & C(c, V, \beta, T) \left(1 + \sum_{j=1}^{\mathbf{s}} \|Y^n_{\cdot,j}\|_{s,t,\beta,n} (t-s)^{\beta} \right) \|X\|_{\beta} (t-s)^{\beta}.
\end{aligned}$$

Summing up above inequalities over $i = 1, \cdots, \mathbf{s}$ and dividing both sides by $(t-s)^{\beta}$, we have

$$\begin{aligned}
\sum_{i=1}^{\mathbf{s}} \|Y^n_{\cdot,i}\|_{s,t,\beta,n} &\leqslant (C(c, \mathbf{s}, V, \beta, T) \vee 1) \left(1 + \sum_{i=1}^{\mathbf{s}} \|Y^n_{\cdot,i}\|_{s,t,\beta,n} (t-s)^{\beta} \right) \|X\|_{\beta} \\
&=: C_1 \left(1 + \sum_{i=1}^{\mathbf{s}} \|Y^n_{\cdot,i}\|_{s,t,\beta,n} (t-s)^{\beta} \right) \|X\|_{\beta}.
\end{aligned}$$

If $n \geqslant 2T(2C_1\|X\|_{\beta})^{1/\beta}$ then there exist $N_0 \in \mathbb{N}_+$ and $N_1 := \frac{N_0 T}{n}$ such that

$$(2C_1\|X\|_{\beta})^{-1/\beta} \leqslant 2N_1 \leqslant 2(2C_1\|X\|_{\beta})^{-1/\beta}.$$

When $t - s = N_1$, considering the choice for N_1, we get

$$\sum_{i=1}^{\mathbf{s}} \|Y^n_{\cdot,i}\|_{s,t,\beta,n} \leqslant 2C_1\|X\|_{\beta}.$$

When $t - s > N_1$,

$$\sum_{i=1}^{\mathbf{s}} \left|Y^n_{t,i} - Y^n_{s,i}\right| \leqslant C\left(\left[\frac{t-s}{N_1}\right] + 1\right) \sup_{r=\lceil r\rceil^n \leqslant t_{n-1}} \sum_{i=1}^{\mathbf{s}} \|Y^n_{\cdot,i}\|_{r,r+N_1,\beta,n} N_1^{\beta},$$

where $[t]$ means the largest integer that is not larger than t. Then we obtain

$$\begin{aligned}\sum_{i=1}^{\mathbf{s}} \|Y^n_{\cdot,i}\|_{s,t,\beta,n} &\leqslant C\left(\frac{T}{N_1}+1\right)^{1-\beta} C_1\|X\|_\beta \\ &\leqslant C\left(\|X\|_\beta^{1/\beta-1}+1\right)\|X\|_\beta \\ &\leqslant C\max\left\{\|X\|_\beta, \|X\|_\beta^{1/\beta}\right\},\end{aligned}$$

where the second inequality is from the choice of N_1. If $n < 2T(2C_1\|X\|_\beta)^{1/\beta}$, the definition of $Y^n_{\cdot,i}$ leads to

$$\sum_{i=1}^{\mathbf{s}} |Y^n_{t,i} - Y^n_{s,i}| \leqslant C(t-s)\left(\frac{T}{n}\right)^{-1+\beta}\|X\|_\beta,$$

and thus

$$\sum_{i=1}^{\mathbf{s}} \|Y^n_{\cdot,i}\|_{s,t,\beta,n} \leqslant C\|X\|_\beta^{1/\beta}.$$

Then for $0 \leqslant s < t \leqslant T$, it holds that

$$\sum_{i=1}^{\mathbf{s}} \|Y^n_{\cdot,i}\|_{s,t,\beta} \leqslant C\|X\|_\beta + \sum_{i=1}^{\mathbf{s}} \|Y^n_{\cdot,i}\|_{\lceil s\rceil^n,\lfloor t\rfloor_n,\beta,n} \leqslant C\max\left\{\|X\|_\beta, \|X\|_\beta^{1/\beta}\right\}.$$

Therefore,

$$\begin{aligned}\|Y^n_{\cdot,i}\|_\beta &\leqslant C(c,\mathbf{s},V,\beta,T)\max\left\{\|X\|_\beta, \|X\|_\beta^{1/\beta}\right\},\\ \|Y^n_{\cdot,i}\|_\infty &\leqslant |y| + C(c,\mathbf{s},V,\beta,T)\max\left\{\|X\|_\beta, \|X\|_\beta^{1/\beta}\right\}.\end{aligned}$$

Moreover, if $C_0 \in (0, 1/C_1)$ then for $0 \leqslant s < t \leqslant T$ such that $\|X\|_\beta |t-s|^\beta \leqslant C_0$, it can be improved to

$$\sum_{i=1}^{\mathbf{s}} \|Y^n_{\cdot,i}\|_{s,t,\beta} \leqslant C(C_1,\beta,T,C_0)\|X\|_\beta.$$

Considering the numerical solution Y^n since

$$
\begin{aligned}
\left|Y_t^n - Y_s^n\right| &\leqslant \sum_{i=1}^{\mathbf{s}} \left| \int_s^t b_i V\left(Y^n_{\lceil r \rceil^n, j}\right) \, \mathrm{d}X_r \right| \\
&\leqslant C\left(1 + \sum_{j=1}^{\mathbf{s}} \|Y^n_{\cdot j}\|_{s,t,\beta,n}(t-s)^\beta\right) \|X\|_\beta (t-s)^\beta \\
&\leqslant C(c, \mathbf{s}, V, \beta, T) \max\left\{\|X\|_\beta, \|X\|_\beta^{1/\beta+1}\right\}(t-s)^\beta,
\end{aligned}
$$

we obtain that

$$
\begin{aligned}
\|Y^n_\cdot\|_\beta &\leqslant C(c, \mathbf{s}, V, \beta, T) \max\left\{\|X\|_\beta, \|X\|_\beta^{1/\beta+1}\right\}, \\
\|Y^n_\cdot\|_\infty &\leqslant |y| + C(c, \mathbf{s}, V, \beta, T) \max\left\{\|X\|_\beta, \|X\|_\beta^{1/\beta+1}\right\},
\end{aligned}
$$

and for $0 \leqslant s < t \leqslant T$ such that $\|X\|_\beta |t-s|^\beta \leqslant C_0 \in (0, 1/C_1)$,

$$
\|Y^n_\cdot\|_{s,t,\beta} \leqslant C(c, \mathbf{s}, V, \beta, T, C_0)\|X\|_\beta.
$$

□

REMARK 4.1 Note that Fernique's lemma implies that $\big\| \|X\|_\beta^N \big\|_{L^p(\Omega)} < \infty$ for any $p \geqslant 1$ and $N \in \mathbb{N}_+$; see e.g., Hu *et al.* (2016a, Remark 3.2).

4.2 A priori *estimates simplified step-N Euler schemes*

Using similar arguments in the proof of Lemmas 4.1–4.3, we obtain Lemmas 4.4–4.6. Then we obtain the bound for $Y^n_\cdot$ defined in (3.13).

LEMMA 4.4 If $V \in \mathscr{C}_b^{N-1}(\mathbb{R}^m; \mathbb{R}^{m\times d})$ then for any $n \in \mathbb{N}_+$ and $1/2 < \beta < H$, $\|Y^n_\cdot\|_{\beta,n}$ are all finite almost surely.

LEMMA 4.5 Let α, β and β' satisfy $1 - \beta < \alpha < \beta'$. Then for any $s, t \in [0, T]$ satisfying $s < t$ and $s = \lceil s \rceil^n$, there exists a constant $C = C(\alpha, \beta, \beta', T)$ such that

$$
\int_s^t (t-r)^{\alpha+\beta-1} \int_s^r \frac{(\lfloor r \rfloor_n - \lfloor u \rfloor_n)^{\beta'}}{(r-u)^{\alpha+1}} \, \mathrm{d}u \, \mathrm{d}r \leqslant K(t-s)^{\beta+\beta'}.
$$

LEMMA 4.6 Let β and β' satisfy $\beta + \beta' > 1$. Let $g \in \mathscr{C}_b^1(\mathbb{R}^m; \mathbb{R})$, $x^l \in \mathscr{C}([s, t]; \mathbb{R})$, $l \in \{1, \cdots, d\}$ and $z \in \mathscr{C}([s, t]; \mathbb{R}^m)$. If $\|x^l\|_\beta$, $\|z\|_{\beta',n}$ are all finite for any $l \in \{1, \cdots, d\}$, $n \in \mathbb{N}_+$ then for $w \geqslant 2$, $s = \lceil s \rceil^n$

and $t = \lceil t \rceil^n$,

$$\left| \int_s^t g(z_{\lfloor r \rfloor_n}) \left(x^{l_w}_{\lceil r \rceil_n} - x^{l_w}_{\lfloor r \rfloor_n} \right) \cdots \left(x^{l_2}_{\lceil r \rceil^n} - x^{l_2}_{\lfloor r \rfloor_n} \right) \, \mathrm{d}X^{l_1}_r \right|$$
$$\leqslant C(g, \beta, \beta', T) \left(1 + \|z\|_{s,t,\beta',n} (t-s)^{\beta'} \right) \|x^{l_w}\|_\beta \cdots \|x^{l_1}\|_\beta (t-s)^{\beta w},$$

where $l_1, \cdots, l_w \in \{1, \cdots, d\}$.

Proof. Let $\Phi_r = g(z_{\lfloor r \rfloor_n})(x^{l_w}_{\lceil r \rceil^n} - x^{l_w}_{\lfloor r \rfloor_n}) \cdots (x^{l_2}_{\lceil r \rceil^n} - x^{l_2}_{\lfloor r \rfloor_n})$. Taking α such that $\alpha < \beta'$ and $\beta + \alpha > 1$ we estimate the left fractional Weyl derivative of Φ via

$$|D^\alpha_{s+} \Phi_r| \leqslant \left| \frac{\Phi_r}{(r-s)^\alpha} \right| + \alpha \int_s^r \frac{|\Phi_r - \Phi_u|}{(r-u)^{\alpha+1}} \, \mathrm{d}u.$$

For the first term we have

$$\left| \frac{\Phi_r}{(r-s)^\alpha} \right| \leqslant C \|x^{l_w}\|_\beta \cdots \|x^{l_2}\|_\beta \left(\frac{T}{n} \right)^{\beta(w-1)} (r-s)^{-\alpha}.$$

For the second term we decompose $\Phi_r - \Phi_u$ into

$$\begin{aligned}
&\Big[g(z_{\lfloor r \rfloor_n}) \cdots \left(x^{l_3}_{\lceil r \rceil^n} - x^{l_3}_{\lfloor r \rfloor_n} \right) \left(x^{l_2}_{\lceil r \rceil^n} - x^{l_2}_{\lfloor r \rfloor_n} \right) - g(z_{\lfloor r \rfloor_n}) \cdots \left(x^{l_3}_{\lceil r \rceil^n} - x^{l_3}_{\lfloor r \rfloor_n} \right) \left(x^{l_2}_{\lceil u \rceil^n} - x^{l_2}_{\lfloor u \rfloor_n} \right) \Big] \\
&\quad + \Big[g(z_{\lfloor r \rfloor_n}) \cdots \left(x^{l_3}_{\lceil r \rceil^n} - x^{l_3}_{\lfloor r \rfloor_n} \right) \left(x^{l_2}_{\lceil u \rceil^n} - x^{l_2}_{\lfloor u \rfloor_n} \right) - g(z_{\lfloor r \rfloor_n}) \cdots \left(x^{l_3}_{\lceil u \rceil^n} - x^{l_3}_{\lfloor u \rfloor_n} \right) \left(x^{l_2}_{\lceil u \rceil^n} - x^{l_2}_{\lfloor u \rfloor_n} \right) \Big] \\
&\quad + \cdots \\
&\quad + \Big[g(z_{\lfloor r \rfloor_n}) \left(x^{l_w}_{\lceil u \rceil^n} - x^{l_w}_{\lfloor u \rfloor_n} \right) \cdots \left(x^{l_2}_{\lceil u \rceil^n} - x^{l_2}_{\lfloor u \rfloor_n} \right) - g(z_{\lfloor u \rfloor_n}) \left(x^{l_w}_{\lceil u \rceil^n} - x^{l_w}_{\lfloor u \rfloor_n} \right) \cdots \left(x^{l_2}_{\lceil u \rceil^n} - x^{l_2}_{\lfloor u \rfloor_n} \right) \Big] \\
&\qquad =: I_1 + I_2 \cdots + I_w.
\end{aligned}$$

We analyse each of them by

$$|I_1| \leqslant C \|x^{l_w}\|_\beta \cdots \|x^{l_2}\|_\beta \left(\frac{T}{n} \right)^{\beta(w-2)} \Big[(\lceil r \rceil^n - \lceil u \rceil^n)^\beta + (\lfloor r \rfloor_n - \lfloor u \rfloor_n)^\beta \Big],$$

$$|I_2| \leqslant C \|x^{l_w}\|_\beta \cdots \|x^{l_2}\|_\beta \left(\frac{T}{n} \right)^{\beta(w-2)} \Big[(\lceil r \rceil^n - \lceil u \rceil^n)^\beta + (\lfloor r \rfloor_n - \lfloor u \rfloor_n)^\beta \Big],$$

$$\cdots$$

$$|I_{w-1}| \leqslant C \|x^{l_w}\|_\beta \cdots \|x^{l_2}\|_\beta \left(\frac{T}{n} \right)^{\beta(w-2)} \Big[(\lceil r \rceil^n - \lceil u \rceil^n)^\beta + (\lfloor r \rfloor_n - \lfloor u \rfloor_n)^\beta \Big]$$

and

$$
\begin{aligned}
|I_w| \leqslant& \left| (g(z_{\lfloor r \rfloor_n}) - g(z_{\lfloor u \rfloor_n})) \left(x^{l_w}_{\lceil u \rceil^n} - x^{l_w}_{\lfloor u \rfloor_n} \right) \cdots \left(x^{l_2}_{\lceil u \rceil^n} - x^{l_2}_{\lfloor u \rfloor_n} \right) \right| \\
\leqslant& C \left| (z_{\lfloor r \rfloor_n} - z_{\lfloor u \rfloor_n}) \left(x^{l_w}_{\lceil u \rceil^n} - x^{l_w}_{\lfloor u \rfloor_n} \right) \cdots \left(x^{l_2}_{\lceil u \rceil^n} - x^{l_2}_{\lfloor u \rfloor_n} \right) \right| \\
\leqslant& C \|z\|_{s,t,\beta',n} (\lfloor r \rfloor_n - \lfloor u \rfloor_n)^{\beta'} \|x^{l_w}\|_\beta \cdots \|x^{l_2}\|_\beta \left(\frac{T}{n} \right)^{\beta(w-1)}.
\end{aligned}
$$

Combining Lemma 4.5 and arguments in Lemma 4.3 we conclude the proof. □

PROPOSITION 4.2 Let $1/2 < \beta < H$. If $V \in \mathscr{C}_b^N(\mathbb{R}^m; \mathbb{R}^{m\times d})$, then for any $n \in \mathbb{N}_+$,

$$
\|Y^n\|_\beta \leqslant C(N, V, \beta, T) \max \left\{ \|X\|_\beta, \|X\|_\beta^{N-1+1/\beta} \right\},
$$

$$
\|Y^n\|_\infty \leqslant |y| + C(N, V, \beta, T) \max \left\{ \|X\|_\beta, \|X\|_\beta^{N-1+1/\beta} \right\}.
$$

Moreover, there exists $C_0 > 0$ such that $0 \leqslant s < t \leqslant T$ and $\|X\|_\beta |t-s|^\beta \leqslant C_0$ imply

$$
\|Y^n\|_{s,t,\beta} \leqslant C(N, V, \beta, T, C_0) \|X\|_\beta.
$$

Proof. Take $s, t \in [0, T]$ such that $s = \lceil s \rceil^n$ and $t = \lceil t \rceil^n$. Lemma 4.6 yields

$$
\begin{aligned}
\left| Y^n_t - Y^n_s \right| \leqslant& C(N, V, \beta, T) \sum_{w=1}^N \|X\|_\beta^w (t-s)^{\beta w} \left[1 + \|Y^n\|_{s,t,\beta,n} (t-s)^\beta \right] \\
=:& C_1 \sum_{w=1}^N \|X\|_\beta^w (t-s)^{\beta w} \left[1 + \|Y^n\|_{s,t,\beta,n} (t-s)^\beta \right],
\end{aligned}
$$

since $\|X^{l_w}\|_\beta \cdots \|X^{l_1}\|_\beta \leqslant \|X\|_\beta^w$. Dividing both sides by $(t-s)^\beta$, we have

$$
\|Y^n\|_{s,t,\beta,n} \leqslant C_1 \sum_{w=1}^N \|X\|_\beta^w (t-s)^{\beta(w-1)} \left[1 + \|Y^n\|_{s,t,\beta,n} (t-s)^\beta \right].
$$

If $n \geqslant 2T(2NC_1\|X\|_\beta)^{1/\beta}$ then there exist $N_0 \in \mathbb{N}_+$ and $N_1 = \frac{N_0 T}{n}$ such that

$$
(2NC_1\|X\|_\beta)^{-1/\beta} \leqslant 2N_1 \leqslant 2(2NC_1\|X\|_\beta)^{-1/\beta}.
$$

When $t - s = N_1$ we get

$$
C_1 \|X\|_\beta^w (t-s)^{\beta w} \leqslant \frac{1}{2N}, \quad \forall\, w = 1, \cdots, N,
$$

which shows

$$\|Y^n\|_{s,t,\beta,n} \leqslant 2C_1 \sum_{w=1}^{N} \|X\|_\beta^w (t-s)^{\beta(w-1)}.$$

When $t - s > N_1$,

$$\left|Y_t^n - Y_s^n\right| \leqslant C\left(\left[\frac{t-s}{N_1}\right] + 1\right) \sup_{r=\lceil r\rceil^n \leqslant t_{n-1}} \|Y^n\|_{r,r+N_1,\beta,n} N_1^\beta.$$

We derive that if $n \geqslant 2T(2NC_1\|X\|_\beta)^{1/\beta}$,

$$\|Y^n\|_{s,t,\beta,n} \leqslant C \max\left\{\|X\|_\beta, \|X\|_\beta^{N-1+1/\beta}\right\}.$$

If $n < 2T(2NC_1\|X\|_\beta)^{1/\beta}$ the definition of Y^n leads to

$$\|Y^n\|_{s,t,\beta,n} \leqslant C\|X\|_\beta^{1/\beta}.$$

Therefore, for $0 \leqslant s < t \leqslant T$,

$$\|Y^n\|_{s,t,\beta} \leqslant C\|X\|_\beta^N + \|Y^n\|_{\lceil s\rceil^n, \lfloor t\rfloor_n,\beta,n} \leqslant C \max\left\{\|X\|_\beta, \|X\|_\beta^{N-1+1/\beta}\right\},$$

which we conclude the estimates

$$\|Y^n\|_\beta \leqslant C(c, \mathbf{s}, V, \beta, T) \max\left\{\|X\|_\beta, \|X\|_\beta^{N-1+1/\beta}\right\},$$

$$\|Y^n\|_\infty \leqslant |y| + C(c, \mathbf{s}, V, \beta, T) \max\left\{\|X\|_\beta, \|X\|_\beta^{N-1+1/\beta}\right\}.$$

Moreover, if $C_0 \in (0, (C_1 N)^{-1})$ then for $0 \leqslant s < t \leqslant T$ such that $\|X\|_\beta |t-s|^\beta \leqslant C_0$, we obtain

$$\|Y^n\|_{s,t,\beta} \leqslant C(N, V, \beta, T, C_0)\|X\|_\beta.$$

□

The continuous dependence of two-stage values $Y^n_{\cdot,1}$ and $Y^n_{\cdot,2}$ introduced by (3.17) and (3.18) follows from Proposition 4.2.

PROPOSITION 4.3 Let $1/2 < \beta < H$. If $V \in \mathscr{C}_b^N(\mathbb{R}^m; \mathbb{R}^{m\times d})$ then for any $n \in \mathbb{N}_+$,

$$\|Y^n_{\cdot,2}\|_\beta \leqslant C(N,V,\beta,T)\max\left\{\|X\|_\beta, \|X\|_\beta^{N-1+1/\beta}\right\},$$

$$\|Y^n_{\cdot,2}\|_\infty \leqslant |y| + C(N,V,\beta,T)\max\left\{\|X\|_\beta, \|X\|_\beta^{N-1+1/\beta}\right\}.$$

Moreover, there exists some $C_0 > 0$ such that $0 \leqslant s < t \leqslant T$ and $\|X\|_\beta |t-s|^\beta \leqslant C_0$ imply

$$\|Y^n_{\cdot,2}\|_{s,t,\beta} \leqslant C(N,V,\beta,T,C_0)\|X\|_\beta.$$

4.3 *Proof of Theorems* 3.2–3.3

Now we are in position for the proof of our main theorems, Theorems 3.2 and 3.3.

Proof of Theorem 3.2. Notice that condition (3.10) deduces the expression of R^{lead} (3.9). Together with Proposition 4.1, Lemmas 2.3 and 2.4 lead to

$$\left\|R_t^{lead} - R_s^{lead}\right\|_{L^p(\Omega)} \leqslant C(t-s)^{\frac{1}{2}} h^{2H-\frac{1}{2}}, \quad \forall\, p \geqslant 1,\ t = \lceil t \rceil^n,\ s = \lceil s \rceil^n. \tag{4.1}$$

Similarly, based on (2.4) and Hu *et al.* (2017, Lemma 6.2), we have

$$\left\|(R_t - R_t^{lead}) - (R_s - R_s^{lead})\right\|_{L^p(\Omega)} \leqslant C(t-s)^{\frac{1}{2}} h^{2H}, \quad \forall\, p \geqslant 1,\ t = \lceil t \rceil^n,\ s = \lceil s \rceil^n.$$

Recall that

$$L_t = \sum_l \int_0^t \int_0^1 \nabla V_l(\theta Y_s + (1-\theta) Y_s^n)(Y_s - Y_s^n)\, \mathrm{d}\theta\, \mathrm{d}X_s^l =: \sum_l \int_0^t S_s^l (Y_s - Y_s^n)\, \mathrm{d}X_s^l.$$

We introduce two linear equations defined through S_s^l. Let matrices Λ^n and Γ^n satisfy the linear equations

$$\Lambda_t^n = I + \sum_l \int_0^t S_s^l \Lambda_s^n \, \mathrm{d}X_s^l,$$

$$\Gamma_t^n = I - \sum_l \int_0^t \Gamma_s^n S_s^l \, \mathrm{d}X_s^l,$$

where $I \in \mathbb{R}^{m\times m}$ denotes the identity matrix. Using the chain rule one can check that $\Lambda^n \Gamma^n = I$. Applying Proposition 4.1, Remark 4.1 and Hu *et al.* (2016a, Lemma 3.1 (*ii*)), we obtain

$$\max\left\{\|\|\Lambda^n\|_\infty\|_{L^p(\Omega)}, \|\|\Lambda^n\|_\beta\|_{L^p(\Omega)}, \|\|\Gamma^n\|_\infty\|_{L^p(\Omega)}, \|\|\Gamma^n\|_\beta\|_{L^p(\Omega)}\right\} \leqslant C, \quad \forall\, p \geqslant 1. \tag{4.2}$$

Since $Y_t - Y_t^n = \Lambda_t^n \int_0^t \Gamma_s^n \,\mathrm{d}R_s$ the Hölder inequality implies that

$$\| \|Y - Y^n\|_\infty \|_{L^p(\Omega)} \leqslant \left\| \|\Lambda^n\|_\infty \left\| \int_0^\cdot \Gamma_s^n \,\mathrm{d}R_s \right\|_\infty \right\|_{L^p(\Omega)}$$

$$\leqslant \| \|\Lambda^n\|_\infty \|_{L^{2p}(\Omega)} \left\| \left\| \int_0^\cdot \Gamma_s^n \,\mathrm{d}R_s \right\|_\infty \right\|_{L^{2p}(\Omega)}, \quad \forall\, p \geqslant 1.$$

Define $f_t^n := n^{2H-\frac{1}{2}} \int_0^t \Gamma_s^n \,\mathrm{d}R_s$. It suffices to prove that $\| \|f^n\|_\infty \|_{L^q(\Omega)} \leqslant C$, for any $q \geqslant 1$.

For $0 \leqslant s < t \leqslant T$ if there exists some $k \in \{1, \cdots, n\}$ such that $0 \leqslant s < t_k < t \leqslant T$, we decompose $\int_s^t \Gamma_u^n \,\mathrm{d}R_u$ into

$$\int_s^t \Gamma_u^n \,\mathrm{d}R_u = \int_s^{\lceil s\rceil^n} \Gamma_u^n \,\mathrm{d}R_u + \int_{\lceil s\rceil^n}^{\lfloor t\rfloor_n} \Gamma_u^n \,\mathrm{d}R_u + \int_{\lfloor t\rfloor_n}^t \Gamma_u^n \,\mathrm{d}R_u$$

$$= \int_s^{\lceil s\rceil^n} \Gamma_u^n \,\mathrm{d}R_u + \int_{\lceil s\rceil^n}^{\lfloor t\rfloor_n} \Gamma_{\lfloor u\rfloor_n}^n \,\mathrm{d}R_u + \int_{\lceil s\rceil^n}^{\lfloor t\rfloor_n} \int_{\lfloor u\rfloor_n}^u d\Gamma_v^n \,\mathrm{d}R_u + \int_{\lfloor t\rfloor_n}^t \Gamma_u^n \,\mathrm{d}R_u.$$

For the first term, combining the definitions of Y_t^n and R_t, we have

$$\int_s^{\lceil s\rceil^n} \Gamma_u^n \,\mathrm{d}R_u = \int_s^{\lceil s\rceil^n} \Gamma_u^n \Big[V(Y_u^n) - \sum_i b_i V\left(Y^n_{\lceil u\rceil^n, i}\right) \Big] \,\mathrm{d}X_u.$$

By the property of Young's integral (see e.g., Lyons *et al.*, 2007, Theorem 1.16 and Remark 1.17), we obtain for $\frac{1}{2} < \beta < H$,

$$\left| \int_s^{\lceil s\rceil^n} \Gamma_u^n \,\mathrm{d}R_u \right| \leqslant C(\beta, T)\|X\|_\beta |\lceil s\rceil^n - s|^\beta$$

$$\left[\left\| \Gamma_\cdot^n \Big[V(Y_\cdot^n) - \sum_i b_i V(Y^n_{\lceil\cdot\rceil^n, i}) \Big] \right\|_\infty + \left\| \Gamma_\cdot^n \Big[V(Y_\cdot^n) - \sum_i b_i V(Y^n_{\lceil\cdot\rceil^n, i}) \Big] \right\|_\beta \right]$$

$$\leqslant C(\beta, V, T)\|X\|_\beta |\lceil s\rceil^n - s|^\beta$$

$$\Big[\sum_i (\|\Gamma_\cdot^n\|_\infty + \|\Gamma_\cdot^n\|_\beta) \|Y_\cdot^n - Y^n_{\lceil\cdot\rceil^n, i}\|_\infty + \sum_i \|\Gamma_\cdot^n\|_\infty \|Y_\cdot^n - Y^n_{\lceil\cdot\rceil^n, i}\|_\beta \Big],$$

where we use the assumptions $V \in \mathscr{C}_b^1$ and $\sum_i b_i = 1$ in the last inequality. It follows from (3.1) to (3.4) that

$$\left| \int_s^{\lceil s\rceil^n} \Gamma_u^n \,\mathrm{d}R_u \right| \leqslant C(\beta, V, T)\|X\|_\beta^2 (\|\Gamma^n\|_\infty + \|\Gamma^n\|_\beta) |\lceil s\rceil^n - s|^\beta h^\beta$$

$$\leqslant C(\beta, V, T)\|X\|_\beta^2 (\|\Gamma^n\|_\infty + \|\Gamma^n\|_\beta) |t - s|^{\frac{1}{2} - 2(H-\beta)} n^{-(2H - \frac{1}{2})}.$$

For the second term, according to (4.1) and (4.2) and Lemma 2.4, we have

$$\left\| n^{2H-\frac{1}{2}} \int_{\lceil s\rceil^n}^{\lfloor t\rfloor_n} \Gamma^n_{\lfloor u\rfloor_n} \, \mathrm{d}R_u \right\|_{L^q(\Omega)} \leqslant C|t-s|^{\frac{1}{2}}.$$

For the third term, combining the definitions of Γ^n_t and R_t, we know that it contains the third-level iterated integrals of X. Then

$$\left\| n^{2H} \int_{\lceil s\rceil^n}^{\lfloor t\rfloor_n} \int_{\lfloor u\rfloor_n}^{u} \mathrm{d}\Gamma^n_v \, \mathrm{d}R_u \right\|_{L^q(\Omega)} \leqslant C|t-s|^{\frac{1}{2}}.$$

For the fourth term, using similar arguments as the estimate for the first term, we get

$$\left| \int_{\lfloor t\rfloor_n}^{t} \Gamma^n_u \, \mathrm{d}R_u \right| \leqslant C(\beta, V, T)\|X\|^2_\beta (\|\Gamma^n\|_\infty + \|\Gamma^n\|_\beta)|t-s|^{\frac{1}{2}-2(H-\beta)} n^{-(2H-\frac{1}{2})}.$$

If $t_k \leqslant s < t \leqslant t_{k+1}$ it holds that

$$\left| \int_{s}^{t} \Gamma^n_u \, \mathrm{d}R_u \right| \leqslant C(\beta, V, T)\|X\|^2_\beta (\|\Gamma^n\|_\infty + \|\Gamma^n\|_\beta)|t-s|^{\frac{1}{2}-2(H-\beta)} n^{-(2H-\frac{1}{2})}.$$

Therefore, for $0 \leqslant s < t \leqslant T$ and $q \geqslant 1$, we obtain

$$\|f^n_t - f^n_s\|_{L^q(\Omega)} \leqslant C\big(|t-s|^{\frac{1}{2}} + |t-s|^{\frac{1}{2}-2(H-\beta)}\big).$$

For $q > 4$ we take β such that $\max\{\frac{1}{2}, H - \frac{1}{2q}\} < \beta < H$ and take $\alpha = \frac{1}{2} - \frac{1}{q}$ such that $\alpha \in (\frac{1}{q}, \frac{1}{2})$. Lemma 2.1 yields that

$$\begin{aligned}
\mathbb{E}\Big[\|f^n\|^q_\infty\Big] &\leqslant T^{\alpha q-1}\mathbb{E}\Big[\|f^n\|^q_{\alpha-\frac{1}{q}}\Big] \\
&\leqslant C \int_0^T \int_0^T \frac{\mathbb{E}\big[|f^n_t - f^n_s|^q\big]}{|t-s|^{1+q\alpha}} \, \mathrm{d}s \, \mathrm{d}t \\
&\leqslant C \int_0^T \int_0^T \frac{|t-s|^{\frac{q}{2}} + |t-s|^{\frac{q}{2}-2q(H-\beta)}}{|t-s|^{1+q\alpha}} \, \mathrm{d}s \, \mathrm{d}t \\
&\leqslant C.
\end{aligned}$$

□

Proof of Theorem 3.3. According to (3.21) we have that R^{lead} for the simplified step-N Euler scheme is the same as that in (3.9) for the Runge–Kutta method in the proof of Theorem 3.2. Then similar techniques yield

$$\|R^{lead}_t - R^{lead}_s\|_{L^p(\Omega)} \leqslant C(t-s)^{\frac{1}{2}} h^{2H-\frac{1}{2}}, \quad \forall\, p \geqslant 1,\ t = \lceil t\rceil^n,\ s = \lceil s\rceil^n$$

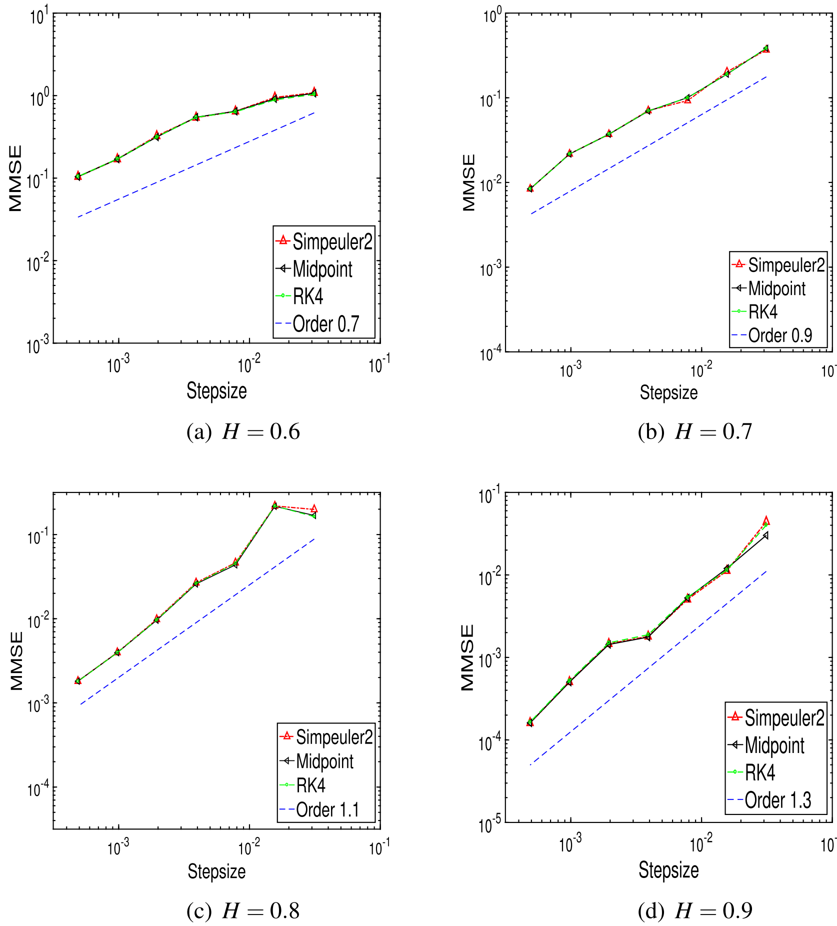

FIG. 1. Maximum mean-square error (MMSE) vs. step size.

and

$$\left\|(R_t - R_t^{lead}) - (R_s - R_s^{lead})\right\|_{L^p(\Omega)} \leqslant C(t-s)^{\frac{1}{2}} h^{2H}, \quad \forall\, p \geqslant 1,\ t = \lceil t \rceil^n,\ s = \lceil s \rceil^n.$$

Repeating the arguments in the proof of Theorem 3.2 we conclude the result. □

5. Numerical experiments

In this section we give an example to verify our results. Consider

$$dY_t = 3\sin(Y_t)dt + 3\cos(Y_t)\mathrm{d}X_t^2 + 3\sin(Y_t)\mathrm{d}X_t^3, \quad t \in (0, 1],$$
$$Y_0 = 5,$$

where X^2 and X^3 are independent fBms with Hurst parameter $H > \frac{1}{2}$. We compare the following three numerical schemes: simplified step-2 Euler scheme and two Runge–Kutta methods with coefficients defined by the Butcher tableaus:

$$\begin{array}{c|c} 1/2 & 1/2 \\ \hline & 1 \end{array},$$

$$\begin{array}{c|cccc} 0 & & & & \\ 1/2 & 1/2 & & & \\ 1/2 & 0 & 1/2 & & \\ 1 & 0 & 0 & 1 & \\ \hline & 1/6 & 2/6 & 2/6 & 1/6 \end{array}$$

In other words the first Runge–Kutta method is the midpoint scheme and the second one is a four-stage Runge–Kutta method satisfying conditions for order 4 in deterministic case. Both of them satisfy condition (3.10). Theorems 3.2 and 3.3 indicate that their maximum mean-square convergence rate is $2H - \frac{1}{2}$, i.e.,

$$\left\| \max_{1\leqslant k\leqslant n} |Y_{t_k} - Y^n_{t_k}| \right\|_{L^2(\Omega)} \leqslant Ch^{2H-\frac{1}{2}},$$

which is illustrated by numerical results in Fig. 1. For each scheme the numerical solution with time step size $h = 2^{-13}$ is taken as the approximated 'exact solution' for comparison and the average of 1000 sample paths is used as an approximation of the expectation.

REMARK 5.1 As mentioned in the introduction the convergence rate $2H - \frac{1}{2}$ is optimal since only increments of fBms are used in the methods under study. This fact is illustrated in Fig. 1 that the four-stage Runge–Kutta method shows the same rate as the other two schemes. It is still an open problem to construct numerical schemes with strong convergence rates higher than $2H - \frac{1}{2}$, in which case efficient simulation of iterated integrals of multi-dimensional fBms should also be taken into consideration to make schemes implementable.

Funding

National Natural Science Foundation of China (11971470, 11871068).

REFERENCES

BAYER, C., FRIZ, P. K., RIEDEL, S. & SCHOENMAKERS, J. (2016) From rough path estimates to multilevel Monte Carlo. *SIAM J. Numer. Anal.*, **54**, 1449–1483.

CAO, Y., HONG, J. & LIU, Z. (2017) Approximating stochastic evolution equations with additive white and rough noises. *SIAM J. Numer. Anal.*, **55**, 1958–1981.

CAO, Y., HONG, J. & LIU, Z. (2018) Finite element approximations for second-order stochastic differential equation driven by fractional Brownian motion. *IMA J. Numer. Anal.*, **38**, 184–197.

CHRONOPOULOU, A. & VIENS, F. G. (2012) Estimation and pricing under long-memory stochastic volatility. *Ann. Finance*, **8**, 379–403.

DAVIE, A. M. (2007) Differential equations driven by rough paths: an approach via discrete approximation. *Appl. Math. Res. Express*, **40**, abm009.

DENK, G. & WINKLER, R. (2007) Modelling and simulation of transient noise in circuit simulation. *Math. Comput. Model. Dyn. Syst.*, **13**, 383–394.

DEYA, A., NEUENKIRCH, A. & TINDEL, S. (2012) A Milstein-type scheme without Lévy area terms for SDEs driven by fractional Brownian motion. *Ann. Inst. Henri Poincaré Probab. Stat.*, **48**, 518–550.

FRIZ, P. K. & RIEDEL, S. (2014) Convergence rates for the full Gaussian rough paths. *Ann. Inst. Henri Poincaré Probab. Stat.*, **50**, 154–194.

FRIZ, P. K. & VICTOIR, N. B. (2010) *Multidimensional Stochastic Processes as Rough Paths: Theory and Applications*. Cambridge Studies in Advanced Mathematics, vol. 120. Cambridge: Cambridge University Press, pp. xiv+656.

GARNIER, J. & SØLNA, K. (2017) Correction to black-Scholes formula due to fractional stochastic volatility. *SIAM J. Financial Math.*, **8**, 560–588.

GRADINARU, M. & NOURDIN, I. (2009) Milstein's type schemes for fractional SDEs. *Ann. Inst. Henri Poincaré Probab. Stat.*, **45**, 1085–1098.

HONG, J., HUANG, C. & WANG, X. (2018) Symplectic Runge-Kutta methods for Hamiltonian systems driven by Gaussian rough paths. *Appl. Numer. Math.*, **129**, 120–136.

HU, Y., LIU, Y. & NUALART, D. (2016a) Rate of convergence and asymptotic error distribution of Euler approximation schemes for fractional diffusions. *Ann. Appl. Probab.*, **26**, 1147–1207.

HU, Y., LIU, Y. & NUALART, D. (2016b) Taylor schemes for rough differential equations and fractional diffusions. *Discrete Contin. Dyn. Syst. Ser. B*, **21**, 3115–3162.

HU, Y., LIU, Y. & NUALART, D. (2017) Crank-Nicolson scheme for stochastic differential equations driven by fractional Brownian motions (in press). arXiv: 1709.01614.

HÜSLER, J., PITERBARG, V. & SELEZNJEV, O. (2003) On convergence of the uniform norms for Gaussian processes and linear approximation problems. *Ann. Appl. Probab.*, **13**, 1615–1653.

KOU, S. C. (2008) Stochastic modeling in nanoscale biophysics: subdiffusion within proteins. *Ann. Appl. Stat.*, **2**, 501–535.

LIU, Y. & TINDEL, S. (2019) First-order Euler scheme for SDEs driven by fractional Brownian motions: the rough case. *Ann. Appl. Probab.*, **29**, 758–826.

LYONS, T. J., CARUANA, M. & LÉVY, T. (2007) *Differential Equations Driven by Rough Paths*. Lecture Notes in Mathematics, vol. 1908. Berlin: Springer, pp. xviii+109. Lectures from the 34th Summer School on Probability Theory held in Saint-Flour, July 6–24, 2004, With an Introduction Concerning the Summer School by Jean Picard.

MANDELBROT, B. B. & VAN NESS, J. W. (1968) Fractional Brownian motions, fractional noises and applications. *SIAM Rev.*, **10**, 422–437.

MILSTEIN, G. N. & TRETYAKOV, M. V. (2004) *Stochastic Numerics for Mathematical Physics*. Scientific Computation. Berlin: Springer, pp. xx+594.

MISHURA, Y. & SHEVCHENKO, G. (2008) The rate of convergence for Euler approximations of solutions of stochastic differential equations driven by fractional Brownian motion. *Stochastics*, **80**, 489–511.

NEUENKIRCH, A. (2008) Optimal pointwise approximation of stochastic differential equations driven by fractional Brownian motion. *Stochastic Process. Appl.*, **118**, 2294–2333.

NEUENKIRCH, A., TINDEL, S. & UNTERBERGER, J. (2010) Discretizing the fractional Lévy area. *Stochastic Process. Appl.*, **120**, 223–254.

NEUENKIRCH, A. & SHALAIKO, T. (2016) The maximum rate of convergence for the approximation of the fractional Lévy area at a single point. *J. Complexity*, **33**, 107–117.

ZÄHLE, M. (1998) Integration with respect to fractal functions and stochastic calculus. I. *Probab. Theory Related Fields*, **111**, 333–374.