ESAIM: M2AN 57 (2023) 1981–2006
https://doi.org/10.1051/m2an/2023015

ESAIM: Mathematical Modelling and Numerical Analysis
www.esaim-m2an.org

STRONG ERROR ANALYSIS OF EULER METHODS FOR OVERDAMPED GENERALIZED LANGEVIN EQUATIONS WITH FRACTIONAL NOISE: NONLINEAR CASE

Xinjie Dai[1,2], Jialin Hong[1,2], Derui Sheng[1,2,*] and Tau Zhou[1,2]

Abstract. This paper considers the strong error analysis of the Euler and fast Euler methods for nonlinear overdamped generalized Langevin equations driven by the fractional noise. The main difficulty lies in handling the interaction between the fractional Brownian motion and the singular kernel, which is overcome by means of the Malliavin calculus and fine estimates of several multiple singular integrals. Consequently, these two methods are proved to be strongly convergent with order nearly $\min\{2(H+\alpha-1),\alpha\}$, where $H\in(1/2,1)$ and $\alpha\in(1-H,1)$ respectively characterize the singularity levels of fractional noises and singular kernels in the underlying equation. This result improves the existing convergence order $H+\alpha-1$ of Euler methods for the nonlinear case, and gives a positive answer to the open problem raised in Fang and Li [*ESAIM Math. Model. Numer. Anal.* **54** (2020) 431–463]. As an application of the theoretical findings, we further investigate the complexity of the multilevel Monte Carlo simulation based on the fast Euler method, which turns out to behave better performance than the standard Monte Carlo simulation when computing the expectation of functionals of the considered equation. Finally, numerical experiments are carried out to support the theoretical results.

Mathematics Subject Classification. 65C20, 65C30, 65C05, 60H07.

Received February 4, 2022. Accepted February 13, 2023.

1. Introduction

The *generalized Langevin equation* (GLE) was originally introduced by Mori [21] and later used extensively to describe the subdiffusion within a single protein molecule [14,15], the motion of microparticles moving randomly in viscoelastic fluids [4,20], and so on. To be specific, the position $x(t)$ of a moving particle with mass m in the energy potential V at time t can be modelled by the GLE

$$m\ddot{x}(t)=-\nabla V(x(t))-\int_0^t K(t-s)\dot{x}(s)\,\mathrm{d}s+\eta(t).$$

Keywords and phrases. Generalized Langevin equation, fractional Brownian motion, fast Euler method, multilevel Monte Carlo simulation, Malliavin calculus.

[1] LSEC, ICMSEC, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, P.R. China.

[2] School of Mathematical Sciences, University of Chinese Academy of Sciences, Beijing 100049, P.R. China.

*Corresponding author: sdr@lsec.cc.ac.cn

Here, dot denotes the derivative on time, and the convolutional kernel $K(t)$ of the friction (dissipation) is related to the random force (fluctuation) $\eta(t)$ through the fluctuation-dissipation theorem

$$\mathbb{E}\big[\eta(t)\eta(s)\big] = k_{\mathrm{B}}T_A K(t-s), \quad \text{for } s \le t,$$

where k_{B} is Boltzmann's constant and T_A is the absolute temperature (see *e.g.*, [16]). To capture the ubiquitous memory phenomena in biology and physics, the fluctuation $\eta(t)$ is often characterized by the fractional noise, and then the fluctuation-dissipation theorem reveals the memory kernel $K(t)$ being proportional to a power law $t^{-\alpha}$ with some $\alpha > 0$ (see *e.g.*, [14, 15]). In the overdamped regime ($m \ll 1$), the GLE with fractional noise reduces to the following fractional *stochastic differential equation* (SDE)

$$D_c^{\alpha} x(t) = b(x(t)) + \sigma \dot{W}_H(t), \tag{1.1}$$

where $b := -\nabla V$, $D_c^{\alpha} x(t) := \frac{1}{\Gamma(1-\alpha)} \int_0^t (t-s)^{-\alpha} \dot{x}(s)\,\mathrm{d}s$ is the Caputo fractional derivative with Γ being the Gamma function, $\sigma > 0$ is the noise intensity, and W_H denotes the *fractional Brownian motion* (fBm) with Hurst index $H \in (1/2, 1)$. Equation (1.1), also known as the overdamped GLE with fractional noise, shall be mathematically interpreted by its integral form (see (2.1) for details). It is a class of *stochastic Volterra integral equations* (SVIEs), and we refer to [17, 18] for more theoretical results on the well-posedness and long-time behavior of the exact solution.

This paper is concerned with discrete-time simulations of equation (1.1), in view of the absence of closed-form solutions. Generally, the memory kernels will result in expensive costs when performing standard time discretizations such as the Euler method. As an appropriate candidate, the fast Euler method shares a satisfactory computational efficiency, which is constructed by combining the Euler method with the sum-of-exponentials approximation (see *e.g.*, [1]). The prerequisite of the strong error analysis of the fast Euler method is to estimate the strong error of the Euler method. Following the arguments in Proposition 3.1 of [5], the strong convergence order $H + \alpha - 1$ of the Euler method is available for equation (1.1) with $\alpha \in (1 - H, 1)$, which exactly coincides with the mean square Hölder continuity exponent of the exact solution. Clearly, one can expect a higher convergence order in terms of equation (1.1) since the driven noise is additive. For the harmonic potential case, which corresponds to the linear external force case, it is firstly proved in [5] that the Euler method applied to equation (1.1) with $\alpha = 2 - 2H$ is strongly convergent with the sharp order nearly $\min\{3 - 3H, 3/2 - H\}$. This convergence order result was later extended by Dai and Xiao [2] for general $\alpha \in (1 - H, 1)$. For the nonlinear external force case, the authors of [5], Page 440 left the improvement of the strong convergence order of the Euler method as an open problem, which is exactly one main goal of the present paper.

The key-point of the strong convergence analysis for the Euler method consists in the upper bound estimate of

$$\mathbb{E}\big[b'\big(\xi_s^{\theta}\big) b'\big(\xi_{\tau}^{\lambda}\big)(G(s) - G(\hat{s}))(G(\tau) - G(\hat{\tau}))\big], \tag{1.2}$$

where $\xi_s^{\theta} := (1-\theta)x(\hat{s}) + \theta x(s)$ with $\hat{s}$ denoting the maximal grid point before s, and $G(\cdot) := \frac{\sigma}{\Gamma(\alpha)} \int_0^{\cdot} (\cdot - s)^{\alpha - 1}\,\mathrm{d}W_H(s)$ is a singular stochastic integral; see Proposition 4.1 for more related notations. In contrast with the cases of SDEs with fBm and SVIEs with standard Brownian motion, the treatment of (1.2) is more difficult due to the interaction between the fBm and singular kernels, even for the linear external force case; see Lemma 1 of [2] for more details. For the nonlinear external force case, we adopt the dual formula in Malliavin calculus to convert stochastic integrals in (1.2) into deterministic ones, whose integrands involve the first and second order Malliavin derivatives of the exact solution. It turns out that the Malliavin derivatives of the exact solution are bounded by some quantities associated to the corresponding singular kernel, rather than by some constant in the case of SDEs (see *e.g.*, [13]). Consequently, more complicated multiple singular integrals need to be handled in our case. By delicately estimating these multiple singular integrals, we attain the strong convergence order nearly $\min\{2(H + \alpha - 1), \alpha\}$ of the Euler method for equation (1.1) with $\alpha \in (1 - H, 1)$. On this basis, we can also read that the strong error of the fast Euler method with tolerance ϵ is bounded by that of the Euler method plus some multiples of ϵ. The new strong convergence order nearly $\min\{2(H+\alpha-1),\alpha\}$ improves the existing convergence order $H+\alpha-1$ of Euler methods for equation (1.1) in the nonlinear case, and particularly reproduces the corresponding result of Theorem 1 from [13] for SDEs with fBm.

An important application of strong error analysis of numerical methods is to analyze the complexity of the *multilevel Monte Carlo* (MLMC) simulation that was originally developed by Giles [6] to approximate the expectation of functionals of SDEs with standard Brownian motion. Compared with the standard Monte Carlo simulation, the MLMC simulation has better performance when computing such quantities, and has been applied successively to different equations with various noises (see [13] for SDEs with fBm, Richard *et al.* [24] for SVIEs with standard Brownian motion, and so on). By computing corrections using multiple levels of grids, the MLMC simulation reduces the variance of the estimator to achieve the fine grid accuracy at a relatively low cost, where the variance has a close relationship with the strong convergence order of the time discretization. Taking into account that the fast Euler method is more efficient than the Euler method, we apply the MLMC simulation based on the fast Euler method to compute the expectation of functionals of equation (1.1). The corresponding complexity analysis is meanwhile investigated; see Theorem 2.3 for more details.

The paper is organized as follows. Section 2 presents the theoretical findings on the Euler method, fast Euler method and MLMC simulation when they are applied to equation (1.1). In order to facilitate the proof of main results, we study the first and second order Malliavin derivatives of the exact solution, and establish the estimates of several multiple singular integrals in Section 3. Section 4 provides the detailed proof of main results. Numerical examples are given in Section 5 to illustrate the theoretical results.

Notations. Denote $a\vee b:=\max\{a,b\}$ and $a\wedge b:=\min\{a,b\}$ for $a,\, b\in\mathbb{R}$. For the integer $m\geq 1$, denote by C_b^m the space of not necessarily bounded real-valued functions that have continuous and bounded derivatives up to order m, and by C_p^m the space of m times continuously differentiable real-valued functions whose derivatives up to order m are of at most polynomial growth. For the integer $l\geq 2$, let $C_{b,p}^{1,l}:=C_b^1\cap C_p^l$. For any $q\geq 1$, $\|\cdot\|_q$ denotes the $L^q(\Omega;\mathbb{R})$-norm, and particularly $\|\cdot\|:=\|\cdot\|_2$. Denote by $\langle\cdot,\cdot\rangle$ the $L^2(\Omega;\mathbb{R})$-inner product. Let $\mathbf{1}_S(\cdot)$ be the indicator function of the set S. Use C as a generic constant and use $C(\cdot)$ if necessary to mention the parameters it depends on, whose values are always independent of the stepsize h and may change when it appears in different places.

2. Main results

Throughout this paper, we restrict ourselves to the case of 1-dimension for the simplicity of notations, and remark that all results could be extended to the multi-dimension case. Then the overdamped GLE (1.1) is mathematically interpreted by

$$x(t)=x_0+\frac{1}{\Gamma(\alpha)}\int_0^t(t-s)^{\alpha-1}b(x(s))\,\mathrm{d}s+\frac{\sigma}{\Gamma(\alpha)}\int_0^t(t-s)^{\alpha-1}\,\mathrm{d}W_H(s) \tag{2.1}$$

for $t\in[0,T]$, where W_H is a 1-dimensional fBm on some complete filtered probability space $(\Omega,\mathscr{F},\{\mathscr{F}_t\}_{t\in[0,T]},\mathbb{P})$. We always assume that $b:\mathbb{R}\to\mathbb{R}$ is a Lipschitz continuous function and the initial value $x_0\in\mathbb{R}$ is deterministic. In this setting, equation (2.1) has a unique strong solution for $H\in(1/2,1)$ and $\alpha\in(1-H,1)$ (see [18], Thm. 1).

In order to solve equation (2.1) numerically, we introduce the Euler method and fast Euler method, where the latter is more computationally efficient. Meanwhile, the corresponding strong error analysis is established.

2.1. Euler method and fast Euler method

For a fixed integer $N\geq 2$, let $\{t_n:=nh,\, n=0,1,\cdots,N\}$ be a uniform partition of $[0,T]$ with the stepsize $h:=T/N$. As introduced in [5], the Euler method for equation (2.1) can be formulated as

$$x_n=x_0+\frac{1}{\Gamma(\alpha)}\sum_{j=1}^n b(x_{j-1})\int_{t_{j-1}}^{t_j}(t_n-s)^{\alpha-1}\,\mathrm{d}s+G(t_n), \tag{2.2}$$

for $n = 1, 2, \cdots, N$, in which

$$G(t) := \frac{\sigma}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} \, \mathrm{d}W_H(s), \qquad \forall\, t \in [0, T]. \tag{2.3}$$

Now we are in a position to present our first main result on the strong convergence order of the Euler method (2.2) for equation (2.1) in Theorem 2.1, whose proof is deferred to Section 4.

Theorem 2.1. *Let $H \in (1/2, 1)$ and $\alpha \in (1-H, 1)$. If $b \in C_{b,p}^{1,3}$, then there exists some positive constant C such that the strong error of the Euler method (2.2) can be controlled as*

$$\sup_{n \le T/h} \|x_n - x(t_n)\| \le C\mathcal{R}_{H,\alpha}(h),$$

where

$$\mathcal{R}_{H,\alpha}(h) := \begin{cases} h^{2(H+\alpha-1)}, & \text{if } \alpha \in (1-H, 2-2H), \\ (|\ln h| \vee \ln T) h^{2-2H}, & \text{if } \alpha = 2-2H, \\ h^{\alpha}, & \text{if } \alpha \in (2-2H, 1). \end{cases}$$

Remark 2.1. For the case $\alpha = 1$, the model (2.1) reduces to the SDE

$$\mathrm{d}x(t) = b(x(t)) \, \mathrm{d}t + \sigma \, \mathrm{d}W_H(t), \qquad t \in [0, T] \tag{2.4}$$

with $x(0) = x_0$. When α tends to 1, Theorem 2.1 reproduces Theorem 1 in [13], which presented that the Euler method for equation (2.4) is of first-order strong convergence under a slightly stronger assumption $b \in C_b^3$.

For the nonlinear model (2.1) with $\alpha \in (1-H, 1)$, the convergence order of the Euler method (2.2) obtained by Fang and Li [5] coincides with the mean square Hölder continuity exponent $H + \alpha - 1$ of the exact solution. The convergence order in Theorem 2.1 is nearly $\min\{2(H+\alpha-1), \alpha\}$ (up to a logarithmic factor for the case $\alpha = 2-2H$). Therefore, we improve the existing convergence rate in the nonlinear case from order $H + \alpha - 1$ to order $2(H + \alpha - 1)$ when $\alpha \in (1-H, 2-2H)$, and to order α when $\alpha \in (2-2H, 1)$. Moreover, the convergence order in Theorem 2.1 is twice of the mean square Hölder continuity exponent of the exact solution when $\alpha \in (1-H, 2-2H)$, which exhibits the same phenomenon as in the case of SDEs driven additively by standard Brownian motion.

Since the overdamped GLE (2.1) is an SVIE with memory, the Euler method (2.2) needs the computational cost of $\mathcal{O}(N^2)$ for a single sample path, which is too expensive in practical calculations. To improve the computational efficiency, Fang and Li [5] proposed the fast Euler method by using the following sum-of-exponentials approximation.

Lemma 2.1 (Sum-of-exponentials approximation [5,12])**.** *For $\alpha \in (0, 1)$, tolerance $\epsilon > 0$ and truncation $\kappa > 0$, there exist positive numbers τ_i and ω_i with $1 \le i \le M_{\exp}$ such that*

$$\left| t^{\alpha-1} - \sum_{i=1}^{M_{\exp}} \omega_i e^{-\tau_i t} \right| \le \epsilon, \qquad \forall t \in [\kappa, T], \tag{2.5}$$

where

$$M_{\exp} = \mathcal{O}\left(\log \frac{1}{\epsilon} \left(\log \log \frac{1}{\epsilon} + \log \frac{T}{\kappa} \right) + \log \frac{1}{\kappa} \left(\log \log \frac{1}{\epsilon} + \log \frac{1}{\kappa} \right) \right).$$

Using the sum-of-exponentials approximation (2.5) with a given tolerance $\epsilon \ll 1$, the Euler method (2.2) can be directly modified to

$$y_n = y_0 + \sum_{j=1}^{n-1} \frac{b(y_{j-1})}{\Gamma(\alpha)} \int_{t_{j-1}}^{t_j} \sum_{i=1}^{M_{\exp}} \omega_i e^{-\tau_i(t_n - s)}\, \mathrm{d}s + \frac{b(y_{n-1})}{\Gamma(\alpha)} \int_{t_{n-1}}^{t_n} (t_n - s)^{\alpha-1}\, \mathrm{d}s + G(t_n).$$

Then, exchanging the summations order obtains the fast Euler method

$$y_n = y_0 + \sum_{i=1}^{M_{\exp}} \omega_i \zeta_i^n + \frac{h^\alpha}{\Gamma(\alpha+1)} b(y_{n-1}) + G(t_n), \tag{2.6}$$

where

$$\zeta_i^n := \begin{cases} 0, & \text{if } n = 1, \\ \frac{1}{\Gamma(\alpha)} \sum_{j=1}^{n-1} b(y_{j-1}) \int_{t_{j-1}}^{t_j} e^{-\tau_i(t_n - s)}\, \mathrm{d}s, & \text{if } n \geq 2 \end{cases}$$

satisfies the recurrence formula

$$\zeta_i^{n+1} = e^{-\tau_i h} \zeta_i^n + \frac{1}{\tau_i \Gamma(\alpha)} \left(e^{-\tau_i h} - e^{-2\tau_i h}\right) b(y_{n-1}).$$

If one takes $\epsilon = h^\alpha$ and $\kappa = h$, then $M_{\exp} = \mathcal{O}((\log N)^2)$, and the above recurrence formula indicates that the fast Euler method (2.6) only needs a computational cost of $\mathcal{O}(N(\log N)^2)$ for a single sample path, so it is more efficient than the Euler method (2.2); see also [5, 12] for more details. For the fast Euler method (2.6), we provide the following strong convergence theorem.

Theorem 2.2. *Under the assumptions of Theorem 2.1, there exists a positive constant C independent of $0 < \epsilon \ll 1$ such that the strong error of the fast Euler method (2.6) can be bounded by*

$$\sup_{n \leq T/h} \|y_n - x(t_n)\| \leq C\mathcal{R}_{H,\alpha}(h) + C\epsilon,$$

where $\mathcal{R}_{H,\alpha}(h)$ is defined in Theorem 2.1.

Proof. Based on the result of Theorem 2.1, the proof can be completed similar to that of Theorem 4.4 from [5]. Thus, the details are omitted. □

For the fast Euler method (2.6) with $\epsilon = \mathcal{R}_{H,\alpha}(h)$, Theorem 2.2 implies that it is strongly convergent with order nearly $\min\{2(H+\alpha-1), \alpha\}$, which improves the existing convergence order $H + \alpha - 1$ in the nonlinear case (see [5], Thm. 4.4).

Remark 2.2. In Theorems 2.1 and 2.2, we impose the global Lipschitz condition for the coefficient b and obtain the strong convergence orders of the Euler method (2.2) and the fast Euler method (2.6). When b satisfies the local Lipschitz condition and linear growth condition, the existence and uniqueness of the solution to the model (2.1) would be obtained by the truncation argument (see [19] for the case $\alpha = 1$), and the convergence of the Euler method (2.2) would be also obtained by the stopping time argument (see [9], Thm. 2.2 for the case $\alpha = 1$). However, the rate of the convergence is unknown. For more general nonlinear function b (*e.g.*, b satisfies the local Lipschitz condition and polynomial growth condition), the existence and uniqueness of the exact solution of the model (2.1) is unknown. And in this setting, the Euler method (2.2) may be divergent for the model (2.1), since it has been shown in Theorem 2.1 of [11] that the Euler–Maruyama method is divergent for the SDE driven by standard Brownian motion (corresponds to the model (2.1) with $\alpha = 1$ and $H = \frac{1}{2}$) with polynomial growth coefficients. It is interesting and challenging to construct convergent and efficient numerical methods for the model (2.1) with non-globally Lipschitz coefficients, and we leave it as a future work.

2.2. Multilevel Monte Carlo simulation

We are also interested in the calculation of $\mathbb{E}[f(x(T))]$ for some Lipschitz continuous function $f:\ \mathbb{R}\to\mathbb{R}$, because this quantity receives a lot of attention in applications (see *e.g.*, [7]). With Theorem 2.2 in hand, one can construct the MLMC simulation combined with the fast Euler method (2.6) to compute $\mathbb{E}[f(x(T))]$, where the MLMC simulation was originally proposed by Giles [6] to improve the efficiency of Monte Carlo simulations.

Without loss of generality, we assume that $T=1$ in this subsection. For a fixed integer $M\geq 2$, define different stepsizes $h_l=M^{-l}$ $(l=0,1,\cdots,L)$. For convenience, denote

$$Q=f(x(1)),\qquad P_l=f\left(y^{h_l}_{M^l}\right),\quad l=0,1,\cdots,L, \tag{2.7}$$

where $y^{h_l}_{M^l}$ is the approximation of $x(1)$ by using the fast Euler method (2.6) with stepsize h_l and tolerance h_l^{α}. Based on the trivial identity

$$\mathbb{E}[P_L]=\mathbb{E}[P_0]+\sum_{l=1}^{L}\mathbb{E}[P_l-P_{l-1}],$$

the MLMC estimator can be formulated as

$$Z:=\frac{1}{N_0}\sum_{n=1}^{N_0}P_0^{(n)}+\sum_{l=1}^{L}\frac{1}{N_l}\sum_{n=1}^{N_l}\left[P_l^{(n)}-P_{l-1}^{(n)}\right], \tag{2.8}$$

where $\frac{1}{N_0}\sum_{n=1}^{N_0}P_0^{(n)}$ with N_0 i.i.d. copies $\{P_0^{(n)}\}_{n=1}^{N_0}$ of P_0 is used to estimate $\mathbb{E}[P_0]$, $\frac{1}{N_l}\sum_{n=1}^{N_l}\left[P_l^{(n)}-P_{l-1}^{(n)}\right]$ with N_l i.i.d. copies $\left\{P_l^{(n)},P_{l-1}^{(n)}\right\}_{n=1}^{N_l}$ of $\{P_l,P_{l-1}\}$ is used to estimate $\mathbb{E}[P_l-P_{l-1}]$ for $l=1,2,\cdots,L$, and the positive integers L and N_l $(l=0,1,\cdots,L)$ are related to the following complexity theorem.

Theorem 2.3. *Let $H\in(1/2,1)$ and $\alpha=2-2H$. Under the assumptions of Theorem 2.1, there exist suitable positive integers L and N_l $(l=0,1,\cdots,L)$ such that the mean square error of the MLMC estimator (2.8) can be controlled by a specified accuracy ε $(0<\varepsilon\ll 1)$, i.e.,*

$$\|Z-\mathbb{E}[f(x(1))]\|\leq\varepsilon$$

with the computational cost C_{MLMC} satisfying

$$C_{\mathrm{MLMC}}\leq\begin{cases}C\varepsilon^{-2}, & \text{if } H\in\left(\frac{1}{2},\frac{3}{4}\right),\\ C\varepsilon^{-2}|\ln\varepsilon|^{6}, & \text{if } H=\frac{3}{4},\\ C\varepsilon^{-\left(2+\frac{4H-3}{2-2H-\rho}\right)}|\ln\varepsilon|^{4}, & \text{if } H\in\left(\frac{3}{4},1\right),\end{cases}$$

where $\rho\in(0,1-H)$ can be arbitrarily small, and the positive constant C is independent of ε.

Proof. The variance formula $\mathrm{Var}(X)=\|X\|^2-|\mathbb{E}[X]|^2$, for $X\in L^2(\Omega;\mathbb{R})$, yields

$$\begin{aligned}\|Z-\mathbb{E}[f(x(1))]\|^2&=|\mathbb{E}[Z-\mathbb{E}[f(x(1))]]|^2+\mathrm{Var}(Z)\\&=|\mathbb{E}[P_L-Q]|^2+\frac{1}{N_0}\mathrm{Var}(P_0)+\sum_{l=1}^{L}\frac{1}{N_l}\mathrm{Var}(P_l-P_{l-1})=:\mathrm{I}+\mathrm{II}+\mathrm{III},\end{aligned} \tag{2.9}$$

where Q and P_l are defined by (2.7). Firstly, Theorem 2.2 and the Lipschitz continuity of f show that

$$\mathrm{I}:=|\mathbb{E}[P_L-Q]|^2\leq\|P_L-Q\|^2=\left\|f\left(y^{h_L}_{M^L}\right)-f(x(1))\right\|^2$$

$$\leq C\left\|y^{h_L}_{M^L} - x(1)\right\|^2 \leq Ch_L^{4-4H}|\ln h_L|^2 \leq C_1 h_L^{4-4H-2\rho} \leq \frac{\varepsilon^2}{3} \tag{2.10}$$

provided $\rho \in (0, 1-H)$ and

$$L = \left\lceil \frac{1}{4-4H-2\rho} \log_M\left(3C_1\varepsilon^{-2}\right)\right\rceil, \tag{2.11}$$

where the notation $\lceil\cdot\rceil$ denotes the ceiling function. Secondly,

$$\mathrm{II} := \frac{1}{N_0}\mathrm{Var}(P_0) \leq \frac{\|P_0\|^2}{N_0} \leq \frac{C_2}{N_0} \leq \frac{\varepsilon^2}{3} \tag{2.12}$$

provided

$$N_0 = \lceil 3C_2\varepsilon^{-2}\rceil. \tag{2.13}$$

Thirdly, for $l = 1, 2, \cdots, L$, similar to the estimation of (2.10), one can obtain

$$\mathrm{Var}(P_l - Q) \leq \|P_l - Q\|^2 \leq Ch_l^{4-4H}|\ln h_l|^2 \quad \text{and} \quad \mathrm{Var}(Q - P_{l-1}) \leq Ch_l^{4-4H}|\ln h_l|^2,$$

so

$$\mathrm{Var}(P_l - P_{l-1}) = \mathrm{Var}(P_l - Q + Q - P_{l-1}) \leq \left(\sqrt{\mathrm{Var}(P_l - Q)} + \sqrt{\mathrm{Var}(Q - P_{l-1})}\right)^2 \leq C_3 h_l^{4-4H}|\ln h_l|^2.$$

Thus,

$$\mathrm{III} := \sum_{l=1}^{L} \frac{1}{N_l}\mathrm{Var}(P_l - P_{l-1}) \leq \sum_{l=1}^{L} \frac{C_3}{N_l} h_l^{4-4H}|\ln h_l|^2 \leq \frac{\varepsilon^2}{3} \tag{2.14}$$

provided

$$N_l = \left\lceil 3C_3\varepsilon^{-2} h_l^{\frac{5-4H}{2}} |\ln h_l|^2 \sum_{l'=1}^{L} h_{l'}^{\frac{3-4H}{2}}\right\rceil, \quad \text{for } l = 1, 2, \cdots, L. \tag{2.15}$$

Combining (2.9), (2.10), (2.12) and (2.14), one can take L and N_l $(l = 0, 1, \cdots, L)$ satisfying (2.11), (2.13) and (2.15) such that $\|Z - \mathbb{E}[f(x(1))]\| \leq \varepsilon$ holds.

It remains to analyze the computational cost C_{MLMC} of the MLMC estimator (2.8). Actually, the fact that the fast Euler method (2.6) sampling a path needs a computational cost of $\mathcal{O}\big(|\ln h_l|^2 h_l^{-1}\big)$ (see [5], Thm. 4.4) indicates

$$\begin{aligned}
C_{\mathrm{MLMC}} &\leq CN_0 + C\sum_{l=1}^{L} N_l |\ln h_l|^2 h_l^{-1} \\
&\leq C\big(3C_2\varepsilon^{-2} + 1\big) + \sum_{l=1}^{L}\left(3C_3\varepsilon^{-2} h_l^{\frac{5-4H}{2}} |\ln h_l|^2 \sum_{l'=1}^{L} h_{l'}^{\frac{3-4H}{2}} + 1\right) C|\ln h_l|^2 h_l^{-1} \\
&\leq C\varepsilon^{-2} + C\varepsilon^{-2}\sum_{l=1}^{L} |\ln h_l|^4 h_l^{\frac{3-4H}{2}} \sum_{l'=1}^{L} h_{l'}^{\frac{3-4H}{2}} \\
&\leq \begin{cases} C\varepsilon^{-2}, & \text{if } H \in \left(\frac{1}{2}, \frac{3}{4}\right), \\ C\varepsilon^{-2}|\ln\varepsilon|^6, & \text{if } H = \frac{3}{4}, \\ C\varepsilon^{-\left(2+\frac{4H-3}{2-2H-\rho}\right)}|\ln\varepsilon|^4, & \text{if } H \in \left(\frac{3}{4}, 1\right), \end{cases}
\end{aligned}$$

where the last step used the subsequent Lemma 2.2 and the following relations

$$L = \log_M \varepsilon^{-\frac{1}{2-2H-\rho}} + C \le C|\ln\varepsilon|, \quad h_L = M^{-L} = C\varepsilon^{\frac{1}{2-2H-\rho}} \quad \text{and} \quad |\ln h_L| \le C|\ln\varepsilon|.$$

The proof is completed. □

Lemma 2.2. *Let $M \ge 2$, $h_l = M^{-l}$ $(l = 1, 2, \cdots, L)$, $\beta \ge 0$ and $\gamma \in \mathbb{R}$. Then, there exists some positive constant C (only depends on M, β, γ, but not on L) such that*

$$\sum_{l=1}^{L} |\ln h_l|^\beta h_l^\gamma \le \begin{cases} C, & \text{if } \gamma > 0, \\ |\ln h_L|^\beta L, & \text{if } \gamma = 0, \\ C|\ln h_L|^\beta h_L^\gamma, & \text{if } \gamma < 0. \end{cases}$$

Proof. When $\gamma > 0$, it follows from $M \ge 2 > 1$ that

$$\sum_{l=1}^{L} |\ln h_l|^\beta h_l^\gamma = \sum_{l=1}^{L} \big(l \ln M\big)^\beta M^{-\gamma l} \le C \sum_{l=1}^{+\infty} l^\beta M^{-\gamma l} \le C.$$

Note that the result is trivial for the case $\gamma = 0$. While $\gamma < 0$, the fact that $|\ln h_l|$ is increasing with respect to l shows

$$\sum_{l=1}^{L} |\ln h_l|^\beta h_l^\gamma \le |\ln h_L|^\beta \sum_{l=1}^{L} M^{(L-l)\gamma} M^{-L\gamma} \le |\ln h_L|^\beta h_L^\gamma \sum_{l=1}^{+\infty} M^{\gamma l} \le C|\ln h_L|^\beta h_L^\gamma,$$

which completes the proof. □

3. Preliminaries for the proof of Theorem 2.1

To facilitate the proof of Theorem 2.1, this section is devoted to establishing the Malliavin differentiability with respect to the fBm of the exact solution and the sharp estimates of some singular integrals. In the following, the Beta function will be frequently used, which is defined by

$$B(a,b) := \int_0^1 u^{a-1}(1-u)^{b-1}\,\mathrm{d}u, \qquad \text{for } a, b > 0. \tag{3.1}$$

3.1. Malliavin calculus with respect to the fBm

Let us start with some basic definitions and Malliavin calculus with respect to the fBm; see [10, 22, 23] for more details. Throughout this paper, we always assume $H \in (1/2, 1)$, and in this case, the covariance of the fBm $\{W_H(t)\}_{t\in[0,T]}$ possesses the following form

$$\mathrm{Cov}(t,s) := \mathbb{E}\big[W_H(t)W_H(s)\big] = H(2H-1)\int_0^t \int_0^s |u-v|^{2H-2}\,\mathrm{d}u\,\mathrm{d}v, \quad \forall s, t \in [0,T].$$

Denote by $\mathcal{E}$ the set of real-valued step functions on $[0,T]$ and let $\mathcal{H}$ be the Hilbert space defined as the closure of $\mathcal{E}$ with respect to the inner product $\langle \mathbf{1}_{[0,t]}(\cdot), \mathbf{1}_{[0,s]}(\cdot)\rangle_{\mathcal{H}} := \mathrm{Cov}(t,s)$. The map $\mathbf{1}_{[0,t]}(\cdot) \mapsto W_H(t)$ can be extended to an isometry between $\mathcal{H}$ and a closed subspace of $L^2(\Omega, \mathcal{F}, \mathbb{P})$. More precisely, denote this isometry by $\varphi \mapsto W_H(\varphi)$. For any $\varphi, \phi \in \{\psi \in L^1_{loc}([0,T]) : \int_0^T \int_0^T |\psi(u)||\psi(v)||u-v|^{2H-2}\,\mathrm{d}u\,\mathrm{d}v < \infty\}$,

$$\langle \varphi, \phi\rangle_{\mathcal{H}} = H(2H-1)\int_0^T \int_0^T \varphi(u)\phi(v)|u-v|^{2H-2}\,\mathrm{d}u\,\mathrm{d}v = \mathbb{E}\Big[W_H(\varphi)W_H(\phi)\Big].$$

Denote by $\mathcal{S}$ the class of smooth real-valued random variables such that $F \in \mathcal{S}$ has the form

$$F = f(W_H(\varphi_1), \dots, W_H(\varphi_n)),$$

where $f \in C_p^\infty(\mathbb{R}^n;\mathbb{R})$, $\varphi_i \in \mathcal{H}$, $i = 1,\dots,n$, $n \in \mathbb{N}_+$. Here, $C_p^\infty(\mathbb{R}^n;\mathbb{R})$ is the space of all real-valued smooth functions on $\mathbb{R}^n$ with polynomial growth. The Malliavin derivative with respect to the fBm of $F \in \mathcal{S}$ is an $\mathcal{H}$-valued random variable defined by

$$DF = \sum_{i=1}^{n} \frac{\partial}{\partial x_i} f(W_H(\varphi_1), \dots, W_H(\varphi_n))\varphi_i,$$

which is also a stochastic process $DF = \{D_r F\}_{r\in[0,T]}$ with

$$D_r F = \sum_{i=1}^{n} \frac{\partial}{\partial x_i} f(W_H(\varphi_1), \dots, W_H(\varphi_n))\varphi_i(r).$$

For any $q \geq 1$, we denote the domain of D in $L^q(\Omega;\mathbb{R})$ by $\mathbb{D}^{1,q}$, meaning that $\mathbb{D}^{1,q}$ is the closure of $\mathcal{S}$ with respect to the norm

$$\|F\|_{\mathbb{D}^{1,q}} = (\mathbb{E}[|F|^q] + \mathbb{E}[\|DF\|^q_{\mathcal{H}}])^{\frac{1}{q}}.$$

In a similar manner, for $F \in \mathcal{S}$, the iterated derivative $D^k F$ ($k \in \mathbb{N}_+$) is defined as a random variable with values in $\mathcal{H}^{\otimes k}$. For every $q \geq 1$ and $k \in \mathbb{N}_+$, denote by $\mathbb{D}^{k,q}$ the completion of $\mathcal{S}$ with respect to the norm

$$\|F\|_{\mathbb{D}^{k,q}} = \left(\mathbb{E}\left[|F|^q + \sum_{j=1}^{k} \|D^j F\|^q_{\mathcal{H}^{\otimes j}}\right]\right)^{\frac{1}{q}}.$$

We also denote $\mathbb{D}^{k,\infty} := \bigcap_{q\in[1,\infty)} \mathbb{D}^{k,q}$ for simplicity. For $F \in \mathbb{D}^{2,\infty}$ and $\phi \in C_p^2$, it follows from Remark 3.4 of [25] that $\phi(F) \in \mathbb{D}^{2,\infty}$, $D(\phi(F)) = \phi'(F)DF$ and $D^2(\phi(F)) = \phi''(F)DF \otimes DF + \phi'(F)D^2F$.

Next, we introduce the adjoint operator δ of the derivative operator D, which is also known as the Skorohod integral. If an $\mathcal{H}$-valued random variable $\varphi \in L^2(\Omega;\mathcal{H})$ satisfies

$$\left|\mathbb{E}[\langle \varphi, DF\rangle_{\mathcal{H}}]\right| \leq C(\varphi)\|F\|, \quad \forall F \in \mathbb{D}^{1,2},$$

then $\varphi \in \mathrm{Dom}(\delta)$ and $\delta(\varphi) \in L^2(\Omega;\mathbb{R})$ is characterized by the dual formula

$$\mathbb{E}[\langle \varphi, DF\rangle_{\mathcal{H}}] = \mathbb{E}[F\delta(\varphi)], \quad \forall F \in \mathbb{D}^{1,2}.$$

In particular, when $\varphi \in \mathcal{H}$ is deterministic, the Skorohod integral $\delta(\varphi)$ coincides with the Riemann–Stieltjes integral $\int_0^T \varphi(u)\,\mathrm{d}W_H(u)$.

3.2. Malliavin differentiability of the exact solution

In this part, we consider the first and second order Malliavin derivatives of the exact solution in Theorem 3.1, which turns out to be bounded by some quantities related to the singular kernel in the model (2.1), rather than by a constant, compared to the case of SDEs with fBm (see *e.g.*, [13]). The following extension of Grönwall's lemma is established in Lemma 15 of [3].

Lemma 3.1. *Let $\{f_n\}_{n\in\mathbb{N}}$ be a sequence of nonnegative real-valued functions on $[0,T]$ and k_1, k_2 be nonnegative real numbers such that for all $n \geq 1$, $t \in [0,T]$ and some $g \in L^1([0,T];\mathbb{R}_+)$,*

$$f_n(t) \leq k_1 + \int_0^t \big(k_2 + f_{n-1}(s)\big)g(t-s)\,\mathrm{d}s.$$

Then, there is a sequence $\{a_n\}_{n\in\mathbb{N}_+}$ *of nonnegative real numbers satisfying* $\sum_{n=1}^{\infty} a_n < \infty$ *with the property: If* $\sup_{t\in[0,T]} f_0(t) = M$*, then for all* $n \geq 1$*,* $t \in [0,T]$*,*

$$f_n(t) \leq k_1 + (k_1 + k_2)\sum_{i=1}^{n-1} a_i + (k_2 + M)a_n,$$

which implies $\sup_{n\geq 0}\sup_{t\in[0,T]} f_n(t) < \infty$*. Moreover, if further* $k_1 = k_2 = 0$*, then* $\sum_{n\geq 0} f_n(t)$ *converges uniformly on* $[0,T]$*.*

Theorem 3.1. *Let* $H \in (1/2, 1)$ *and* $\alpha \in (1-H, 1)$*. If* $b \in C_{b,p}^{1,2}$*, then* $x(t) \in \mathbb{D}^{2,\infty}$ *for any* $t \in [0,T]$*, and there exists some constant* $C = C(\alpha, H, \sigma, T)$ *such that for* $r \in [0,T]$*,*

$$|D_r x(t)| \leq C(t-r)^{\alpha-1}\mathbf{1}_{[0,t)}(r), \quad \text{a.s.} \tag{3.2}$$

Moreover, for any $q \geq 1$*, there exists some constant* $C = C(\alpha, H, \sigma, T, q)$ *such that for* $r_1, r_2 \in [0,T]$*,*

$$\|D_{r_2}D_{r_1}x(t)\|_q \leq C\int_{r_1\vee r_2}^{t} (t-s)^{\alpha-1}(s-r_1)^{\alpha-1}(s-r_2)^{\alpha-1}\,\mathrm{d}s\mathbf{1}_{[0,t)}(r_1\vee r_2). \tag{3.3}$$

Proof. Consider the following standard Picard iteration sequence

$$x^{(n+1)}(t) = x_0 + \frac{1}{\Gamma(\alpha)}\int_0^t (t-s)^{\alpha-1}b\big(x^{(n)}(s)\big)\,\mathrm{d}s + G(t), \quad n \geq 0 \tag{3.4}$$

with $x^{(0)}(t) = x_0$ for all $t \in [0,T]$. Here, G is defined by (2.3). To prove the statement $x(t) \in \mathbb{D}^{2,\infty}$ for all $t \in [0,T]$, it suffices to show that $\{x^{(n)}(t)\}_{n=0}^{\infty}$ converges to $x(t)$ in $L^q(\Omega;\mathbb{R})$ and $\sup_{n\geq 0}\|x^{(n)}(t)\|_{\mathbb{D}^{2,q}} < \infty$ for any $q \geq 1$, in view of Lemma 1.5.3 from [22].

Claim 1. $\{x^{(n)}(t)\}_{n=0}^{\infty}$ *converges to* $x(t)$ *in* $L^q(\Omega;\mathbb{R})$ *for any* $q \geq 1$*.*

Let $q \geq 1$. Since the law of $G(t)$ is Gaussian, it follows from Lemma 2 of [18] that for any $t \in [0,T]$,

$$\|G(t)\|_q \leq C(q)\|G(t)\| \leq C\big\|(t-\cdot)^{\alpha-1}\mathbf{1}_{[0,t)}(\cdot)\big\|_{\mathcal{H}} \leq C(\alpha, H, T, \sigma, q). \tag{3.5}$$

Then, by (3.4) and (3.5),

$$\Big\|x^{(1)}(t) - x^{(0)}(t)\Big\|_q = \Big\|b(x_0)\frac{1}{\Gamma(\alpha+1)}t^{\alpha} + G(t)\Big\|_q \leq C(\alpha, H, T, \sigma, q, x_0).$$

For each $n \geq 1$, the assumption $b \in C_b^1$ reads

$$\Big\|x^{(n+1)}(t) - x^{(n)}(t)\Big\|_q \leq C\int_0^t (t-s)^{\alpha-1}\Big\|x^{(n)}(s) - x^{(n-1)}(s)\Big\|_q\,\mathrm{d}s.$$

Applying Lemma 3.1 to the sequence $\{f_n\}_{n\geq 0}$ with $f_n(t) := \|x^{(n+1)}(t) - x^{(n)}(t)\|_q$ shows

$$\sum_{n=0}^{\infty}\sup_{t\in[0,T]}\Big\|x^{(n+1)}(t) - x^{(n)}(t)\Big\|_q < \infty,$$

which implies that $x^{(n)}(t) = x^{(0)}(t) + \sum_{k=0}^{n-1}\big(x^{(k+1)}(t) - x^{(k)}(t)\big)$ converges uniformly in $L^q(\Omega;\mathbb{R})$ with respect to $t \in [0,T]$, and

$$\sup_{t\in[0,T]}\Big\|x^{(n)}(t)\Big\|_q < \infty. \tag{3.6}$$

Putting $n \to \infty$ on both sides of (3.4), the limit of $\{x^{(n)}(t)\}_{n\geq 0}$ satisfies equation (2.1), and

$$\sup_{t\in[0,T]}\|x(t)\|_q < \infty. \tag{3.7}$$

Claim 2. $\sup_{n\geq 0}\big\|x^{(n)}(t)\big\|_{\mathbb{D}^{2,q}} < \infty$ *for any* $q \geq 1$.

Obviously, $x^{(0)}(t) \in \mathbb{D}^{2,\infty}$. Assume by induction that $x^{(n)}(t) \in \mathbb{D}^{2,\infty}$. Then, it follows from $b \in C^{1,2}_{b,p}$ that $x^{(n+1)}(t) \in \mathbb{D}^{2,\infty}$. Thus, we have $x^{(n)}(t) \in \mathbb{D}^{2,\infty}$, for any integer $n \geq 1$. Taking the Malliavin derivative on both sides of (3.4), the chain rule of the Malliavin derivative gives that for $r < t$,

$$D_r x^{(n+1)}(t) = \frac{1}{\Gamma(\alpha)}\int_0^t (t-s)^{\alpha-1} b'(x^{(n)}(s)) D_r x^{(n)}(s)\mathbf{1}_{[0,s)}(r)\,\mathrm{d}s + \frac{\sigma}{\Gamma(\alpha)}(t-r)^{\alpha-1}, \tag{3.8}$$

and for $r \in [t,T]$, $D_r x^{(n+1)}(t) = 0$. Then, using (3.5), (3.8) and the assumption $b \in C^1_b$, we have

$$\begin{aligned}\Big\|Dx^{(n+1)}(t)\Big\|_{L^q(\Omega;\mathcal{H})} &\leq C\int_0^t (t-s)^{\alpha-1}\Big\|Dx^{(n)}(s)\Big\|_{L^q(\Omega;\mathcal{H})}\,\mathrm{d}s + C\big\|(t-\cdot)^{\alpha-1}\mathbf{1}_{[0,t)}(\cdot)\big\|_{\mathcal{H}} \\ &\leq C\int_0^t (t-s)^{\alpha-1}\Big\|Dx^{(n)}(s)\Big\|_{L^q(\Omega;\mathcal{H})}\,\mathrm{d}s + C(\alpha,H,T,\sigma).\end{aligned}$$

Applying Lemma 3.1 to the sequence $\{f_n\}_{n\geq 0}$ with $f_n(t) := \|Dx^{(n)}(t)\|_{L^q(\Omega;\mathcal{H})}$ obtains

$$\sup_{n\geq 0}\sup_{t\in[0,T]}\Big\|Dx^{(n)}(t)\Big\|_{L^q(\Omega;\mathcal{H})} \leq C(\alpha,H,T,\sigma), \quad \forall q \geq 1. \tag{3.9}$$

Taking the Malliavin derivative on both sides of (3.8), we also have for r_1, $r_2 \in [0,T]$,

$$\begin{aligned}D_{r_2}D_{r_1}x^{(n+1)}(t) &= \frac{1}{\Gamma(\alpha)}\int_0^t (t-s)^{\alpha-1} b''(x^{(n)}(s)) D_{r_2}x^{(n)}(s) D_{r_1}x^{(n)}(s)\mathbf{1}_{[0,s)}(r_1\vee r_2)\,\mathrm{d}s \\ &\quad + \frac{1}{\Gamma(\alpha)}\int_0^t (t-s)^{\alpha-1} b'(x^{(n)}(s)) D_{r_2}D_{r_1}x^{(n)}(s)\mathbf{1}_{[0,s)}(r_1\vee r_2)\,\mathrm{d}s.\end{aligned}$$

Then, using the assumption $b \in C^{1,2}_{b,p}$, Hölder's inequality as well as (3.6) and (3.9), one can derive

$$\begin{aligned}\Big\|D^2x^{(n+1)}(t)\Big\|_{L^q(\Omega;\mathcal{H}^{\otimes 2})} &\leq C\int_0^t (t-s)^{\alpha-1}\Big\|b''(x^{(n)}(s))\Big\|_{2q}\Big\|Dx^{(n)}(s)\Big\|^2_{L^{4q}(\Omega;\mathcal{H})}\,\mathrm{d}s \\ &\quad + C\int_0^t (t-s)^{\alpha-1}\Big\|D^2x^{(n)}(s)\Big\|_{L^q(\Omega;\mathcal{H}^{\otimes 2})}\,\mathrm{d}s \\ &\leq C + C\int_0^t (t-s)^{\alpha-1}\Big\|D^2x^{(n)}(s)\Big\|_{L^q(\Omega;\mathcal{H}^{\otimes 2})}\,\mathrm{d}s.\end{aligned}$$

By means of Lemma 3.1, a similar argument as in the proof of (3.9) will yield

$$\sup_{n\geq 0}\sup_{t\in[0,T]}\Big\|D^2x^{(n)}(t)\Big\|_{L^q(\Omega;\mathcal{H}^{\otimes 2})} \leq C(\alpha,H,T,\sigma),$$

which together with (3.9) proves *Claim 2*. Therefore, $x(t) \in \mathbb{D}^{2,\infty}$ for all $t \in [0,T]$.

It remains to prove the estimates (3.2) and (3.3). By the chain rule of the Malliavin derivative, we have

$$D_r x(t) = \left(\frac{1}{\Gamma(\alpha)}\int_r^t (t-s)^{\alpha-1} b'(x(s)) D_r x(s)\,\mathrm{d}s + \frac{\sigma}{\Gamma(\alpha)}(t-r)^{\alpha-1}\right)\mathbf{1}_{[0,t)}(r), \quad r \in [0,T]. \tag{3.10}$$

Therefore,

$$|D_r x(t)| \leq C\left(\int_r^t (t-s)^{\alpha-1}|D_r x(s)|\,\mathrm{d}s + (t-r)^{\alpha-1}\right)\mathbf{1}_{[0,t)}(r),$$

from which one sees that (3.2) is a direct consequence of the singular Grönwall inequality (see *e.g.*, [26], Thm. 1). Now we turn to proving (3.3). Taking the Malliavin derivative on both sides of (3.10), and taking the estimates (3.2) and (3.7) into account, one obtains that for r_1, $r_2 \in [0,T]$, $q \geq 1$,

$$\begin{aligned}&\|D_{r_2}D_{r_1}x(t)\|_q\\&\quad\leq C\bigg(\int_{r_1\vee r_2}^t (t-s)^{\alpha-1}(s-r_1)^{\alpha-1}(s-r_2)^{\alpha-1}\,\mathrm{d}s+\int_{r_1\vee r_2}^t (t-s)^{\alpha-1}\|D_{r_2}D_{r_1}x(s)\|_q\,\mathrm{d}s\bigg)\mathbf{1}_{[0,t)}(r_1\vee r_2).\end{aligned}$$

Then by the singular Grönwall inequality (see *e.g.*, [26], Thm. 1), one gets

$$\begin{aligned}\|D_{r_2}D_{r_1}x(t)\|_q &\leq C\mathbf{1}_{[0,t)}(r_1\vee r_2)\bigg(\int_{r_1\vee r_2}^t (t-s)^{\alpha-1}(s-r_1)^{\alpha-1}(s-r_2)^{\alpha-1}\,\mathrm{d}s\\&\quad+\int_{r_1\vee r_2}^t\sum_{n=1}^{\infty}\frac{(C\Gamma(\alpha))^n}{\Gamma(n\alpha)}(t-s)^{n\alpha-1}\int_{r_1\vee r_2}^s (s-u)^{\alpha-1}(u-r_1)^{\alpha-1}(u-r_2)^{\alpha-1}\,\mathrm{d}u\,\mathrm{d}s\bigg).\end{aligned}$$

Notice that by the Fubini theorem and the relation $B(n\alpha,\alpha)=\frac{\Gamma(n\alpha)\Gamma(\alpha)}{\Gamma(n\alpha+\alpha)}$,

$$\begin{aligned}&\int_{r_1\vee r_2}^t\sum_{n=1}^{\infty}\frac{(C\Gamma(\alpha))^n}{\Gamma(n\alpha)}(t-s)^{n\alpha-1}\int_{r_1\vee r_2}^s (s-u)^{\alpha-1}(u-r_1)^{\alpha-1}(u-r_2)^{\alpha-1}\,\mathrm{d}u\,\mathrm{d}s\\&\quad=\int_{r_1\vee r_2}^t\sum_{n=1}^{\infty}\frac{(C\Gamma(\alpha))^n}{\Gamma(n\alpha)}\bigg(\int_u^t (t-s)^{n\alpha-1}(s-u)^{\alpha-1}\,\mathrm{d}s\bigg)(u-r_1)^{\alpha-1}(u-r_2)^{\alpha-1}\,\mathrm{d}u\\&\quad=\Gamma(\alpha)\int_{r_1\vee r_2}^t\sum_{n=1}^{\infty}\frac{(C\Gamma(\alpha)(t-u)^{\alpha})^n}{\Gamma(n\alpha+\alpha)}(t-u)^{\alpha-1}(u-r_1)^{\alpha-1}(u-r_2)^{\alpha-1}\,\mathrm{d}u\\&\quad\leq CE_{\alpha,\alpha}(C\Gamma(\alpha)T^{\alpha})\int_{r_1\vee r_2}^t (t-u)^{\alpha-1}(u-r_1)^{\alpha-1}(u-r_2)^{\alpha-1}\,\mathrm{d}u,\end{aligned}$$

where $E_{\alpha,\alpha}(z)=\sum_{n=0}^{\infty}\frac{z^n}{\Gamma(\alpha n+\alpha)}$, $z\in\mathbb{R}$ is the Mittag–Leffler function, and the exchange of the integrals is justified since the integrand is nonnegative. Gathering the above estimates together proves (3.3). The proof is completed. □

3.3. Estimates of singular integrals

In this part, we give some estimates of singular integrals, which are vital to the proof of Theorem 2.1.

Lemma 3.2. *Let $H\in(1/2,1)$ and $\alpha\in(1-H,1)$. Then there exists some constant C such that for any $1\leq i<j\leq N$ and $s\in(t_{i-1},t_i]$,*

$$\int_0^{t_{i-1}}\int_{t_{i-1}}^s (s-u)^{\alpha-1}(u-v)^{2H-2}(t_{j-1}-v)^{\alpha-2H}\,\mathrm{d}u\,\mathrm{d}v\leq Ch^{2\alpha-1},\tag{3.11}$$

$$\int_0^{t_{i-1}}\int_v^{t_{i-1}}\Big((t_{i-1}-u)^{\alpha-1}-(s-u)^{\alpha-1}\Big)(t_{j-1}-v)^{\alpha-2H}(u-v)^{2H-2}\,\mathrm{d}u\,\mathrm{d}v\leq Ch^{2\alpha-1}.\tag{3.12}$$

Proof. For $\beta\in(-2H,1-2H)$, using (3.1) yields

$$\begin{aligned}&\int_0^{t_{i-1}}\int_{t_{i-1}}^s (s-u)^{\alpha-1}(u-v)^{2H-2}(t_{j-1}-v)^{\beta}\,\mathrm{d}u\,\mathrm{d}v\\&\quad\leq\int_{t_{i-1}}^s\int_0^{t_{i-1}}(s-u)^{\alpha-1}(u-v)^{2H-2+\beta}\,\mathrm{d}v\,\mathrm{d}u\leq C(s-t_{i-1})^{\alpha+2H-1+\beta}.\end{aligned}\tag{3.13}$$

We proceed to prove (3.11) by separating the cases $2\alpha \geq 1$ and $2\alpha < 1$. When $2\alpha \geq 1$, applying $s - t_{i-1} \leq h$ and (3.13) with $\beta = \alpha - 2H$ yields (3.11). While $2\alpha < 1$, by $t_{j-1} - v \geq h$, for $v \in [0, t_{i-1}]$, we have

$$\begin{aligned}
&\int_0^{t_{i-1}} \int_{t_{i-1}}^{s} (s-u)^{\alpha-1}(u-v)^{2H-2}(t_{j-1}-v)^{\alpha-2H}\,\mathrm{d}u\,\mathrm{d}v \\
&\quad = \int_{t_{i-1}}^{s} \int_0^{t_{i-1}} (s-u)^{\alpha-1}(u-v)^{2H-2}(t_{j-1}-v)^{2\alpha-1}(t_{j-1}-v)^{1-\alpha-2H}\,\mathrm{d}v\,\mathrm{d}u \\
&\quad \leq h^{2\alpha-1} \int_{t_{i-1}}^{s} \int_0^{t_{i-1}} (s-u)^{\alpha-1}(u-v)^{2H-2}(t_{j-1}-v)^{1-\alpha-2H}\,\mathrm{d}v\,\mathrm{d}u \leq Ch^{2\alpha-1},
\end{aligned}$$

thanks to (3.13) with $\beta = 1 - \alpha - 2H \in (-2H, 1-2H)$.

Now we turn to the proof of (3.12). Note that for any $\theta \in (-1, 0)$, there exists $C = C(\theta)$ such that

$$\int_0^u (u-r)^\theta - (t-r)^\theta\,\mathrm{d}r \leq C(t-u)^{1+\theta}, \qquad \forall 0 \leq u < t \leq T. \tag{3.14}$$

We also claim

$$\int_0^{t_{i-1}} \big((t_{i-1}-v)^{\alpha+2H-2} - (s-v)^{\alpha+2H-2}\big)(t_{j-1}-v)^{\alpha-2H}\,\mathrm{d}v \leq Ch^{2\alpha-1}. \tag{3.15}$$

Indeed, for $\alpha \in [2-2H, 1)$, (3.15) holds trivially since $t_{i-1} < s$ and $\alpha + 2H - 2 \geq 0$. For $\alpha \in (1-H, 2-2H)$, using $t_{j-1} - v \geq h$, for $v \in [0, t_{i-1}]$, yields

$$\begin{aligned}
&\int_0^{t_{i-1}} \Big((t_{i-1}-v)^{\alpha+2H-2} - (s-v)^{\alpha+2H-2}\Big)(t_{j-1}-v)^{\alpha-2H}\,\mathrm{d}v \\
&\qquad \leq h^{\alpha-2H} \int_0^{t_{i-1}} (t_{i-1}-v)^{\alpha+2H-2} - (s-v)^{\alpha+2H-2}\,\mathrm{d}v \leq Ch^{2\alpha-1},
\end{aligned}$$

thanks to (3.14) with $\theta = \alpha + 2H - 2 \in (-1, 0)$, which proves (3.15). Using (3.1) again,

$$\begin{aligned}
&\int_v^{t_{i-1}} \Big((t_{i-1}-u)^{\alpha-1} - (s-u)^{\alpha-1}\Big)(u-v)^{2H-2}\,\mathrm{d}u \\
&\quad = B(\alpha, 2H-1)\Big((t_{i-1}-v)^{\alpha+2H-2} - (s-v)^{\alpha+2H-2}\Big) + \int_{t_{i-1}}^{s} (s-u)^{\alpha-1}(u-v)^{2H-2}\,\mathrm{d}u.
\end{aligned} \tag{3.16}$$

Finally, equation (3.12) follows from (3.11), (3.15) and (3.16). The proof is completed. □

Lemma 3.3. *Let $H \in (1/2, 1)$ and $\alpha \in (1-H, 1)$. Then there exists some constant C such that for any $0 < u \leq t_{k-1} < s \leq t_k$ and $\tau \in (t_{l-1}, t_l]$ with $1 \leq k < l \leq N$,*

$$\int_{t_{k-1}}^{t_{l-1}} \Big((t_{l-1}-v)^{\alpha-1} - (\tau-v)^{\alpha-1}\Big)|v-u|^{2H-2}\,\mathrm{d}v \leq Ch^\alpha(\tau-s)^{2H-2}.$$

Proof. **Case 1: $k < l \leq k+2$.** In this case, $\tau - s \leq 3h$. Then, by (3.1),

$$\begin{aligned}
&\int_{t_{k-1}}^{t_{l-1}} \Big((t_{l-1}-v)^{\alpha-1} - (\tau-v)^{\alpha-1}\Big)|v-u|^{2H-2}\,\mathrm{d}v \\
&\qquad \leq \int_{t_{k-1}}^{t_{l-1}} (t_{l-1}-v)^{\alpha-1}(v-t_{k-1})^{2H-2}\,\mathrm{d}v \leq Ch^{\alpha+2H-2} \leq Ch^\alpha(\tau-s)^{2H-2}.
\end{aligned}$$

Case 2: $k+3 \le l \le N$**.** In this case, $t_{k-1} \le \frac{\tau+s}{2} \le t_{l-1} - \frac{1}{2}h$. For $v \in (t_{k-1}, \frac{\tau+s}{2})$, we obtain

$$\tau - v \le (l-k+1)h \le 4(l-k-2)h = 8(t_{l-1} - \frac{t_l + t_k}{2}) \le 8(t_{l-1} - \frac{\tau+s}{2}) \le 8(t_{l-1} - v),$$

which implies $\tau - s \le 2(\tau - v) \le 16(t_{l-1} - v)$ and thus

$$(t_{l-1} - v)^{\alpha-1} - (\tau - v)^{\alpha-1} \le (1-\alpha)(t_{l-1} - v)^{\alpha-2}h \le Ch(t_{l-1} - v)^{\alpha-2H}(\tau - s)^{2H-2}.$$

Hence, it follows from $u \le t_{k-1}$ and $t_{l-1} - v \ge \frac{1}{2}h$ for $v \in \left(t_{k-1}, \frac{\tau+s}{2}\right)$ that

$$\begin{aligned}
&\int_{t_{k-1}}^{\frac{\tau+s}{2}} \left((t_{l-1} - v)^{\alpha-1} - (\tau - v)^{\alpha-1}\right)|v-u|^{2H-2}\,\mathrm{d}v \\
&\quad \le Ch \int_{t_{k-1}}^{\frac{\tau+s}{2}} (t_{l-1} - v)^{\alpha-2H}(v-u)^{2H-2}\,\mathrm{d}v(\tau - s)^{2H-2} \\
&\quad \le Ch^{\alpha} \int_{t_{k-1}}^{t_{l-1}} (t_{l-1} - v)^{1-2H}(v - t_{k-1})^{2H-2}\,\mathrm{d}v(\tau - s)^{2H-2} \le Ch^{\alpha}(\tau - s)^{2H-2}.
\end{aligned}$$

For $v \in \left(\frac{\tau+s}{2}, t_{l-1}\right)$, we have $\tau - s \le 2(v-s) \le 2(v-u)$, for $u \le t_{k-1} < s$, and then

$$\begin{aligned}
&\int_{\frac{\tau+s}{2}}^{t_{l-1}} \left((t_{l-1} - v)^{\alpha-1} - (\tau - v)^{\alpha-1}\right)|v-u|^{2H-2}\,\mathrm{d}v \\
&\quad \le C \int_{\frac{\tau+s}{2}}^{t_{l-1}} \left((t_{l-1} - v)^{\alpha-1} - (\tau - v)^{\alpha-1}\right)\mathrm{d}v(\tau - s)^{2H-2} \le Ch^{\alpha}(\tau - s)^{2H-2},
\end{aligned}$$

which completes the proof.

□

Lemma 3.4. *Assume either*

$$\text{(C1) } \beta < 0 \text{ and } \gamma < 0 \text{ with } -1 < \beta + \gamma < 0 \quad \text{or} \quad \text{(C2) } \beta \ge 0,\ \gamma \in (-1, 0).$$

Then there exists some constant C such that for any $\tau \in (t_{j-1}, t_j]$ and $\eta \in (0, t_{j-1}]$ with $1 < j \le N$,

$$\int_0^{\eta} (\eta - r)^{\beta}\left((t_{j-1} - r)^{\gamma} - (\tau - r)^{\gamma}\right)\mathrm{d}r \le Ch^{(\beta+\gamma+1)\wedge(\gamma+1)}. \tag{3.17}$$

Proof. We first prove (3.17) under condition (C1). Notice that there exists a unique integer k satisfying $t_{k-1} < \eta \le t_k \le t_{j-1}$. Since $\gamma < 0$, we have for all $r \in (0, t_{k-1})$,

$$(t_{j-1} - r)^{\gamma} - (\tau - r)^{\gamma} \le (t_{j-1} - r)^{\gamma} - (t_j - r)^{\gamma} \le (t_{k-1} - r)^{\gamma} - (t_k - r)^{\gamma},$$

which together with $\beta < 0$ and $\beta + \gamma \in (-1, 0)$ indicates

$$\begin{aligned}
\int_0^{t_{k-1}} (\eta - r)^{\beta}((t_{j-1} - r)^{\gamma} - (\tau - r)^{\gamma})\,\mathrm{d}r &\le \int_0^{t_{k-1}} (t_{k-1} - r)^{\beta}((t_{k-1} - r)^{\gamma} - (t_k - r)^{\gamma})\,\mathrm{d}r \\
&\le \int_0^{t_{k-1}} (t_{k-1} - r)^{\beta+\gamma} - (t_k - r)^{\beta+\gamma}\,\mathrm{d}r \le Ch^{\beta+\gamma+1},
\end{aligned}$$

due to (3.14) with $\theta = \beta + \gamma$. In addition, since $t_{k-1} < \eta \le t_k \le t_{j-1}$ and $\gamma < 0$, it holds that

$$\int_{t_{k-1}}^{\eta} (\eta - r)^{\beta}\left((t_{j-1} - r)^{\gamma} - (\tau - r)^{\gamma}\right)\mathrm{d}r \le \int_{t_{k-1}}^{\eta} (\eta - r)^{\beta+\gamma}\,\mathrm{d}r \le Ch^{\beta+\gamma+1}.$$

Thus, one can conclude that (3.17) holds under condition (C1).

It remains to prove (3.17) under condition (C2). Actually, observe that $\beta \geq 0$ ensures $(\eta - r)^{\beta} \leq T^{\beta}$ for all $r \in [0, \eta]$. Hence, it follows from (3.14) with $\theta = \gamma$ that

$$\int_0^{\eta} (\eta - r)^{\beta}\big((t_{j-1} - r)^{\gamma} - (\tau - r)^{\gamma}\big)\,\mathrm{d}r \leq C\int_0^{t_{j-1}} (t_{j-1} - r)^{\gamma} - (\tau - r)^{\gamma}\,\mathrm{d}r \leq Ch^{\gamma+1}.$$

The proof is completed. □

Lemma 3.5. *Let $H \in (1/2, 1)$ and $\alpha \in (1-H, 1)$. Then there exists some constant C such that for any $\eta \in [0, T]$ and $\tau \in (t_{j-1}, t_j]$ with $1 \leq j \leq N$,*

$$\int_0^{\eta}\int_{t_{j-1}}^{\tau} (\eta - r)^{\alpha-1}(\tau - v)^{\alpha-1}|v - r|^{2H-2}\,\mathrm{d}v\,\mathrm{d}r \leq C\mathcal{R}_{H,\alpha}(h), \tag{3.18}$$

$$\int_0^{\eta}\int_0^{t_{j-1}} (\eta - r)^{\alpha-1}\big((t_{j-1} - v)^{\alpha-1} - (\tau - v)^{\alpha-1}\big)|v - r|^{2H-2}\,\mathrm{d}v\,\mathrm{d}r \leq C\mathcal{R}_{H,\alpha}(h), \tag{3.19}$$

where $\mathcal{R}_{H,\alpha}(h)$ is defined in Theorem 2.1.

Proof. We will first prove (3.18) with $1 < j \leq N$. Specifically, it is divided into the following four cases.

Case 1. $\eta \in [0, t_1]$. The estimate (3.18) can be obtained by

$$\begin{aligned}\int_0^{\eta}\int_{t_{j-1}}^{\tau} (\eta - r)^{\alpha-1}(\tau - v)^{\alpha-1}|v - r|^{2H-2}\,\mathrm{d}v\,\mathrm{d}r &\leq \int_0^{\eta}\int_r^{\tau} (\eta - r)^{\alpha-1}(\tau - v)^{\alpha-1}(v - r)^{2H-2}\,\mathrm{d}v\,\mathrm{d}r\\ &\leq C\int_0^{\eta} (\eta - r)^{\alpha-1}(\tau - r)^{\alpha+2H-2}\,\mathrm{d}r \leq Ch^{\alpha\wedge 2(\alpha+H-1)}.\end{aligned}$$

Case 2. $\eta \in (t_1, t_{j-1}]$. In this case, we have

$$\begin{aligned}\int_0^{\eta-h}\int_{t_{j-1}}^{\tau} (\eta - r)^{\alpha-1}(\tau - v)^{\alpha-1}|v - r|^{2H-2}\,\mathrm{d}v\,\mathrm{d}r &\leq \int_0^{\eta-h}\int_{t_{j-1}}^{\tau} (\eta - r)^{\alpha+2H-3}(\tau - v)^{\alpha-1}\,\mathrm{d}v\,\mathrm{d}r\\ &\leq Ch^{\alpha}\int_0^{\eta-h} (\eta - r)^{\alpha+2H-3}\,\mathrm{d}r\\ &\leq \begin{cases} Ch^{\alpha}h^{\alpha+2H-2}, & \text{if } \alpha \in (1-H, 2-2H),\\ Ch^{\alpha}(|\ln h| \vee \ln T), & \text{if } \alpha = 2-2H,\\ Ch^{\alpha}, & \text{if } \alpha \in (2-2H, 1)\end{cases}\\ &\leq C\mathcal{R}_{H,\alpha}(h).\end{aligned}$$

Besides, in virtue of $\tau > t_{j-1} \geq \eta$,

$$\int_{\eta-h}^{\eta}\int_{t_{j-1}}^{\tau} (\eta - r)^{\alpha-1}(\tau - v)^{\alpha-1}|v - r|^{2H-2}\,\mathrm{d}v\,\mathrm{d}r \leq C\int_{\eta-h}^{\eta} (\eta - r)^{\alpha-1}(\tau - r)^{\alpha+2H-2}\,\mathrm{d}r \leq Ch^{\alpha\wedge 2(\alpha+H-1)}.$$

Thus, the estimate (3.18) holds for Case 2.

Case 3. $\eta \in (t_{j-1}, \tau]$. In view of the result of Case 2, the left hand side of (3.18) can be controlled by

$$C\mathcal{R}_{H,\alpha}(h) + \int_{t_{j-1}}^{\eta}\int_{t_{j-1}}^{\tau} (\eta - r)^{\alpha-1}(\tau - v)^{\alpha-1}|r - v|^{2H-2}\,\mathrm{d}v\,\mathrm{d}r,$$

where the second term can be bounded by

$$\begin{aligned}
&\int_{t_{j-1}}^{\eta}\int_{t_{j-1}}^{r}(\eta-r)^{\alpha-1}(\tau-v)^{\alpha-1}(r-v)^{2H-2}\,\mathrm{d}v\,\mathrm{d}r+\int_{t_{j-1}}^{\eta}\int_{r}^{\tau}(\eta-r)^{\alpha-1}(\tau-v)^{\alpha-1}(v-r)^{2H-2}\,\mathrm{d}v\,\mathrm{d}r\\
&\qquad\le\int_{t_{j-1}}^{\eta}\int_{v}^{\eta}(\tau-v)^{\alpha-1}(\eta-r)^{\alpha-1}(r-v)^{2H-2}\,\mathrm{d}r\,\mathrm{d}v+C\int_{t_{j-1}}^{\eta}(\eta-r)^{\alpha-1}(\tau-r)^{\alpha+2H-2}\,\mathrm{d}r\\
&\qquad\le Ch^{2(\alpha+H-1)}.
\end{aligned}\tag{3.20}$$

Hence, the estimate (3.18) holds for Case 3.

Case 4. $\eta\in(\tau,T]$. Utilizing the result of Case 3, the estimate (3.18) can be attained by

$$\begin{aligned}
&\int_{0}^{\eta}\int_{t_{j-1}}^{\tau}(\eta-r)^{\alpha-1}(\tau-v)^{\alpha-1}|v-r|^{2H-2}\,\mathrm{d}v\,\mathrm{d}r\\
&\quad=\int_{0}^{\tau}\int_{t_{j-1}}^{\tau}(\eta-r)^{\alpha-1}(\tau-v)^{\alpha-1}|v-r|^{2H-2}\,\mathrm{d}v\,\mathrm{d}r+\int_{\tau}^{\eta}\int_{t_{j-1}}^{\tau}(\eta-r)^{\alpha-1}(\tau-v)^{\alpha-1}(r-v)^{2H-2}\,\mathrm{d}v\,\mathrm{d}r\\
&\quad\le C\mathcal{R}_{H,\alpha}(h)+C\int_{t_{j-1}}^{\tau}(\eta-v)^{\alpha+2H-2}(\tau-v)^{\alpha-1}\,\mathrm{d}v\\
&\quad\le C\mathcal{R}_{H,\alpha}(h)+Ch^{\alpha\wedge 2(\alpha+H-1)}\le C\mathcal{R}_{H,\alpha}(h).
\end{aligned}$$

We have proven (3.18) when $1<j\le N$. It remains to prove (3.18) for the case $j=1$. In this case, one can divide $\eta\in[0,T]$ into the cases $\eta\in[0,\tau]$ and $\eta\in(\tau,T]$, which will fall into the above Cases 1 and 4, respectively. Thus, the proof of (3.18) is completed.

Next, we turn to proving (3.19) under the settings $\eta\in[0,t_{j-1}]$ and $\eta\in(t_{j-1},T]$, separately. When $\eta\in[0,t_{j-1}]$, it follows from Lemma 3.4 that

$$\int_{0}^{\eta}(\eta-r)^{\alpha+2H-2}\big((t_{j-1}-r)^{\alpha-1}-(\tau-r)^{\alpha-1}\big)\,\mathrm{d}r\le Ch^{\alpha\wedge 2(\alpha+H-1)}.\tag{3.21}$$

Similarly, for $\alpha<2-2H$, Lemma 3.4 also implies that

$$\int_{0}^{\eta}(\eta-r)^{\alpha-1}\big((t_{j-1}-r)^{\alpha+2H-2}-(\tau-r)^{\alpha+2H-2}\big)\,\mathrm{d}r\le Ch^{2(\alpha+H-1)},\tag{3.22}$$

and (3.22) holds trivially for $\alpha\ge 2-2H$ since $t_{j-1}\le\tau$. The left hand side of (3.19) is equal to

$$\begin{aligned}
\mathcal{T}_1+\mathcal{T}_2:=&\int_{0}^{\eta}\int_{0}^{r}(\eta-r)^{\alpha-1}\big((t_{j-1}-v)^{\alpha-1}-(\tau-v)^{\alpha-1}\big)(r-v)^{2H-2}\,\mathrm{d}v\,\mathrm{d}r\\
&+\int_{0}^{\eta}\int_{r}^{t_{j-1}}(\eta-r)^{\alpha-1}\big((t_{j-1}-v)^{\alpha-1}-(\tau-v)^{\alpha-1}\big)(v-r)^{2H-2}\,\mathrm{d}v\,\mathrm{d}r.
\end{aligned}$$

By (3.1) and (3.21),

$$\mathcal{T}_1\le C\int_{0}^{\eta}(\eta-v)^{\alpha+2H-2}\big((t_{j-1}-v)^{\alpha-1}-(\tau-v)^{\alpha-1}\big)\,\mathrm{d}v\le Ch^{\alpha\wedge 2(\alpha+H-1)},$$

and using a similar argument of (3.16) gives

$$\mathcal{T}_2\le C\int_{0}^{\eta}(\eta-r)^{\alpha-1}\Big((t_{j-1}-r)^{\alpha+2H-2}-(\tau-r)^{\alpha+2H-2}+\int_{t_{j-1}}^{\tau}(\tau-v)^{\alpha-1}(v-r)^{2H-2}\,\mathrm{d}v\Big)\,\mathrm{d}r,$$

which can be further bounded by $C\mathcal{R}_{H,\alpha}(h)$, in view of (3.18) and (3.22). Thus, (3.19) holds for $\eta \in [0, t_{j-1}]$.

When $\eta \in (t_{j-1}, T]$, the left hand side of (3.19) can be bounded by

$$\begin{aligned}\mathcal{T}_1' + \mathcal{T}_2' := &\int_0^{t_{j-1}} \int_0^{t_{j-1}} (t_{j-1} - r)^{\alpha-1} \big((t_{j-1} - v)^{\alpha-1} - (\tau - v)^{\alpha-1}\big) |v - r|^{2H-2} \, \mathrm{d}v \, \mathrm{d}r \\ &+ \int_{t_{j-1}}^{\eta} \int_0^{t_{j-1}} (\eta - r)^{\alpha-1} \big((t_{j-1} - v)^{\alpha-1} - (\tau - v)^{\alpha-1}\big) (r - v)^{2H-2} \, \mathrm{d}v \, \mathrm{d}r,\end{aligned}$$

in which $\mathcal{T}_1' \leq Ch^{\alpha \wedge 2(\alpha+H-1)}$ by the obtained result with $\eta = t_{j-1}$. Besides,

$$\mathcal{T}_2' \leq C \int_0^{t_{j-1}} (\eta - v)^{\alpha+2H-2} \big((t_{j-1} - v)^{\alpha-1} - (\tau - v)^{\alpha-1}\big) \, \mathrm{d}v \leq \begin{cases} Ch^{\alpha}, & \text{if } \alpha \in [2 - 2H, 1), \\ Ch^{2(\alpha+H-1)}, & \text{if } \alpha \in (1 - H, 2 - 2H), \end{cases}$$

in which (3.14) with $\theta = \alpha - 1 \in (-1, 0)$ is used to deal with the case $\alpha \in [2 - 2H, 1)$, while the inequality

$$\begin{aligned}&\int_0^{t_{j-1}} (\eta - v)^{\alpha+2H-2} \big((t_{j-1} - v)^{\alpha-1} - (\tau - v)^{\alpha-1}\big) \, \mathrm{d}v \\ &\quad \leq C \int_0^{t_{j-1}} (t_{j-1} - v)^{\alpha+2H-2} \big((t_{j-1} - v)^{\alpha-1} - (\tau - v)^{\alpha-1}\big) \, \mathrm{d}v\end{aligned}$$

and (3.21) are used to handle the case $\alpha \in (1 - H, 2 - 2H)$. Hence, (3.19) holds for $\eta \in (t_{j-1}, T]$. The proof is completed. □

4. Proofs

In this section, we provide the detailed proof of Theorem 2.1.

4.1. Proof of Theorem 2.1

We denote $\hat{s} := t_{j-1}$ for $s \in (t_{j-1}, t_j]$ with $j = 1, 2, \cdots, N$. By introducing

$$R_n := \sum_{j=1}^{n} \int_{t_{j-1}}^{t_j} (t_n - s)^{\alpha-1} \big(b(x(s)) - b(x(\hat{s}))\big) \, \mathrm{d}s, \quad n \in \{1, 2, \cdots, N\}$$

and according to (2.1) and (2.2), the strong error of the Euler method satisfies

$$\begin{aligned}\|x_n - x(t_n)\| &\leq \frac{1}{\Gamma(\alpha)} \left\| \sum_{j=1}^{n} \int_{t_{j-1}}^{t_j} (t_n - s)^{\alpha-1} (b(x_{j-1}) - b(x(\hat{s}))) \, \mathrm{d}s \right\| \\ &\quad + \frac{1}{\Gamma(\alpha)} \left\| \sum_{j=1}^{n} \int_{t_{j-1}}^{t_j} (t_n - s)^{\alpha-1} (b(x(\hat{s})) - b(x(s))) \, \mathrm{d}s \right\| \\ &\leq C \sum_{j=1}^{n} \int_{t_{j-1}}^{t_j} (t_n - s)^{\alpha-1} \|x_{j-1} - x(\hat{s})\| \, \mathrm{d}s + \frac{1}{\Gamma(\alpha)} \|R_n\|,\end{aligned}$$

for any $n \in \{1, 2, \cdots, N\}$, where the last step used the assumption $b \in C_b^1$. Then, applying the singular Grönwall inequality yields

$$\|x_n - x(t_n)\| \leq C \|R_n\|. \tag{4.1}$$

In order to estimate $\|R_n\|$, the GLE (2.1) is reformulated as

$$x(t) = \zeta(t) + G(t), \qquad \text{where } \zeta(t) := x_0 + \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} b(x(s))\,\mathrm{d}s. \tag{4.2}$$

The mean value theorem implies $b(x(s)) - b(x(\hat{s})) = \big(x(s) - x(\hat{s})\big) \int_0^1 b'\big(\xi_s^\theta\big)\,\mathrm{d}\theta$ with

$$\xi_s^\theta := x(\hat{s}) + \theta\big(x(s) - x(\hat{s})\big). \tag{4.3}$$

This along with (4.2) gives

$$\begin{aligned}\|R_n\| &\le \left\| \sum_{j=1}^n \int_{t_{j-1}}^{t_j} (t_n - s)^{\alpha-1} (\zeta(s) - \zeta(\hat{s})) \int_0^1 b'\big(\xi_s^\theta\big)\,\mathrm{d}\theta\,\mathrm{d}s \right\| \\ &\quad + \left\| \sum_{j=1}^n \int_{t_{j-1}}^{t_j} (t_n - s)^{\alpha-1} (G(s) - G(\hat{s})) \int_0^1 b'\big(\xi_s^\theta\big)\,\mathrm{d}\theta\,\mathrm{d}s \right\| \\ &=: \|R_{n,1}\| + \|R_{n,2}\|. \end{aligned} \tag{4.4}$$

Since $b \in C_b^1$ and ζ is α-Hölder continuous in $L^2(\Omega;\mathbb{R})$ (see [5], Page 456),

$$\|R_{n,1}\| \le C \sum_{j=1}^n \int_{t_{j-1}}^{t_j} (t_n - s)^{\alpha-1} \big\|\zeta(s) - \zeta(\hat{s})\big\|\,\mathrm{d}s \le Ch^\alpha \sum_{j=1}^n \int_{t_{j-1}}^{t_j} (t_n - s)^{\alpha-1}\,\mathrm{d}s \le Ch^\alpha. \tag{4.5}$$

To estimate $\|R_{n,2}\|$, we denote

$$I_{n,2}^j := \int_0^1 \int_{t_{j-1}}^{t_j} (t_n - s)^{\alpha-1} \big(G(s) - G(\hat{s})\big) b'\big(\xi_s^\theta\big)\,\mathrm{d}s\,\mathrm{d}\theta,$$

and make the decomposition

$$\|R_{n,2}\|^2 = \sum_{j=1}^n \left\| I_{n,2}^j \right\|^2 + 2 \sum_{1 \le i < j \le n} \left\langle I_{n,2}^i, I_{n,2}^j \right\rangle.$$

For the first term of the right hand side, using the assumption $b \in C_b^1$ and the fact that G is $(H+\alpha-1)$-Hölder continuous in $L^2(\Omega;\mathbb{R})$ (see [18], Prop. 1) yield

$$\begin{aligned}\sum_{j=1}^n \left\| I_{n,2}^j \right\|^2 &\le C \sum_{j=1}^n \left(\int_{t_{j-1}}^{t_j} (t_n - s)^{\alpha-1} \|G(s) - G(\hat{s})\|\,\mathrm{d}s \right)^2 \\ &\le Ch^{2(H+\alpha-1)} \sum_{j=1}^n \left(\int_{t_{j-1}}^{t_j} (t_n - s)^{\alpha-1}\,\mathrm{d}s \right)^2 \le Ch^{2(H+2\alpha-1)} \left(1 + \sum_{j=1}^{n-1} (n-j)^{2(\alpha-1)} \right) \\ &\le \begin{cases} Ch^{2H+4\alpha-2}, & \text{if } 1 - H < \alpha < \frac{1}{2}, \\ C\big(|\ln h| \vee \ln T\big) h^{2H}, & \text{if } \alpha = \frac{1}{2}, \\ Ch^{2H+2\alpha-1}, & \text{if } \frac{1}{2} < \alpha < 1, \end{cases}\end{aligned}$$

which can be further bounded by $Ch^{2\alpha}$. Thus,

$$\|R_{n,2}\|^2 \le Ch^{2\alpha} + 2 \sum_{1 \le i < j \le n} \left\langle I_{n,2}^i, I_{n,2}^j \right\rangle. \tag{4.6}$$

Therefore, the combination of (4.1), (4.4)–(4.6) shows

$$\|x_n - x(t_n)\|^2 \leq Ch^{2\alpha} + 2 \sum_{1\leq i<j\leq n} \Big\langle I^i_{n,2}, I^j_{n,2} \Big\rangle, \tag{4.7}$$

where $\langle I^i_{n,2}, I^j_{n,2}\rangle$ equals to

$$\int_0^1 \int_0^1 \int_{t_{i-1}}^{t_i} \int_{t_{j-1}}^{t_j} (t_n - s)^{\alpha-1}(t_n-\tau)^{\alpha-1}\mathbb{E}\big[b'\big(\xi^\theta_s\big)b'\big(\xi^\lambda_\tau\big)(G(s)-G(\hat{s}))(G(\tau)-G(\hat{\tau}))\big]\,\mathrm{d}\tau\,\mathrm{d}s\,\mathrm{d}\theta\,\mathrm{d}\lambda. \tag{4.8}$$

Remark 4.1. The key of the rest proof of Theorem 2.1 lies in carefully estimating the expectation

$$\mathbb{E}\big[b'\big(\xi^\theta_s\big)b'\big(\xi^\lambda_\tau\big)\big(G(s)-G(\hat{s})\big)\big(G(\tau)-G(\hat{\tau})\big)\big] \tag{4.9}$$

in the multiple integral (4.8). Due to the correlation among the random variables $b'\big(\xi^\theta_s\big)b'\big(\xi^\lambda_\tau\big)$, $G(s)-G(\hat{s})$ and $G(\tau)-G(\hat{\tau})$, a common way to bound the expectation (4.9) is applying the assumption $b\in C^1_b$, Hölder's inequality and the fact that G is $(H+\alpha-1)$-Hölder continuous in $L^2(\Omega;\mathbb{R})$ to obtain

$$\begin{aligned}\mathbb{E}\big[b'\big(\xi^\theta_s\big)b'\big(\xi^\lambda_\tau\big)\big(G(s)-G(\hat{s})\big)\big(G(\tau)-G(\hat{\tau})\big)\big] &\leq C\|G(s)-G(\hat{s})\|\|G(\tau)-G(\hat{\tau})\| \\ &\leq C(s-\hat{s})^{H+\alpha-1}(\tau-\hat{\tau})^{H+\alpha-1} \\ &\leq Ch^{2(H+\alpha-1)},\end{aligned}$$

which together with (4.7) and (4.8) implies that one can only derive the order $\alpha+H-1$ for the convergence rate of $\|x_n - x(t_n)\|$ in this way. Here, the low convergence order $\alpha+H-1$ is mainly restricted by the low Hölder regularity of G.

Our main approach to improving the convergence order of $\|x_n - x(t_n)\|$ lies in the application of Malliavin calculus to obtain a finer estimate for the expectation (4.9) (see Prop. 4.1 below), which can avoid using the low Hölder regularity of G. To be specific, by noticing that G is actually a Skorohod integral, we use the duality between the Malliavin derivative and the Skorohod integral to formulate an equivalent representation of the expectation (4.9), as shown in (4.11) below. Deriving this equivalent representation uses the covariance structure of the fBm that is characterized by the inner product of $\mathcal{H}$. Since the correlation between the increments of G can be reserved in the right hand side of (4.11), it is possible to obtain a finer estimate for the right hand side of (4.11), compared to applying Hölder's inequality to the left hand side of (4.11) directly.

Proposition 4.1. *For $s\in(t_{i-1},t_i]$, $\tau\in(t_{j-1},t_j]$ with $1\leq i<j\leq N$, and $\theta,\lambda\in(0,1)$, let ξ^θ_s, ξ^λ_τ be given by* (4.3)*. Under the assumptions of Theorem 2.1,*

$$\mathbb{E}\Big[b'\big(\xi^\theta_s\big)b'\big(\xi^\lambda_\tau\big)\big(G(s)-G(\hat{s})\big)\big(G(\tau)-G(\hat{\tau})\big)\Big] \leq C\mathcal{R}^2_{H,\alpha}(h)(\tau-s)^{2H-2},$$

where G is defined by (2.3) *and the constant $C=C(\alpha,H,T,\sigma)$ is independent of $\theta,\lambda,s,\tau,i,j$.*

The proof of Proposition 4.1 is deferred to Section 4.2, and we proceed to complete the proof of Theorem 2.1. With Proposition 4.1 in mind, one can derive

$$\begin{aligned}\sum_{1\leq i<j\leq n}\langle I^i_{n,2}, I^j_{n,2}\rangle &\leq C\mathcal{R}^2_{H,\alpha}(h)\int_0^{t_{n-1}}\int_s^{t_n}(t_n-s)^{\alpha-1}(t_n-\tau)^{\alpha-1}(\tau-s)^{2H-2}\,\mathrm{d}\tau\,\mathrm{d}s \\ &= C\mathcal{R}^2_{H,\alpha}(h)\int_h^{t_n}\int_0^u u^{\alpha-1}v^{\alpha-1}(u-v)^{2H-2}\,\mathrm{d}v\,\mathrm{d}u \\ &\leq C\mathcal{R}^2_{H,\alpha}(h)\int_h^{t_n} u^{2\alpha+2H-3}\,\mathrm{d}u \leq C\mathcal{R}^2_{H,\alpha}(h),\end{aligned} \tag{4.10}$$

where the last step used the fact $2\alpha+2H-3>-1$ since $\alpha\in(1-H,1)$. Substituting (4.10) into (4.7), then the proof of Theorem 2.1 is completed by the relation $h^\alpha\leq C\mathcal{R}_{H,\alpha}(h)$. ☐

4.2. Proof of Proposition 4.1

For all $0 < s < t \le T$, the increment of the diffusion term can be rewritten as

$$G(t) - G(s) = \frac{\sigma}{\Gamma(\alpha)}\Big(\int_0^s (t-u)^{\alpha-1} - (s-u)^{\alpha-1}\,\mathrm{d}W_H(u) + \int_s^t (t-u)^{\alpha-1}\,\mathrm{d}W_H(u)\Big)$$
$$= \frac{\sigma}{\Gamma(\alpha)}\delta\Big(((t-\cdot)^{\alpha-1} - (s-\cdot)^{\alpha-1})\mathbf{1}_{(0,s)}(\cdot) + (t-\cdot)^{\alpha-1}\mathbf{1}_{(s,t)}(\cdot)\Big),$$

where δ is the Skorohod integral introduced in Section 3.1. For convenience, for $s \in (0,T]$, we denote

$$\widehat{A}(\cdot\,;s) := \big((\hat{s}-\cdot)^{\alpha-1} - (s-\cdot)^{\alpha-1}\big)\mathbf{1}_{(0,\hat{s})}(\cdot), \qquad \widetilde{A}(\cdot\,;s) := (s-\cdot)^{\alpha-1}\mathbf{1}_{(\hat{s},s)}(\cdot),$$
$$A(\cdot\,;s) := \widehat{A}(\cdot\,;s) + \widetilde{A}(\cdot\,;s), \qquad U(\cdot\,;s) := -\widehat{A}(\cdot\,;s) + \widetilde{A}(\cdot\,;s).$$

It follows from the dual formula ([13], Eq. (25)) that

$$\mathbb{E}\big[b'(\xi_s^\theta)b'(\xi_\tau^\lambda)(G(s)-G(\hat{s}))(G(\tau)-G(\hat{\tau}))\big] = \frac{\sigma^2}{\Gamma^2(\alpha)}\mathbb{E}\big[b'(\xi_s^\theta)b'(\xi_\tau^\lambda)\delta(U(\cdot;s))\delta(U(\cdot;\tau))\big]$$
$$= \frac{\sigma^2}{\Gamma^2(\alpha)}\mathbb{E}\big[b'(\xi_s^\theta)b'(\xi_\tau^\lambda)\langle U(\cdot;s), U(\cdot;\tau)\rangle_{\mathcal{H}}\big] + \frac{\sigma^2}{\Gamma^2(\alpha)}\mathbb{E}\Big[\big\langle D\big\langle D\big(b'(\xi_s^\theta)b'(\xi_\tau^\lambda)\big), U(\cdot\,;s)\big\rangle_{\mathcal{H}}, U(\cdot\,;\tau)\big\rangle_{\mathcal{H}}\Big]. \tag{4.11}$$

Then,

$$\mathbb{E}\big[b'(\xi_s^\theta)b'(\xi_\tau^\lambda)(G(s)-G(\hat{s}))(G(\tau)-G(\hat{\tau}))\big]$$
$$\le C\int_{[0,T]^2}\big\|b'(\xi_s^\theta)b'(\xi_\tau^\lambda)\big\|_1 A(u;s)A(v;\tau)|u-v|^{2H-2}\,\mathrm{d}u\,\mathrm{d}v$$
$$+ C\int_{[0,T]^4}\big\|D_{r_2}D_{r_1}\big(b'(\xi_s^\theta)b'(\xi_\tau^\lambda)\big)\big\|_1 A(u;s)A(v;\tau)|r_1-u|^{2H-2}|r_2-v|^{2H-2}\,\mathrm{d}r_1\,\mathrm{d}u\,\mathrm{d}r_2\,\mathrm{d}v. \tag{4.12}$$

Applying the product rule and chain rule of the Malliavin derivative obtains

$$D_{r_1}\big(b'(\xi_s^\theta)b'(\xi_\tau^\lambda)\big) = b''(\xi_s^\theta)D_{r_1}\xi_s^\theta b'(\xi_\tau^\lambda) + b'(\xi_s^\theta)b''(\xi_\tau^\lambda)D_{r_1}\xi_\tau^\lambda, \tag{4.13}$$

and

$$D_{r_2}D_{r_1}\big(b'(\xi_s^\theta)b'(\xi_\tau^\lambda)\big) = b'''(\xi_s^\theta)D_{r_2}\xi_s^\theta D_{r_1}\xi_s^\theta b'(\xi_\tau^\lambda) + b''(\xi_s^\theta)D_{r_2}D_{r_1}\xi_s^\theta b'(\xi_\tau^\lambda)$$
$$+ b''(\xi_s^\theta)D_{r_1}\xi_s^\theta b''(\xi_\tau^\lambda)D_{r_2}\xi_\tau^\lambda + b''(\xi_s^\theta)D_{r_2}\xi_s^\theta b''(\xi_\tau^\lambda)D_{r_1}\xi_\tau^\lambda$$
$$+ b'(\xi_s^\theta)b'''(\xi_\tau^\lambda)D_{r_2}\xi_\tau^\lambda D_{r_1}\xi_\tau^\lambda + b'(\xi_s^\theta)b''(\xi_\tau^\lambda)D_{r_2}D_{r_1}\xi_\tau^\lambda. \tag{4.14}$$

Invoking (4.13) and (4.14), it follows from $b \in C_{b,p}^{1,3}$, Hölder's inequality and (3.7) that

$$\big\|D_{r_2}D_{r_1}\big(b'(\xi_s^\theta)b'(\xi_\tau^\lambda)\big)\big\|_1 \le C\big(\|D_{r_2}D_{r_1}\xi_s^\theta\| + \|D_{r_2}D_{r_1}\xi_\tau^\lambda\|\big)$$
$$+ C\big(\|D_{r_1}\xi_s^\theta\|_4 + \|D_{r_1}\xi_\tau^\lambda\|_4\big)\big(\|D_{r_2}\xi_s^\theta\|_4 + \|D_{r_2}\xi_\tau^\lambda\|_4\big).$$

Note that $\big\|b'(\xi_s^\theta)b'(\xi_\tau^\lambda)\big\|_1$ is bounded since $b \in C_b^1$. Thus, by (4.12),

$$\mathbb{E}\big[b'(\xi_s^\theta)b'(\xi_\tau^\lambda)(G(s)-G(\hat{s}))(G(\tau)-G(\hat{\tau}))\big] \le C(\mathcal{J}+\mathcal{K}+\mathcal{L}), \tag{4.15}$$

where

$$\mathcal{J} := \int_{[0,T]^2} A(u;s)A(v;\tau)|u-v|^{2H-2}\,\mathrm{d}u\,\mathrm{d}v, \tag{4.16}$$

$$\begin{aligned}\mathcal{K} := &\int_{[0,T]^2} \big(\|D_{r_1}\xi_s^\theta\|_4 + \|D_{r_1}\xi_\tau^\lambda\|_4\big) A(u;s)|r_1-u|^{2H-2}\,\mathrm{d}r_1\,\mathrm{d}u \\ &\times \int_{[0,T]^2} \big(\|D_{r_2}\xi_s^\theta\|_4 + \|D_{r_2}\xi_\tau^\lambda\|_4\big) A(v;\tau)|r_2-v|^{2H-2}\,\mathrm{d}r_2\,\mathrm{d}v, \end{aligned} \tag{4.17}$$

$$\mathcal{L} := \int_{[0,T]^4} \big(\|D_{r_2}D_{r_1}\xi_s^\theta\| + \|D_{r_2}D_{r_1}\xi_\tau^\lambda\|\big) A(u;s)A(v;\tau)|r_1-u|^{2H-2}|r_2-v|^{2H-2}\,\mathrm{d}r_1\,\mathrm{d}u\,\mathrm{d}r_2\,\mathrm{d}v. \tag{4.18}$$

Finally, the proof of Proposition 4.1 is completed by combining Lemmas 4.1 and 4.2 with (4.15). $\square$

The rest of this section is devoted to proving Lemmas 4.1 and 4.2.

Lemma 4.1. *For $\mathcal{J}$ given by (4.16), there exists some constant C such that*

$$\mathcal{J} \le Ch^{2\alpha}(\tau-s)^{2H-2},$$

for any $s\in(t_{i-1},t_i]$ and $\tau\in(t_{j-1},t_j]$ with $1\le i<j\le N$.

Proof. According to $A=\widehat{A}+\widetilde{A}$ and (4.16), we have $\mathcal{J}=\mathcal{J}_1+\mathcal{J}_2+\mathcal{J}_3+\mathcal{J}_4$ with

$$\mathcal{J}_1 := \int_{[0,T]^2} \widehat{A}(u;s)\widehat{A}(v;\tau)|u-v|^{2H-2}\,\mathrm{d}u\,\mathrm{d}v, \quad \mathcal{J}_2 := \int_{[0,T]^2} \widehat{A}(u;s)\widetilde{A}(v;\tau)|u-v|^{2H-2}\,\mathrm{d}u\,\mathrm{d}v,$$
$$\mathcal{J}_3 := \int_{[0,T]^2} \widetilde{A}(u;s)\widehat{A}(v;\tau)|u-v|^{2H-2}\,\mathrm{d}u\,\mathrm{d}v, \quad \mathcal{J}_4 := \int_{[0,T]^2} \widetilde{A}(u;s)\widetilde{A}(v;\tau)|u-v|^{2H-2}\,\mathrm{d}u\,\mathrm{d}v.$$

Estimate of $\mathcal{J}_1$. We firstly decompose $\mathcal{J}_1$ into

$$\begin{aligned}\mathcal{J}_1 &= \int_0^{t_{i-1}}\int_0^{t_{j-1}} \big((t_{i-1}-u)^{\alpha-1}-(s-u)^{\alpha-1}\big)\big((t_{j-1}-v)^{\alpha-1}-(\tau-v)^{\alpha-1}\big)|v-u|^{2H-2}\,\mathrm{d}v\,\mathrm{d}u \\ &= \int_0^{t_{i-1}}\int_0^{t_{i-1}} \big((t_{i-1}-u)^{\alpha-1}-(s-u)^{\alpha-1}\big)\big((t_{j-1}-v)^{\alpha-1}-(\tau-v)^{\alpha-1}\big)|v-u|^{2H-2}\,\mathrm{d}v\,\mathrm{d}u \\ &\quad + \int_0^{t_{i-1}}\int_{t_{i-1}}^{t_{j-1}} \big((t_{i-1}-u)^{\alpha-1}-(s-u)^{\alpha-1}\big)\big((t_{j-1}-v)^{\alpha-1}-(\tau-v)^{\alpha-1}\big)|v-u|^{2H-2}\,\mathrm{d}v\,\mathrm{d}u \\ &=: \mathcal{J}_{11}+\mathcal{J}_{12}.\end{aligned}$$

In order to estimate $\mathcal{J}_{11}$, we note that for $v\in[0,t_{i-1}]$, one has $\tau-s\le 2(t_{j-1}-v)$, which implies

$$(t_{j-1}-v)^{\alpha-1}-(\tau-v)^{\alpha-1} \le h(t_{j-1}-v)^{\alpha-2} \le Ch(t_{j-1}-v)^{\alpha-2H}(\tau-s)^{2H-2}. \tag{4.19}$$

Besides,

$$\begin{aligned}&\int_0^{t_{i-1}}\int_0^{t_{i-1}} \Big((t_{i-1}-u)^{\alpha-1}-(s-u)^{\alpha-1}\Big)(t_{j-1}-v)^{\alpha-2H}|v-u|^{2H-2}\,\mathrm{d}v\,\mathrm{d}u \\ &\quad\le \int_0^{t_{i-1}}\int_v^{t_{i-1}} \Big((t_{i-1}-u)^{\alpha-1}-(s-u)^{\alpha-1}\Big)(t_{j-1}-v)^{\alpha-2H}(u-v)^{2H-2}\,\mathrm{d}u\,\mathrm{d}v \\ &\qquad + h^{\alpha-1}\int_0^{t_{i-1}}\int_u^{t_{i-1}} \Big((t_{i-1}-u)^{\alpha-1}-(s-u)^{\alpha-1}\Big)(t_{j-1}-v)^{1-2H}(v-u)^{2H-2}\,\mathrm{d}v\,\mathrm{d}u \\ &\quad\le Ch^{2\alpha-1}+Ch^{\alpha-1}\int_0^{t_{i-1}}(t_{i-1}-u)^{\alpha-1}-(s-u)^{\alpha-1}\,\mathrm{d}u \le Ch^{2\alpha-1},\end{aligned} \tag{4.20}$$

in which $(t_{j-1}-v)^{\alpha-1} \le h^{\alpha-1}$ for $v \le t_{i-1} < t_{j-1}$ was used in the first inequality, equation (3.12), $t_{i-1} < t_{j-1}$, and (3.1) were used in the second inequality, and (3.14) with $\theta = \alpha - 1 \in (-1, 0)$ was used in the last inequality. Then, the combination of (4.19) and (4.20) reveals

$$\mathcal{J}_{11} \le Ch^{2\alpha}(\tau - s)^{2H-2}. \tag{4.21}$$

By the virtue of Lemma 3.3 with $k = i$ and $l = j$, as well as (3.14) with $\theta = \alpha - 1$,

$$\mathcal{J}_{12} \le Ch^{\alpha}(\tau - s)^{2H-2} \int_0^{t_{i-1}} (t_{i-1} - u)^{\alpha-1} - (s-u)^{\alpha-1}\,\mathrm{d}u \le Ch^{2\alpha}(\tau - s)^{2H-2}. \tag{4.22}$$

Collecting (4.21) and (4.22), one obtains $\mathcal{J}_1 = \mathcal{J}_{11} + \mathcal{J}_{12} \le Ch^{2\alpha}(\tau - s)^{2H-2}$.

Estimate of $\mathcal{J}_2$. For $v \in [t_{j-1}, \tau]$ and $u \in [0, t_{i-1}]$, $\tau - s = \tau - v + v - s \le h + v - u \le 2(v-u)$, which along with (3.14) with $\theta = \alpha - 1$ implies

$$\begin{aligned}
\mathcal{J}_2 &= \int_0^{t_{i-1}} \int_{t_{j-1}}^{\tau} \big((t_{i-1}-u)^{\alpha-1} - (s-u)^{\alpha-1}\big)(\tau - v)^{\alpha-1}(v-u)^{2H-2}\,\mathrm{d}v\,\mathrm{d}u \\
&\le C \int_0^{t_{i-1}} \int_{t_{j-1}}^{\tau} \big((t_{i-1}-u)^{\alpha-1} - (s-u)^{\alpha-1}\big)(\tau - v)^{\alpha-1}\,\mathrm{d}v\,\mathrm{d}u(\tau - s)^{2H-2} \\
&\le Ch^{2\alpha}(\tau - s)^{2H-2}.
\end{aligned}$$

Estimate of $\mathcal{J}_3$. To facilitate the estimation of $\mathcal{J}_3$, we note that

$$\begin{aligned}
\mathcal{J}_3 &= \int_{t_{i-1}}^{s} \int_0^{t_{j-1}} (s-u)^{\alpha-1}\big((t_{j-1}-v)^{\alpha-1} - (\tau - v)^{\alpha-1}\big)|v-u|^{2H-2}\,\mathrm{d}v\,\mathrm{d}u \\
&= \int_{t_{i-1}}^{s} \int_0^{t_{i-1}} (s-u)^{\alpha-1}\big((t_{j-1}-v)^{\alpha-1} - (\tau - v)^{\alpha-1}\big)(u-v)^{2H-2}\,\mathrm{d}v\,\mathrm{d}u \\
&\quad + \int_{t_{i-1}}^{s} \int_{t_{i-1}}^{t_{j-1}} (s-u)^{\alpha-1}\big((t_{j-1}-v)^{\alpha-1} - (\tau - v)^{\alpha-1}\big)|v-u|^{2H-2}\,\mathrm{d}v\,\mathrm{d}u \\
&=: \mathcal{J}_{3,1} + \mathcal{J}_{3,2}.
\end{aligned}$$

It follows from (4.19) and (3.11) that

$$\mathcal{J}_{3,1} \le Ch \int_{t_{i-1}}^{s} \int_0^{t_{i-1}} (s-u)^{\alpha-1}(u-v)^{2H-2}(t_{j-1}-v)^{\alpha-2H}\,\mathrm{d}v\,\mathrm{d}u(\tau - s)^{2H-2} \le Ch^{2\alpha}(\tau - s)^{2H-2}.$$

The estimate of $\mathcal{J}_{3,2}$ can be divided into two cases: $i < j \le i+2$ and $i+3 \le j \le N$. If $i < j \le i+2$, then similar to (3.20), we have

$$\mathcal{J}_{3,2} \le \int_{t_{i-1}}^{s} \int_{t_{i-1}}^{t_{j-1}} (s-u)^{\alpha-1}(t_{j-1}-v)^{\alpha-1}|v-u|^{2H-2}\,\mathrm{d}v\,\mathrm{d}u \le Ch^{2(\alpha+H-1)} \le Ch^{2\alpha}(\tau - s)^{2H-2}, \tag{4.23}$$

where the last step used $\tau - s \le t_j - t_{i-1} \le 3h$. If $i+3 \le j \le N$, we apply (4.23) to obtain

$$\begin{aligned}
&\int_{t_{i-1}}^{s} \int_{t_{i-1}}^{t_i} (s-u)^{\alpha-1}\big((t_{j-1}-v)^{\alpha-1} - (\tau - v)^{\alpha-1}\big)|v-u|^{2H-2}\,\mathrm{d}v\,\mathrm{d}u \\
&\qquad \le \int_{t_{i-1}}^{s} \int_{t_{i-1}}^{t_i} (s-u)^{\alpha-1}(t_i - v)^{\alpha-1}|v-u|^{2H-2}\,\mathrm{d}v\,\mathrm{d}u
\end{aligned}$$

$$\leq Ch^{2\alpha}(\tau - s)^{2H-2},$$

and utilize Lemma 3.3 with $k = i+1$ and $l = j$ to get

$$\begin{aligned}\int_{t_{i-1}}^{s}\int_{t_i}^{t_{j-1}} (s-u)^{\alpha-1}\big((t_{j-1}-v)^{\alpha-1} - (\tau - v)^{\alpha-1}\big)|v-u|^{2H-2}\,\mathrm{d}v\,\mathrm{d}u \\ \leq Ch^{\alpha}(\tau-s)^{2H-2}\int_{t_{i-1}}^{s}(s-u)^{\alpha-1}\,\mathrm{d}u \\ \leq Ch^{2\alpha}(\tau-s)^{2H-2}.\end{aligned}$$

These imply $\mathcal{J}_{32} \leq Ch^{2\alpha}(\tau-s)^{2H-2}$ for the case $i+3 \leq j \leq N$. Hence, we have $\mathcal{J}_3 \leq Ch^{2\alpha}(\tau-s)^{2H-2}$.

Estimate of $\mathcal{J}_4$. It can be divided into two cases: $j = i+1$ and $i+2 \leq j \leq N$. When $j = i+1$, we have $\tau - s \leq Ch$ and then

$$\begin{aligned}\mathcal{J}_4 &= \int_{t_{i-1}}^{s}\int_{t_i}^{\tau}(s-u)^{\alpha-1}(\tau-v)^{\alpha-1}(v-u)^{2H-2}\,\mathrm{d}v\,\mathrm{d}u \\ &\leq \int_{t_{i-1}}^{s}\int_{s}^{\tau}(s-u)^{\alpha-1}(\tau-v)^{\alpha-1}(v-s)^{2H-2}\,\mathrm{d}v\,\mathrm{d}u \\ &\leq C(\tau-s)^{\alpha+2H-2}\int_{t_{i-1}}^{s}(s-u)^{\alpha-1}\,\mathrm{d}u \\ &\leq Ch^{2\alpha}(\tau-s)^{2H-2}.\end{aligned}$$

When $i+2 \leq j \leq N$, for $v \in [t_{j-1}, \tau]$ and $u \in [t_{i-1}, s]$, we have

$$\tau - s \leq (j-i+1)h \leq 3(j-i-1)h \leq 3(v-u),$$

and thus,

$$\mathcal{J}_4 \leq C\int_{t_{i-1}}^{s}\int_{t_{j-1}}^{\tau}(s-u)^{\alpha-1}(\tau-v)^{\alpha-1}\,\mathrm{d}v\,\mathrm{d}u(\tau-s)^{2H-2} \leq Ch^{2\alpha}(\tau-s)^{2H-2}.$$

Finally, gathering the above estimates of $\mathcal{J}_1$, $\mathcal{J}_2$, $\mathcal{J}_3$ and $\mathcal{J}_4$ together completes the proof. ☐

Lemma 4.2. *For $\mathcal{K}$ and $\mathcal{L}$ respectively given by (4.17) and (4.18), there exists some constant C such that*

$$\mathcal{K} + \mathcal{L} \leq C\mathcal{R}^2_{H,\alpha}(h),$$

for any $s \in (t_{i-1}, t_i]$, $\tau \in (t_{j-1}, t_j]$ with $1 \leq i < j \leq N$, and θ, $\lambda \in (0,1)$.

Proof. According to Lemma 3.5, one has that for any $\tau \in (t_{j-1}, t_j]$ with $1 \leq j \leq N$,

$$\sup_{\mu\in[0,T]}\int_0^{\mu}\int_0^{T}(\mu-r)^{\alpha-1}A(v;\tau)|r-v|^{2H-2}\,\mathrm{d}v\,\mathrm{d}r \leq C\mathcal{R}_{H,\alpha}(h). \tag{4.24}$$

Then, taking (3.2) and (4.24) into account shows

$$\begin{aligned}\mathcal{K} \leq C &\sup_{\mu_1\in[0,T]}\int_0^{\mu_1}\int_0^{T}(\mu_1-r_1)^{\alpha-1}A(u;s)|r_1-u|^{2H-2}\,\mathrm{d}u\,\mathrm{d}r_1 \\ &\times \sup_{\mu_2\in[0,T]}\int_0^{\mu_2}\int_0^{T}(\mu_2-r_2)^{\alpha-1}A(v;\tau)|r_2-v|^{2H-2}\,\mathrm{d}v\,\mathrm{d}r_2 \leq C\mathcal{R}^2_{H,\alpha}(h).\end{aligned}$$

TABLE 1. The mean square error and convergence order of the Euler method (2.2) and the fast Euler method (2.6) with $\epsilon = 10^{-3}$ for equation (5.1) with different Hurst indices $H = 0.6, 0.7, 0.8, 0.9$.

Stepsize	$H = 0.6$		$H = 0.7$		$H = 0.8$		$H = 0.9$	
	Euler	fast Euler	Euler	fast Euler	Euler	fast Euler	Euler	fast Euler
2^{-5}	2.5137e–02	2.5132e–02	4.4350e–02	4.4344e–02	9.3196e–02	9.3192e–02	2.1048e–01	2.1048e–01
2^{-6}	1.4001e–02	1.4000e–02	2.7913e–02	2.7908e–02	6.6820e–02	6.6817e–02	1.8097e–01	1.8097e–01
2^{-7}	7.8632e–03	7.8674e–03	1.7398e–02	1.7396e–02	4.7490e–02	4.7487e–02	1.5594e–01	1.5594e–01
2^{-8}	4.4328e–03	4.4383e–03	1.0449e–02	1.0449e–02	3.3728e–02	3.3725e–02	1.3369e–01	1.3369e–01
Order	0.8343	0.8336	0.6939	0.6938	0.4892	0.4892	0.2179	0.2179
residual	0.0061	0.0064	0.0242	0.0241	0.0051	0.0051	0.0022	0.0022
Theoretical order	0.8	0.8	0.6	0.6	0.4	0.4	0.2	0.2

TABLE 2. The CPU time(s) of the Euler method (2.2) and the fast Euler method (2.6) with different tolerances $\epsilon = 10^{-3}, 10^{-6}, 10^{-9}, 10^{-12}, 10^{-15}$ for equation (5.1) with $H = 0.6$.

Stepsize	Euler method	fast Euler method				
		$\epsilon = 10^{-3}$	$\epsilon = 10^{-6}$	$\epsilon = 10^{-9}$	$\epsilon = 10^{-12}$	$\epsilon = 10^{-15}$
2^{-5}	0.275227	0.043581	0.092371	0.114639	0.137905	0.153910
2^{-6}	1.064096	0.132072	0.188943	0.236102	0.283379	0.315494
2^{-7}	4.316762	0.266233	0.378769	0.497053	0.583575	0.686940
2^{-8}	16.687388	0.565725	0.804654	1.033466	1.241114	1.469896
2^{-9}	65.982834	1.221633	1.756703	2.190837	2.616509	3.083545
2^{-10}	258.604845	2.600641	3.799821	4.586589	5.616637	6.513161

Besides, equations (3.3) and (4.24) display

$$\begin{aligned}\mathcal{L} \le C \sup_{\mu\in[0,T]} \int_0^{\mu} \int_{[0,T]^4} (\mu-\eta)^{\alpha-1}(\eta-r_1)^{\alpha-1}(\eta-r_2)^{\alpha-1}\mathbf{1}_{[0,\eta)}(r_1\vee r_2) \\ \times A(u;s)A(v;\tau)|r_1-u|^{2H-2}|r_2-v|^{2H-2}\,\mathrm{d}r_1\,\mathrm{d}u\,\mathrm{d}r_2\,\mathrm{d}v\,\mathrm{d}\eta \\ \le C\mathcal{R}^2_{H,\alpha}(h) \sup_{\mu\in[0,T]} \int_0^{\mu} (\mu-\eta)^{\alpha-1}\,\mathrm{d}\eta \le C\mathcal{R}^2_{H,\alpha}(h).\end{aligned}$$

The proof is completed. □

5. NUMERICAL EXPERIMENTS

In this section, we provide numerical experiments to verify the theoretical results. Consider the following nonlinear equation

$$x(t) = 1 + \frac{1}{\Gamma(\alpha)}\int_0^t (t-s)^{\alpha-1}\frac{\cos(x^3(s))}{1+x^2(s)}\,\mathrm{d}s + \frac{\sigma}{\Gamma(\alpha)}\int_0^t (t-s)^{\alpha-1}\,\mathrm{d}W_H(s), \quad t\in[0,1], \tag{5.1}$$

where $\alpha = 2-2H$ and $\sigma = \sqrt{2/\Gamma(2H+1)}$, so the fluctuation-dissipation theorem in physics rigorously holds (see *e.g.*, [18]). Note that the drift function $b(x) = \cos(x^3)/(1+x^2)$ belongs to $C^{1,3}_{b,p}$, which indicates that the assumption of Theorems 2.1–2.3 holds.

TABLE 3. The mean square error and CPU time(s) of the standard Monte Carlo (MC) estimator combined with the fast Euler method (2.6) and the MLMC estimator (2.8) with $M = 4$ for the evaluation of $\mathbb{E}\big[f(x(1))\big]$ with $H = 0.65, 0.75, 0.85$ and $f(x) = |x|, x^2$.

H	Estimator	$f(x) = \|x\|$		$f(x) = x^2$	
		Error	CPU time(s)	Error	CPU time(s)
0.65	MC ($N = 4$)	0.026307	1.607621	0.132467	1.596519
	MC ($N = 16$)	0.005913	2.751750	0.030681	2.732153
	MC ($N = 64$)	0.003326	7.877532	0.013319	7.879020
	MLMC ($\varepsilon = 0.08$)	0.046910	0.306292	0.050134	1.342879
	MLMC ($\varepsilon = 0.05$)	0.033677	0.466086	0.032200	1.724913
	MLMC ($\varepsilon = 0.01$)	0.006697	2.122597	0.007652	21.38446
0.75	MC ($N = 4$)	0.034087	1.455264	0.163116	1.474097
	MC ($N = 16$)	0.011039	2.216749	0.051972	2.228000
	MC ($N = 64$)	0.005364	5.856083	0.020095	5.867424
	MLMC ($\varepsilon = 0.08$)	0.049924	0.190690	0.052996	1.045602
	MLMC ($\varepsilon = 0.05$)	0.039557	0.348829	0.035100	1.742142
	MLMC ($\varepsilon = 0.01$)	0.006067	2.594411	0.009111	36.30176
0.85	MC ($N = 4$)	0.041153	1.565133	0.206330	1.544437
	MC ($N = 16$)	0.017986	2.738336	0.093683	2.692867
	MC ($N = 64$)	0.008507	7.883607	0.045557	7.922303
	MLMC ($\varepsilon = 0.08$)	0.064472	0.718679	0.052700	3.193825
	MLMC ($\varepsilon = 0.05$)	0.031722	1.009783	0.033251	8.184939
	MLMC ($\varepsilon = 0.01$)	0.006871	13.74156	0.009602	130.7832

In all numerical tests, we use the Euler solution obtained by (2.2) with small stepsize $h = 2^{-12}$ to replace the unknown exact solution. In addition, we use 10 000 sample paths for the simulation of expectations. We perform the following three experiments:

(1) Applying the Euler method (2.2) and the fast Euler method (2.6) with tolerance $\epsilon = 10^{-3}$ to equation (5.1) with different Hurst indices $H = 0.6, 0.7, 0.8, 0.9$, the computational mean square error and convergence order are listed in Table 1, where the convergence order is measured by the least squares fit (see *e.g.*, [8]). As shown in Table 1, the numerical results are consistent with Theorems 2.1 and 2.2.

(2) Using the Euler method (2.2) and the fast Euler method (2.6) with different tolerances $\epsilon = 10^{-3}$, 10^{-6}, 10^{-9}, 10^{-12}, 10^{-15} to equation (5.1) with Hurst index $H = 0.6$, the corresponding CPU time is displayed in Table 2. This result shows that the fast Euler method (2.6) is more computationally efficient than the Euler method (2.2).

(3) Utilizing the standard Monte Carlo estimator combined with the fast Euler method (2.6) and the MLMC estimator (2.8) with $M = 4$ to evaluate $\mathbb{E}\big[f(x(1))\big]$ for different Hurst indices $H = 0.65, 0.75, 0.85$ and different test functions $f(x) = |x|$, x^2, the corresponding mean square error and CPU time are reported in Table 3. This result illustrates that the MLMC estimator has better performance than the Monte Carlo estimator in the simulation of $\mathbb{E}\big[f(x(1))\big]$.

Acknowledgements. Deep thanks go to the editor and anonymous referees for their insightful comments and helpful suggestions, which are crucial to the improvements of the manuscript. The authors are grateful to Professor Lei Li for valuable discussions. This work is supported by National key R&D Program of China under Grant No. 2020YFA0713701, National Natural Science Foundation of China (Nos. 11971470, 11871068, 12031020, 12022118), and China Postdoctoral Science Foundation (No. 2022M713313).

References

[1] X. Dai and A. Xiao, Lévy-driven stochastic Volterra integral equations with doubly singular kernels: existence, uniqueness, and a fast EM method. *Adv. Comput. Math.* **46** (2020) 29.

[2] X. Dai and A. Xiao, A note on Euler method for the overdamped generalized Langevin equation with fractional noise. *Appl. Math. Lett.* **111** (2021) 106669.

[3] R.C. Dalang, Extending the martingale measure stochastic integral with applications to spatially homogeneous S.P.D.E.'s. *Electron. J. Probab.* **4** (1999) 1–29.

[4] G. Didier and H. Nguyen, Asymptotic analysis of the mean squared displacement under fractional memory kernels. *SIAM J. Math. Anal.* **52** (2020) 3818–3842.

[5] D. Fang and L. Li, Numerical approximation and fast evaluation of the overdamped generalized Langevin equation with fractional noise. *ESAIM Math. Model. Numer. Anal.* **54** (2020) 431–463.

[6] M.B. Giles, Multilevel Monte Carlo path simulation. *Oper. Res.* **56** (2008) 607–617.

[7] M.B. Giles, Multilevel Monte Carlo methods. *Acta Numer.* **24** (2015) 259–328.

[8] D.J. Higham, An algorithmic introduction to numerical simulation of stochastic differential equations. *SIAM Rev.* **43** (2001) 525–546.

[9] D.J. Higham, X. Mao and A.M. Stuart, Strong convergence of Euler-type methods for nonlinear stochastic differential equations. *SIAM J. Numer. Anal.* **40** (2002) 1041–1063.

[10] J. Hong, C. Huang, M. Kamrani and X. Wang, Optimal strong convergence rate of a backward Euler type scheme for the Cox–Ingersoll–Ross model driven by fractional Brownian motion. *Stochastic Process. Appl.* **130** (2020) 2675–2692.

[11] M. Hutzenthaler, A. Jentzen and P.E. Kloeden, Strong and weak divergence in finite time of Euler's method for stochastic differential equations with non-globally Lipschitz continuous coefficients. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.* **467** (2011) 1563–1576.

[12] S. Jiang, J. Zhang, Q. Zhang and Z. Zhang, Fast evaluation of the Caputo fractional derivative and its applications to fractional diffusion equations. *Commun. Comput. Phys.* **21** (2017) 650–678.

[13] P.E. Kloeden, A. Neuenkirch and R. Pavani, Multilevel Monte Carlo for stochastic differential equations with additive fractional noise. *Ann. Oper. Res.* **189** (2011) 255–276.

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

[15] S.C. Kou and X. Sunney Xie, Generalized Langevin equation with fractional Gaussian noise: subdiffusion within a single protein molecule. *Phys. Rev. Lett.* **93** (2004) 180603.

[16] R. Kubo, The fluctuation–dissipation theorem. *Rep. Prog. Phys.* **29** (1966) 255–284.

[17] L. Li and J.-G. Liu, A discretization of Caputo derivatives with application to time fractional SDEs and gradient flows. *SIAM J. Numer. Anal.* **57** (2019) 2095–2120.

[18] L. Li, J.-G. Liu and J. Lu, Fractional stochastic differential equations satisfying fluctuation-dissipation theorem. *J. Stat. Phys.* **169** (2017) 316–339.

[19] X. Mao, Stochastic Differential Equations and Applications, 2nd edition. Horwood Publishing Limited, Chichester (2008).

[20] S.A. McKinley and H.D. Nguyen, Anomalous diffusion and the generalized Langevin equation. *SIAM J. Math. Anal.* **50** (2018) 5119–5160.

[21] H. Mori, Transport, collective motion, and Brownian motion. *Progr. Theoret. Phys.* **33** (1965) 423–455.

[22] D. Nualart, The Malliavin Calculus and Related Topics, 2nd edition. Springer-Verlag, Berlin (2006).

[23] V. Pipiras and M.S. Taqqu, Integration questions related to fractional Brownian motion. *Probab. Theory Relat. Fields* **118** (2000) 251–291.

[24] A. Richard, X. Tan and F. Yang, Discrete-time simulation of stochastic Volterra equations. *Stochastic Process. Appl.* **141** (2021) 109–138

[25] M. Sanz-Solé, Malliavin Calculus with Applications to Stochastic Partial Differential Equations. EPFL Press, distributed by CRC Press (2005).

[26] H. Ye, J. Gao and Y. Ding, A generalized Gronwall inequality and its application to a fractional differential equation. *J. Math. Anal. Appl.* **328** (2007) 1075–1081.