IMA Journal of Numerical Analysis (2025) **00**, 1–44
https://doi.org/10.1093/imanum/draf034

Perturbation estimates for order-one strong approximations of SDEs without globally monotone coefficients

LEI DAI AND XIAOJIE WANG*
School of Mathematics and Statistics, HNP-LAMA, Central South University, Changsha, China
*Corresponding author. x.j.wang7@csu.edu.cn AQ6

[Received on 18 June 2024; revised on 19 February 2025; accepted on 19 March 2025]

To obtain strong convergence rates of numerical schemes, an overwhelming majority of existing works impose a global monotonicity condition on coefficients of stochastic differential equations (SDEs). Nevertheless, there are still many SDEs from applications that do not have globally monotone coefficients. As a recent breakthrough, the authors of (Hutzenthaler and Jentzen 2020, On a perturbation theory and on strong convergence rates for stochastic ordinary and partial differential equations with nonglobally monotone coefficients. *Ann. Prob.*, **48**, 53–93) originally presented a perturbation theory for SDEs, which is crucial to recovering strong convergence rates of numerical schemes in a non-globally monotone setting. However, only a convergence rate of order $1/2$ was obtained there for time-stepping schemes such as a stopped increment-tamed Euler–Maruyama (SITEM) method. An interesting question arises, also raised by the aforementioned work, as to whether a higher convergence rate than $1/2$ can be obtained when higher order schemes are used. The present work attempts to give a positive answer to this question. To this end, we develop some new perturbation estimates that are able to reveal the order-one strong convergence of numerical methods. As the first application of the newly developed estimates, we identify the expected order-one pathwise uniformly strong convergence of the SITEM method for additive noise driven SDEs and multiplicative noise driven second-order SDEs with non-globally monotone coefficients. As the other application, we propose and analyze a positivity preserving explicit Milstein-type method for Lotka–Volterra competition model driven by multi-dimensional noise, with a pathwise uniformly strong convergence rate of order one recovered under mild assumptions. These obtained results are completely new and significantly improve the existing theory. Numerical experiments are also provided to confirm the theoretical findings.

Keywords: SDEs with non-globally monotone coefficients; explicit method; exponential integrability properties; pathwise uniformly strong convergence; order-one strong convergence; stochastic Lotka–Volterra competition model.

1. Introduction

In order to describe the time evolution of many dynamical processes under random environmental effects, stochastic differential equations (SDEs)

$$X_t = X_0 + \int_0^t f(X_s)\,\mathrm{d}s + \int_0^t g(X_s)\,\mathrm{d}W_s,\ t \in [0, T] \tag{1.1}$$

are widely used in various science and engineering fields such as finance, chemistry, physics and biology. In practice, the closed-form solutions of non-linear SDEs are rarely available and one usually falls back on their numerical approximations. For SDEs possessing globally Lipschitz coefficients, the monographs Kloeden & Platen (1992); Milstein & Tretyakov (2004) established a fundamental framework to analyze a batch of numerical schemes including typical methods such as the explicit Euler–Maruyama (EM) method and explicit Milstein method. Nevertheless, a majority of SDEs arising from applications have superlinearly growing coefficients and the globally Lipschitz condition is violated. A natural question thus arises as to whether the traditional numerical methods designed in the globally Lipschitz setting are still able to perform well when used to solve SDEs with superlinearly growing coefficients. Unfortunately, the authors of Hutzenthaler *et al.* (2011) gave a negative answer, by showing that the popularly used EM method is divergent in the sense of both strong and weak convergence, when used to solve a large class of SDEs with superlinearly growing coefficients. Therefore, special care must be taken to construct and analyze convergent numerical schemes in the absence of the Lipschitz regularity of coefficients. In recent years, a prospering growth of relevant works is devoted to numerical approximations of SDEs with non-globally Lipschitz coefficients. Roughly speaking, people either rely on implicit Euler/Milstein schemes Higham (2000); Higham *et al.* (2002); Alfonsi (2013); Mao & Szpruch (2013); Neuenkirch & Szpruch (2014); Andersson & Kruse (2017); Zong *et al.* (2018); Wang *et al.* (2020); Wang (2023) or some explicit schemes based on modifications of the traditional explicit EM/Milstein methods Li & Mao (2009); Hutzenthaler *et al.* (2012); Liu & Mao (2013); Tretyakov & Zhang (2013); Wang & Gan (2013); Hutzenthaler & Jentzen (2015); Mao (2015); Sabanis (2016); Beyn *et al.* (2017); Kumar & Sabanis (2019); Fang & Giles (2020); Brehier (2023); Kelly *et al.* (2023) for SDEs with superlinearly growing coefficients. To get the desired convergence rates of numerical schemes, a frequently used argument is based on Gronwall's lemma together with the popular global monotonicity condition, for all $x, y \in \mathbb{R}^d$,

$$\langle x - y, f(x) - f(y)\rangle + \tfrac{q}{2}\|g(x) - g(y)\|^2 \leq K|x - y|^2, \tag{1.2}$$

where f and g are the drift and diffusion coefficients of SDEs (1.1), respectively, and K is a positive constant independent of x, y. Indeed, an overwhelming majority of existing works on convergence rates carry out the error analysis under the global monotonicity condition (1.2).

However, such a condition is still restrictive and many momentous SDEs from applications fail to obey (1.2). Examples include stochastic van der Pol oscillator, stochastic Lorenz equation, stochastic Langevin dynamics and stochastic Lotka–Volterra (LV) competition model (see, e.g., Mao (2007); Hutzenthaler & Jentzen (2015)). What if we did not have the condition (1.2) available? In fact, the analysis of the convergence rates of numerical schemes without the global monotonicity condition turns out to be highly non-trivial (see Hutzenthaler & Jentzen (2015, 2020)). As a recent breakthrough, Hutzenthaler and Jentzen in Hutzenthaler & Jentzen (2020) made significant progress in this direction and originally developed a framework known as perturbation theory for SDEs beyond the global monotonicity assumption (1.2). This theory, combined with exponential integrability properties of both numerical solutions and exact solutions (see Hutzenthaler *et al.* (2018); Cox *et al.* (2024)) enables one to reveal strong convergence rates of numerical schemes in a non-globally monotone setting. Following this argument, the authors of Hutzenthaler & Jentzen (2020) analyzed the pointwise strong error

$$\sup_{t\in[0,T]} \left\|X_t - Y_t\right\|_{L^r(\Omega;\mathbb{R}^d)} \tag{1.3}$$

of an explicit stopped increment-tamed EM (SITEM) method $\{Y_t\}_{t\in[0,T]}$ proposed by Hutzenthaler *et al.* (2018) (cf. (4.1) therein), which was shown there to inherit exponential integrability properties of SDEs. Successfully, the authors identified the pointwise strong convergence rate of order $\frac{1}{2}$ for the SITEM method. An interesting question arises as to whether a higher convergence rate than order $\frac{1}{2}$ can be obtained when the considered SDEs are driven by additive noise or when high-order (Milstein-type) schemes are used, which is also expected by Hutzenthaler & Jentzen (2020) (see Remark 3.1 therein). Unfortunately, following (Hutzenthaler & Jentzen, 2020, Theorem 1.2), the convergence rates of any schemes would not exceed order $\frac{1}{2}$, which is nothing but the order of the Hölder regularity of the approximation process.

In the present article, we attempt to present some new perturbation estimates that can be used to reveal the order-one pathwise uniformly strong convergence of numerical methods for several SDEs with non-globally monotone coefficients (see Theorem 3.2). Different from Hutzenthaler & Jentzen (2020), we use the Itô formula to expand the difference for the drift term, i.e., $f(Y_s) - a(s)$ in Lemma 3.1 and rely on Burkholder–Davis–Gundy type inequalities to carefully treat the related terms (see estimates of $\mathbb{S}_2$ and $\mathbb{T}_2$ in Theorem 3.2). This approach essentially enables one to attain order-one strong convergence for the error analysis of numerical schemes.

As the first application of the newly developed estimates, we identify the expected order-one pathwise uniformly strong convergence of the SITEM method for some SDEs with nonglobally monotone coefficients (see Theorem 4.2 and subsequent example models), including the additive noise driven SDEs (e.g., the stochastic Lorenz equation with additive noise, Brownian dynamics and Langevin dynamics) and multiplicative noise driven second-order SDEs (i.e., second-order ordinary differential equations perturbed by multiplicative white noise) such as the stochastic van der Pol oscillator and stochastic Duffing–van der Pol oscillator:

$$\Big\| \sup_{t\in[0,T]} |X_t - Y_t| \Big\|_{L^r(\Omega;\mathbb{R})} \le Ch. \tag{1.4}$$

Here $\{Y_t\}_{t\in[0,T]}$ is produced by the SITEM method (4.1), $r > 0$ is an arbitrary constant and $h > 0$ is the uniform stepsize. These findings thus fill the gap left by Hutzenthaler & Jentzen (2020) and also significantly improve relevant convergence results in Hutzenthaler & Jentzen (2020), where the pointwise strong convergence rate of only order $\frac{1}{2}$ was obtained for the SITEM method applied to these models.

As the other application, we propose and analyze a positivity preserving explicit Milstein-type method (5.4) for stochastic LV competition model driven by multi-dimensional noises, with a pathwise uniformly strong convergence rate of order one recovered (Theorem 5.6). To the best of our knowledge, this is the first paper to obtain the order-one pathwise uniformly strong convergence of an explicit positivity preserving scheme for the general LV competition model.

The paper is structured as follows. In the next section, we introduce some notations and inequalities that may be used later. In Section 3, we present new perturbation estimates for SDEs beyond the global monotonicity assumption. Equipped with these estimates, we derive the order-one strong convergence of the SITEM method for some additive noise driven or second-order SDE models. In Section 5, we propose and analyze an explicit Milstein method for the LV competition model with multi-dimensional noise. Some numerical experiments are provided to confirm the theoretical findings and a short conclusion is made in Section 6.

2. Preliminaries

2.1 *Notations*

Throughout this paper, unless otherwise specified, the following notations are used. Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\ge 0}, \mathbb{P})$ be a complete probability space with a filtration $\{\mathcal{F}_t\}_{t\ge 0}$ fulfilling the usual conditions, that is, the filtration is right continuous and increasing, and $\mathcal{F}_0$ contains all $\mathbb{P}$-null sets. Let $\{W_t\}_{t\geq 0}$ be an m-dimensional standard Brownian motion defined on $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$. For $a \in \mathbb{R}$, we define $\frac{a}{\infty} := 0$ and for $a \in \mathbb{R}\backslash\{0\}$, $\frac{a}{0} := \infty$. For a fixed integer number $d \geq 1$ and a vector $x \in \mathbb{R}^d$, $x^{(i)}, i = 1, ..., d$ denotes the ith component of x and $|x|$ denotes the Euclidean norm induced by the vector inner product $\langle\cdot,\cdot\rangle$. For a matrix $A \in \mathbb{R}^{d\times m}, d, m \in \mathbb{N}$, $A^{(i)}, i = 1, ..., m$ denotes the ith column of A and $A^{(ij)}, i = 1, ..., d, j = 1, ..., m$ represents the element at ith row and jth column of A. Let A^* be the transpose of A and $\|A\| := \sqrt{\text{trace}(A^*A)}$ be the Hilbert–Schmidt norm induced by Hilbert–Schmidt inner product $\langle\cdot,\cdot\rangle_{HS}$. For a random variable $\xi : \Omega \to \mathbb{H}$, where $\mathbb{H}$ is a separable Banach space endowed with norm $\|\cdot\|_{\mathbb{H}}$, $\mathbb{E}[\xi]$ denotes its expectation and for any $r > 0$, $\|\xi\|_{L^r(\Omega;\mathbb{H})} := (\mathbb{E}[\|\xi\|^r_{\mathbb{H}}])^{1/r}$. For $f = (f^{(1)}, ..., f^{(d)})^* \in C^2(\mathbb{R}^d, \mathbb{R}^d)$, we use $f'(x)$ to denote the Jacobian matrix of $f(x)$, in which the ith row is $f^{(i)\prime}(x) := \left(\nabla f^{(i)}(x)\right)^* : \mathbb{R}^d \to \mathbb{R}^{1\times d}$. The notation $\text{Hess}_x\left(f(x)\right)$ is a generalized Hessian matrix including d components where the ith is the Hessian matrix $\text{Hess}_x\left(f^{(i)}(x)\right) : \mathbb{R}^d \to \mathbb{R}^{d\times d}$ of $f^{(i)}(x)$. Further, for $f : \mathbb{R}^d \to \mathbb{R}^d, g : \mathbb{R}^d \to \mathbb{R}^{d\times m}$ and $U \in C^2(\mathbb{R}^d, \mathbb{R})$, we denote

$$(\mathcal{A}_{f,g}U)(x) := U'(x)f(x) + \tfrac{1}{2}\,\text{trace}\left(g(x)g(x)^*\,\text{Hess}_x\left(U(x)\right)\right). \tag{2.1}$$

Let $T \in (0, \infty), N \in \mathbb{Z}^+$, let a uniform mesh

$$0 = t_0 < t_1 < \cdots < t_N = T \tag{2.2}$$

be constructed with the time step $h = T/N$, and for $s \in [0, T]$, define

$$\lfloor s\rfloor_N := \sup_{n=0,...,N}\{t_n : t_n \leq s\}.$$

Moreover, we introduce a class of functions $\mathcal{C}^3_{\mathcal{D}}(\mathbb{R}^d, \mathbb{R})$ as follows:

$$\mathcal{C}^3_{\mathcal{D}}(\mathbb{R}^d, \mathbb{R}) := \left\{ \Lambda \in C^2\left(\mathbb{R}^d, \mathbb{R}\right) : \begin{array}{l} \text{Every element of } \text{Hess}_x\left(\Lambda(x)\right) \text{ is locally Lipschitz} \\ \text{continuous and for } i \in \{1, 2, 3\}, a.s. \text{ (Lebesgue} \\ \text{measure) } x \in \mathbb{R}^d, \text{there exist } p, c \geq 3 \text{ such that} \\ \|\Lambda^{[i]}(x)\|_{L^{[i]}(\mathbb{R}^d,\mathbb{R})} \leq c\left(1 + |\Lambda(x)|\right)^{1-i/p} \end{array} \right\}. \tag{2.3}$$

Here for $i = 1, 2, 3$, we denote

$$\|\Lambda^{[i]}(x)\|_{L^{[i]}(\mathbb{R}^d,\mathbb{R})} := \sup_{v_1,...,v_i\in\mathbb{R}^d\backslash\{0\}} \frac{|\Lambda^{[i]}(x)(v_1, ..., v_i)|}{|v_1|\cdots|v_i|}, \tag{2.4}$$

where

$$\Lambda^{[i]}(x)(v_1, \ldots, v_i) = \sum_{l_1,...,l_i=1}^{d} \left(\tfrac{\partial^i \Lambda}{\partial x_{l_1}\ldots\partial x_{l_i}}\right)(x)\cdot v_1^{(l_1)}\cdot v_2^{(l_2)}\cdot\ldots\cdot v_i^{(l_i)}. \tag{2.5}$$

Note that $\mathcal{C}^3_{\mathcal{D}}(\mathbb{R}^d, \mathbb{R})$ forms a linear space containing a batch of functions such as

$$\Lambda(x) = \left(\sum_{i=1}^{d} x_i^{2c_i}\right)^r,\ c_i \geq 1,\ r \geq 1.$$

For a metric space (E,ρ), we say $\Lambda \in \mathcal{C}^1_{\mathcal{P}}(\mathbb{R}^d, E)$ with constants K_Λ, c_Λ, if $\Lambda \in C\big(\mathbb{R}^d, E\big)$ and there exist constants $K_\Lambda, c_\Lambda \geq 0$ such that

$$\rho(\Lambda(x),\Lambda(y)) \leq K_\Lambda(1+|x|+|y|)^{c_\Lambda}|x-y| \tag{2.6}$$

holds for all $x, y \in \mathbb{R}^d$. One can easily see that if $\Lambda : \mathbb{R}^d \to \mathbb{R}$ is differentiable and $\Lambda \in \mathcal{C}^1_{\mathcal{P}}(\mathbb{R}^d,\mathbb{R})$ with constants K_Λ, c_Λ, then there exists some constant K_1 such that for all $x \in \mathbb{R}^d$,

$$|\Lambda(x)| \leq K_1(1+|x|)^{c_\Lambda+1}, |\Lambda'(x)| \leq K_1(1+|x|)^{c_\Lambda}.$$

Finally, we use C (resp. C with some subscripts) to denote a generic positive constant independent of the time step (resp. dependent on the subscripts), which may differ from one place to another.

2.2 *Burkholder–Davis–Gundy type inequalities*

In what follows we recall two Burkholder–Davis–Gundy type inequalities, which are frequently used in the subsequent analysis.

LEMMA 2.1. Let $S : [0,T]\times\Omega \to \mathbb{R}^{d\times m}$ be a predictable stochastic process satisfying $\mathbb{P}(\int_0^T \|S_t\|^2 \, \mathrm{d}t < \infty) = 1$ and let $\{W_t\}_{t\geq 0}$ be a m-dimensional standard Brownian motion. Then for any $p \geq 2$,

$$\Big\| \sup_{t\in[0,T]} \Big| \int_0^t S_r \, \mathrm{d}W_r \Big| \Big\|_{L^p(\Omega;\mathbb{R})} \leq p \left(\int_0^T \sum_{i=1}^m \|S_r^{(i)}\|^2_{L^p(\Omega;\mathbb{R}^d)} \, \mathrm{d}r \right)^{1/2}. \tag{2.7}$$

LEMMA 2.2. Let $M \in \mathbb{N}$ and $S_1, ..., S_M : \Omega \to \mathbb{R}$ be random variables satisfying $\sup_{i\in\{1,...,M\}} \|S_i\|_{L^2(\Omega;\mathbb{R})} < \infty$ and for any $i \in \{1, ..., M-1\}$, $\mathbb{E}[S_{i+1}|S_1, ..., S_i] = 0$. Then for any $p \geq 2$, there exists some positive constant C_p such that

$$\|S_1 + \cdots + S_M\|_{L^p(\Omega;\mathbb{R})} \leq C_p \left(\|S_1\|^2_{L^p(\Omega;\mathbb{R})} + \cdots + \|S_M\|^2_{L^p(\Omega;\mathbb{R})} \right)^{1/2}. \tag{2.8}$$

The first lemma can be found in (Wang & Gan, 2013, Lemma 2.7) and the other one is quoted from (Hutzenthaler & Jentzen, 2011, Lemma 4.1).

3. New perturbation estimates for SDEs

In this section, let us focus on the following SDEs of Itô type:

$$\begin{cases} X_t - X_0 = \int_0^t f(X_s)\,\mathrm{d}s + \int_0^t g(X_s)\,\mathrm{d}W_s, \ t \in [0,T], \\ X_0 = \xi_X, \end{cases} \tag{3.1}$$

where $f : \mathbb{R}^d \to \mathbb{R}^d$ stands for the drift coefficient, $g : \mathbb{R}^d \to \mathbb{R}^{d\times m}$ the diffusion coefficient and $\xi_X : \Omega \to \mathbb{R}^d$ the initial data. Also, we consider an approximation process given by

$$\begin{cases} Y_t - Y_0 = \int_0^t a(s)\,\mathrm{d}s + \int_0^t b(s)\,\mathrm{d}W_s,\ t \in [0,T], \\ Y_0 = \xi_Y, \end{cases} \tag{3.2}$$

where $a : \Omega \times [0,T] \to \mathbb{R}^d$ and $b : \Omega \times [0,T] \to \mathbb{R}^{d\times m}$ are two stochastic processes that are integrable in the sense of Lebesgue integral and Itô stochastic integral, respectively. This can be regarded as a perturbation of the solution process of the original SDE (3.1). For example, when the Euler–Maruyama method is used to approximate (3.1) on a uniform grid $\{t_n = nh\}_{0\le n\le N}$ with stepsize $h = \frac{T}{N}$, one can get a continuous version of the approximation as

$$Y_t - Y_0 = \int_0^t f(Y_{\lfloor s\rfloor_N})\,\mathrm{d}s + \int_0^t g(Y_{\lfloor s\rfloor_N})\,\mathrm{d}W_s,\ t \in [0,T],\ Y_0 = \xi_X, \tag{3.3}$$

where in the notation of (3.2) we have $a(s) = f(Y_{\lfloor s\rfloor_N}), b(s) = g(Y_{\lfloor s\rfloor_N})$ and $\xi_Y = \xi_X$.

The following lemma provides two estimates, which will be essentially used later to obtain the desired perturbation estimates. The first assertion can be regarded as a modification of (Hutzenthaler & Jentzen, 2020, Proposition 2.9) and the second one is new.

LEMMA 3.1. Let $f : \mathbb{R}^d \to \mathbb{R}^d, g : \mathbb{R}^d \to \mathbb{R}^{d\times m}$ be measurable functions. Let $a : [0,T] \times \Omega \to \mathbb{R}^d, b : [0,T] \times \Omega \to \mathbb{R}^{d\times m}$ be predictable stochastic processes and let $\tau : \Omega \to [0,T]$ be a stopping time. Let $\{X_s\}_{s\in[0,T]}$ and $\{Y_s\}_{s\in[0,T]}$ be defined by (3.1) and (3.2) with continuous sample paths, respectively. Assume that $\int_0^T |a(s)| + \|b(s)\|^2 + |f(X_s)| + \|g(X_s)\|^2 + |f(Y_s)| + \|g(Y_s)\|^2\,\mathrm{d}s < \infty$ $\mathbb{P}$-a.s. and for $\varepsilon \in (0,\infty), p \ge 2$ with $\mathbb{P}$-a.s.

$$\int_0^\tau \left[\frac{\langle X_s - Y_s, f(X_s) - f(Y_s)\rangle + \frac{(1+\varepsilon)(p-1)}{2}\|g(X_s) - g(Y_s)\|^2}{|X_s - Y_s|^2}\right]^+ \mathrm{d}s < \infty. \tag{3.4}$$

Then for any $u \in [0,T]$ it holds

$$\begin{aligned}
&\sup_{t\in[0,u]} \left\| \frac{|X_{t\wedge\tau} - Y_{t\wedge\tau}|}{\exp(\int_0^{t\wedge\tau} \frac{1}{p}\eta_{p,r}\,\mathrm{d}r)} \right\|_{L^p(\Omega;\mathbb{R})} \\
&\le \sup_{t\in[0,u]} \Bigg[\|\xi_X - \xi_Y\|^p_{L^p(\Omega;\mathbb{R}^d)} + \underbrace{\mathbb{E}\left[\int_0^{t\wedge\tau} \frac{p|X_s - Y_s|^{p-2}}{\exp(\int_0^s \eta_{p,r}\,\mathrm{d}r)} \big\langle X_s - Y_s, \big(g(X_s) - b(s)\big)\,\mathrm{d}W_s\big\rangle\right]}_{=:\mathbb{S}_1} \\
&+ \underbrace{\mathbb{E}\left[\int_0^{t\wedge\tau} \frac{p|X_s - Y_s|^{p-2}\langle X_s - Y_s, f(Y_s) - a(s)\rangle}{\exp(\int_0^s \eta_{p,r}\,\mathrm{d}r)}\,\mathrm{d}s\right]}_{=:\mathbb{S}_2} \\
&+ \underbrace{\mathbb{E}\left[\int_0^{t\wedge\tau} \frac{p|X_s - Y_s|^{p-2}\frac{(p-1)(1+1/\varepsilon)}{2}\|g(Y_s) - b(s)\|^2}{\exp(\int_0^s \eta_{p,r}\,\mathrm{d}r)}\,\mathrm{d}s\right]}_{=:\mathbb{S}_3} \Bigg]^{1/p}.
\end{aligned} \tag{3.5}$$

Furthermore, we have

$$\begin{aligned}
&\Big\| \sup_{t\in[0,u]} \tfrac{|X_{t\wedge\tau}-Y_{t\wedge\tau}|}{\exp(\int_0^{t\wedge\tau}\frac{1}{2}\eta_{2,r}\,\mathrm{d}r)} \Big\|_{L^p(\Omega;\mathbb{R})} \\
&\le \Bigg[\|\xi_X-\xi_Y\|^2_{L^p(\Omega;\mathbb{R}^d)} + \underbrace{\Big\| \sup_{t\in[0,u]} \int_0^{t\wedge\tau} \tfrac{2\langle X_s-Y_s,(g(X_s)-b(s))\,\mathrm{d}W_s\rangle}{\exp(\int_0^s\eta_{2,r}\,\mathrm{d}r)} \Big\|_{L^{p/2}(\Omega;\mathbb{R})}}_{=:\mathbb{T}_1} \\
&\quad + \underbrace{\Big\| \sup_{t\in[0,u]} \int_0^{t\wedge\tau} \tfrac{2\langle X_s-Y_s,f(Y_s)-a(s)\rangle}{\exp(\int_0^s\eta_{2,r}\,\mathrm{d}r)}\,\mathrm{d}s \Big\|_{L^{p/2}(\Omega;\mathbb{R})}}_{=:\mathbb{T}_2} \\
&\quad + \underbrace{\Big\| \sup_{t\in[0,u]} \int_0^{t\wedge\tau} \tfrac{(1+\frac{1}{\varepsilon})\|g(Y_s)-b(s)\|^2}{\exp(\int_0^s\eta_{2,r}\,\mathrm{d}r)}\,\mathrm{d}s \Big\|_{L^{p/2}(\Omega;\mathbb{R})}}_{=:\mathbb{T}_3} \Bigg]^{1/2}.
\end{aligned} \tag{3.6}$$

Here for $z\ge 2$, we denote

$$\eta_{z,r} := z\mathbb{1}_{r\le\tau}(\omega)\left[\frac{\langle X_r-Y_r,f(X_r)-f(Y_r)\rangle+\frac{(z-1)(1+\varepsilon)}{2}\|g(X_r)-g(Y_r)\|^2}{|X_r-Y_r|^2}\right]^+. \tag{3.7}$$

Proof. For fixed $p\ge 2$, it is easy to validate that $\eta_{p,r}:[0,T]\times\Omega$ defined by (3.7) is well-defined due to (3.4). The Itô formula, the Itô product rule and the inequality $(a+b)^2\le(1+\varepsilon)a^2+(1+\frac{1}{\varepsilon})b^2$ yield

$$\begin{aligned}
\frac{|X_{t\wedge\tau}-Y_{t\wedge\tau}|^p}{\exp(\int_0^{t\wedge\tau}\eta_{p,r}\,\mathrm{d}r)} &= |\xi_X-\xi_Y|^p + \int_0^{t\wedge\tau} \frac{p|X_s-Y_s|^{p-2}}{\exp(\int_0^s\eta_{p,r}\,\mathrm{d}r)}\,\langle X_s-Y_s,(g(X_s)-b(s))\,\mathrm{d}W_s\rangle \\
&\quad + \int_0^{t\wedge\tau} \frac{p|X_s-Y_s|^{p-2}\langle X_s-Y_s,f(X_s)-a(s)\rangle-|X_s-Y_s|^p\eta_{p,s}}{\exp(\int_0^s\eta_{p,r}\,\mathrm{d}r)}\,\mathrm{d}s \\
&\quad + \int_0^{t\wedge\tau} \frac{\frac{p(p-2)}{2}|X_s-Y_s|^{p-4}|(X_s-Y_s)^*(g(X_s)-b(s))|^2+\frac{p}{2}|X_s-Y_s|^{p-2}\|g(X_s)-b(s)\|^2}{\exp(\int_0^s\eta_{p,r}\,\mathrm{d}r)}\,\mathrm{d}s \\
&\le |\xi_X-\xi_Y|^p + \int_0^{t\wedge\tau} \frac{p|X_s-Y_s|^{p-2}}{\exp(\int_0^s\eta_{p,r}\,\mathrm{d}r)}\,\langle X_s-Y_s,(g(X_s)-b(s))\,\mathrm{d}W_s\rangle \\
&\quad + \int_0^{t\wedge\tau} \frac{p|X_s-Y_s|^{p-2}\Big(\frac{(p-1)(1+\varepsilon)}{2}\|g(X_s)-g(Y_s)\|^2+\langle X_s-Y_s,f(X_s)-f(Y_s)\rangle\Big)-|X_s-Y_s|^p\eta_{p,s}}{\exp(\int_0^s\eta_{p,r}\,\mathrm{d}r)}\,\mathrm{d}s \\
&\quad + \int_0^{t\wedge\tau} \frac{p|X_s-Y_s|^{p-2}\Big(\frac{(p-1)(1+1/\varepsilon)}{2}\|g(Y_s)-b(s)\|^2+\langle X_s-Y_s,f(Y_s)-a(s)\rangle\Big)}{\exp(\int_0^s\eta_{p,r}\,\mathrm{d}r)}\,\mathrm{d}s.
\end{aligned} \tag{3.8}$$

Then we arrive at (3.5) by taking expectation of both sides of (3.8). Similarly, to show (3.6), letting $p=2$ in (3.8) deduces that

$$\begin{aligned}
\frac{|X_{t\wedge\tau}-Y_{t\wedge\tau}|^2}{\exp(\int_0^{t\wedge\tau}\eta_{2,r}\,\mathrm{d}r)} &\le |\xi_X-\xi_Y|^2 + \int_0^{t\wedge\tau} \frac{2\langle X_s-Y_s,(g(X_s)-b(s))\,\mathrm{d}W_s\rangle}{\exp(\int_0^s\eta_{2,r}\,\mathrm{d}r)} \\
&\quad + \int_0^{t\wedge\tau} \frac{2\langle X_s-Y_s,f(Y_s)-a(s)\rangle+(1+\frac{1}{\varepsilon})\|g(Y_s)-b(s)\|^2}{\exp(\int_0^s\eta_{2,r}\,\mathrm{d}r)}\,\mathrm{d}s.
\end{aligned} \tag{3.9}$$

This clearly implies (3.6), after taking supremum and $\|\cdot\|_{L^{p/2}(\Omega;\mathbb{R}^d)}$-norm of both sides of (3.9). □

As a consequence of Lemma 3.1, we state the main results of this section.

THEOREM 3.2. Let $f:\mathbb{R}^d\to\mathbb{R}^d, g:\mathbb{R}^d\to\mathbb{R}^{d\times m}$ be measurable functions with $f\in C^2(\mathbb{R}^d,\mathbb{R}^d)$, let $f\in\mathcal{C}^1_{\mathcal{P}}(\mathbb{R}^d,\mathbb{R}^d)$ with constants K_f, c_f and $g\in\mathcal{C}^1_{\mathcal{P}}(\mathbb{R}^d,\mathbb{R}^{d\times m})$ with constants K_g, c_g and let $a:[0,T]\times\Omega\to\mathbb{R}^d, b:[0,T]\times\Omega\to\mathbb{R}^{d\times m}$ be predictable stochastic processes. Let $\tau_N:\Omega\to[0,T]$ be a stopping time which may depend on N, let $\{X_s\}_{s\in[0,T]}$ and $\{Y_s\}_{s\in[0,T]}$ be defined by (3.1) and (3.2) with continuous sample paths, respectively, and let the uniform mesh be constructed by (2.2). Assume that $\int_0^T|a(s)|+\|b(s)\|^2+|f(X_s)|+\|g(X_s)\|^2+|f(Y_s)|+\|g(Y_s)\|^2\,\mathrm{d}s<\infty$ $\mathbb{P}$-a.s. and for $\varepsilon\in(0,\infty)$ with $\mathbb{P}$-a.s.

$$\int_0^{\tau_N}\left[\frac{\langle X_s-Y_s,f(X_s)-f(Y_s)\rangle+\frac{(1+\varepsilon)(p-1)}{2}\|g(X_s)-g(Y_s)\|^2}{|X_s-Y_s|^2}\right]^+\mathrm{d}s<\infty. \tag{3.10}$$

Moreover, let $K_{sup}>0$ be some constant that is independent of N, let $\xi_0:=(\xi_X-\xi_Y)\in L^p(\Omega;\mathbb{R}^d), p\geq 4$, and suppose that

(a) $\{s\leq\tau_N\}\in\mathcal{F}_{\lfloor s\rfloor_N}$;

(b) $\sup_{s\in[0,T]}\left\|\mathbb{1}_{s\leq\tau_N}\left[\frac{\langle X_s-Y_s,f(X_s)-f(Y_s)\rangle+\frac{(1+\varepsilon)(p-1)}{2}\|g(X_s)-g(Y_s)\|^2}{|X_s-Y_s|^2}\right]^+\right\|_{L^{3p}(\Omega;\mathbb{R})}\leq K_{sup}$;

(c) for any $i=1,...,d$, $\sup_{s\in[0,T]}\|\mathrm{Hess}_x(f^{(i)}(Y_s))\|_{L^{3p}(\Omega;\mathbb{R}^{d\times d})}\leq K_{sup}$ and

$$\begin{aligned}&\sup_{s\in[0,T]}\|X_s\|_{L^{6pc_g\vee 3pc_f\vee 3p}(\Omega;\mathbb{R}^d)}\bigvee\sup_{s\in[0,T]}\|Y_s\|_{L^{6pc_g\vee 3pc_f}(\Omega;\mathbb{R}^d)}\bigvee\\&\sup_{s\in[0,T]}\|a(s)\|_{L^{3p}(\Omega;\mathbb{R}^d)}\bigvee\sup_{s\in[0,T]}\|b(s)\|_{L^{3p}(\Omega;\mathbb{R}^{d\times m})}\leq K_{sup}.\end{aligned} \tag{3.11}$$

1. Then for any $u\in[0,T], v\in(0,\infty), q\in(0,\infty]$ with $\frac{1}{p}+\frac{1}{q}=\frac{1}{v}$, there exists some positive constant C that might depend on $p,\varepsilon,d,m,T,K_f,c_f,K_g,c_g,K_{sup}$, but do not depend on h, such that

$$\begin{aligned}\sup_{t\in[0,u]}\|X_{t\wedge\tau_N}-Y_{t\wedge\tau_N}\|_{L^v(\Omega;\mathbb{R}^d)}\leq&\Bigg[\|\xi_0\|^p_{L^p(\Omega;\mathbb{R}^d)}\\&+C\Bigg(h^p+\mathbb{E}\Big[\int_0^u\mathbb{1}_{s\leq\tau_N}\|g(Y_s)-b(s)\|^p\,\mathrm{d}s\Big]+h^{\frac{p}{2}-1}\int_0^u\int_{\lfloor s\rfloor_N}^s\Big(\mathbb{E}\Big[\mathbb{1}_{r\leq\tau_N}\big\|g(Y_r)-b(r)\big\|^p\Big]\Big)^{\frac{1}{2}}\mathrm{d}r\,\mathrm{d}s\\&+\mathbb{E}\Big[\int_0^u\mathbb{1}_{s\leq\tau_N}|f(Y_{\lfloor s\rfloor_N})-a(s)|^p\,\mathrm{d}s\Big]+h^{\frac{p}{2}-1}\mathbb{E}\Big[\int_0^u\int_{\lfloor s\rfloor_N}^s\mathbb{1}_{r\leq\tau_N}|f(Y_{\lfloor r\rfloor_N})-a(r)|^p\,\mathrm{d}r\,\mathrm{d}s\Big]\Bigg)\Bigg]^{\frac{1}{p}}\\&\times\left\|\exp\left(\int_0^{\tau_N}\left[\frac{\langle X_s-Y_s,f(X_s)-f(Y_s)\rangle+\frac{(1+\varepsilon)(p-1)}{2}\|g(X_s)-g(Y_s)\|^2}{|X_s-Y_s|^2}\right]^+\mathrm{d}s\right)\right\|_{L^q(\Omega;\mathbb{R})}.\end{aligned} \tag{3.12}$$

2. In addition to the same settings as (1), if g is Lipschitz, then it holds

$$\Big\| \sup_{t\in[0,u]} |X_{t\wedge\tau_N} - Y_{t\wedge\tau_N}| \Big\|_{L^\nu(\Omega;\mathbb{R})} \leq \bigg[\|\xi_0\|^2_{L^p(\Omega;\mathbb{R}^d)} + C\bigg(h^2 + \int_0^u \Big\| \mathbb{1}_{s\leq\tau_N} \|g(Y_s) - b(s)\| \Big\|^2_{L^p(\Omega;\mathbb{R})} \mathrm{d}s$$

$$+ h^{\frac{1}{2}} \int_0^u \bigg(\int_{\lfloor s\rfloor_N}^s \Big\| \mathbb{1}_{r\leq\tau_N} \|g(Y_r) - b(r)\| \Big\|^2_{L^p(\Omega;\mathbb{R})} \mathrm{d}r \bigg)^{\frac{1}{2}} \mathrm{d}s$$

$$+ \int_0^u \Big\| \mathbb{1}_{s\leq\tau_N} |f(Y_{\lfloor s\rfloor_N}) - a(s)| \Big\|^2_{L^p(\Omega;\mathbb{R})} \mathrm{d}s + h^{\frac{1}{2}} \int_0^u \int_{\lfloor s\rfloor_N}^s \Big\| \mathbb{1}_{r\leq\tau_N} |f(Y_{\lfloor r\rfloor_N}) - a(r)| \Big\|_{L^p(\Omega;\mathbb{R})} \mathrm{d}r\, \mathrm{d}s \bigg) \bigg]^{\frac{1}{2}}$$

$$\times \bigg\| \exp\bigg(\int_0^{\tau_N} \Big[\tfrac{\langle X_s - Y_s, f(X_s) - f(Y_s)\rangle + \frac{1+\varepsilon}{2}\|g(X_s) - g(Y_s)\|^2}{|X_s - Y_s|^2} \Big]^+ \mathrm{d}s \bigg) \bigg\|_{L^q(\Omega;\mathbb{R})}. \tag{3.13}$$

Proof. In the following exposition, we write η_s to represent $\eta_{z,s}$ for short. For any $u \in [0,T]$, by the Hölder inequality one infers that

$$\sup_{t\in[0,u]} \big\| |X_{t\wedge\tau_N} - Y_{t\wedge\tau_N}| \big\|_{L^\nu(\Omega;\mathbb{R})}$$

$$\leq \sup_{t\in[0,u]} \bigg\| \frac{|X_{t\wedge\tau_N} - Y_{t\wedge\tau_N}|}{\exp(\int_0^{t\wedge\tau_N} \frac{1}{p}\eta_r\, \mathrm{d}r)} \bigg\|_{L^p(\Omega;\mathbb{R})} \cdot \bigg\| \exp\bigg(\int_0^{\tau_N} \frac{1}{p}\eta_r\, \mathrm{d}r \bigg) \bigg\|_{L^q(\Omega;\mathbb{R})} \tag{3.14}$$

and

$$\Big\| \sup_{t\in[0,u]} |X_{t\wedge\tau_N} - Y_{t\wedge\tau_N}| \Big\|_{L^\nu(\Omega;\mathbb{R})}$$

$$\leq \bigg\| \sup_{t\in[0,u]} \frac{|X_{t\wedge\tau_N} - Y_{t\wedge\tau_N}|}{\exp(\int_0^{t\wedge\tau_N} \frac{1}{2}\eta_r\, \mathrm{d}r)} \bigg\|_{L^p(\Omega;\mathbb{R})} \cdot \bigg\| \exp\bigg(\int_0^{\tau_N} \frac{1}{2}\eta_r\, \mathrm{d}r \bigg) \bigg\|_{L^q(\Omega;\mathbb{R})}. \tag{3.15}$$

Therefore, it suffices to estimate terms $\mathbb{S}_1, \mathbb{S}_2, \mathbb{S}_3$ in (3.5) and $\mathbb{T}_1, \mathbb{T}_2, \mathbb{T}_3$ in (3.6). Observe that the term $\mathbb{S}_1$ vanishes, due to the condition (c) in Theorem 3.2 and the growth of g. For $\mathbb{S}_3$, by Young's inequality,

$$\mathbb{S}_3 \leq \mathbb{E}\bigg[\int_0^t C_{p,\varepsilon} \mathbb{1}_{s\leq\tau_N} \frac{|X_s - Y_s|^p}{\exp(\int_0^s \eta_r\, \mathrm{d}r)} + C_p \mathbb{1}_{s\leq\tau_N} \|g(Y_s) - b(s)\|^p \, \mathrm{d}s \bigg]$$

$$\leq C \int_0^t \mathbb{E}\bigg[\frac{|X_{s\wedge\tau_N} - Y_{s\wedge\tau_N}|^p}{\exp(\int_0^{s\wedge\tau_N} \eta_r\, \mathrm{d}r)} \bigg] \mathrm{d}s + C\mathbb{E}\bigg[\int_0^t \mathbb{1}_{s\leq\tau_N} \|g(Y_s) - b(s)\|^p \, \mathrm{d}s \bigg]. \tag{3.16}$$

Concerning $\mathbb{S}_2$, one can expand $f(Y_s) - f(Y_{\lfloor s\rfloor_N})$ by Itô's formula and then use the Young inequality and condition (c) in Theorem 3.2 to infer

$$\mathbb{S}_2 = \mathbb{E}\bigg[\int_0^{t\wedge\tau_N} \frac{p|X_s - Y_s|^{p-2} \big\langle X_s - Y_s, \int_{\lfloor s\rfloor_N}^s \langle f'(Y_r), a(r)\rangle + \frac{1}{2}\,\mathrm{trace}(b(r)^* \,\mathrm{Hess}_x(f(Y_r)) b(r))\, \mathrm{d}r \big\rangle}{\exp(\int_0^s \eta_r\, \mathrm{d}r)} \mathrm{d}s \bigg]$$

$$+ \mathbb{E}\bigg[\int_0^{t\wedge\tau_N} \frac{p|X_s - Y_s|^{p-2} \big\langle X_s - Y_s, \int_{\lfloor s\rfloor_N}^s \langle f'(Y_r), b(r)\, \mathrm{d}W_r\rangle \big\rangle}{\exp(\int_0^s \eta_r\, \mathrm{d}r)} \mathrm{d}s \bigg]$$

$$
\begin{aligned}
&+\mathbb{E}\bigg[\int_0^{t\wedge\tau_N} \frac{p|X_s-Y_s|^{p-2}\langle X_s-Y_s, f(Y_{\lfloor s\rfloor_N})-a(s)\rangle}{\exp(\int_0^s \eta_r\,\mathrm{d}r)}\,\mathrm{d}s\bigg]\\
&\le \mathbb{E}\bigg[\int_0^{t\wedge\tau_N} \frac{p|X_s-Y_s|^{p-1}\left|\int_{\lfloor s\rfloor_N}^s \langle f'(Y_r), a(r)\rangle+\frac{1}{2}\operatorname{trace}(b(r)^*\operatorname{Hess}_x(f(Y_r))b(r))\,\mathrm{d}r\right|}{\exp(\int_0^s \eta_r\,\mathrm{d}r)}\,\mathrm{d}s\bigg]\\
&\quad+\mathbb{E}\bigg[\int_0^{t\wedge\tau_N} \frac{p|X_s-Y_s|^{p-2}\left\langle X_s-Y_s, \int_{\lfloor s\rfloor_N}^s \langle f'(Y_r), b(r)\,\mathrm{d}W_r\rangle\right\rangle}{\exp(\int_0^s \eta_r\,\mathrm{d}r)}\,\mathrm{d}s\bigg]\\
&\quad+\mathbb{E}\bigg[\int_0^{t\wedge\tau_N} \frac{p|X_s-Y_s|^{p-1}|f(Y_{\lfloor s\rfloor_N})-a(s)|}{\exp(\int_0^s \eta_r\,\mathrm{d}r)}\,\mathrm{d}s\bigg]\\
&\le C\int_0^t \mathbb{E}\bigg[\frac{|X_{s\wedge\tau_N}-Y_{s\wedge\tau_N}|^p}{\exp(\int_0^{s\wedge\tau_N}\eta_r\,\mathrm{d}r)}\bigg]\mathrm{d}s + C\mathbb{E}\bigg[\int_0^t \mathbb{1}_{s\le\tau_N}|f(Y_{\lfloor s\rfloor_N})-a(s)|^p\,\mathrm{d}s\bigg]\\
&\quad+\mathbb{E}\bigg[\int_0^{t\wedge\tau_N} \frac{p|X_s-Y_s|^{p-2}\langle X_s-Y_s, \int_{\lfloor s\rfloor_N}^s \langle f'(Y_r), b(r)\,\mathrm{d}W_r\rangle\rangle}{\exp(\int_0^s \eta_r\,\mathrm{d}r)}\,\mathrm{d}s\bigg] + Ch^p. \qquad (3.17)
\end{aligned}
$$

To estimate the last but one term for $p \ge 4$, we expand the left item of the inner product by Itô's formula and Itô's product rule to obtain

$$
\begin{aligned}
&\frac{p|X_s-Y_s|^{p-2}(X_s-Y_s)}{\exp(\int_0^s \eta_r\,\mathrm{d}r)}\\
&= \frac{p|X_{\lfloor s\rfloor_N}-Y_{\lfloor s\rfloor_N}|^{p-2}(X_{\lfloor s\rfloor_N}-Y_{\lfloor s\rfloor_N})}{\exp(\int_0^{\lfloor s\rfloor_N}\eta_r\,\mathrm{d}r)} + \int_{\lfloor s\rfloor_N}^s \frac{p|X_r-Y_r|^{p-2}(f(X_r)-a(r))}{\exp(\int_0^r \eta_\iota\,\mathrm{d}\iota)}\,\mathrm{d}r\\
&\quad+\int_{\lfloor s\rfloor_N}^s \frac{-p|X_r-Y_r|^{p-2}(X_r-Y_r)\eta_r}{\exp(\int_0^r \eta_\iota\,\mathrm{d}\iota)}\,\mathrm{d}r + \int_{\lfloor s\rfloor_N}^s \frac{p|X_r-Y_r|^{p-2}(g(X_r)-b(r))}{\exp(\int_0^r \eta_\iota\,\mathrm{d}\iota)}\,\mathrm{d}W_r\\
&\quad+\int_{\lfloor s\rfloor_N}^s \frac{p(p-2)(X_r-Y_r)}{\exp(\int_0^r \eta_\iota\,\mathrm{d}\iota)}|X_r-Y_r|^{p-4}\langle X_r-Y_r, f(X_r)-a(r)\rangle\,\mathrm{d}r\\
&\quad+\int_{\lfloor s\rfloor_N}^s \frac{p(p-2)(X_r-Y_r)}{\exp(\int_0^r \eta_\iota\,\mathrm{d}\iota)}|X_r-Y_r|^{p-4}\left\langle X_r-Y_r, \big(g(X_r)-b(r)\big)\,\mathrm{d}W_r\right\rangle\\
&\quad+\int_{\lfloor s\rfloor_N}^s \frac{p(p-2)(p-4)(X_r-Y_r)}{2\exp(\int_0^r \eta_\iota\,\mathrm{d}\iota)}|X_r-Y_r|^{p-6}\left|(X_r-Y_r)\big(g(X_r)-b(r)\big)\right|^2\mathrm{d}r\\
&\quad+\int_{\lfloor s\rfloor_N}^s \frac{p(p-2)(X_r-Y_r)}{2\exp(\int_0^r \eta_\iota\,\mathrm{d}\iota)}|X_r-Y_r|^{p-4}\left\|g(X_r)-b(r)\right\|^2\mathrm{d}r\\
&\quad+\int_{\lfloor s\rfloor_N}^s \frac{p(p-2)|X_r-Y_r|^{p-4}}{\exp(\int_0^r \eta_\iota\,\mathrm{d}\iota)}\big(g(X_r)-b(r)\big)\big(g(X_r)-b(r)\big)^*(X_r-Y_r)\,\mathrm{d}r. \qquad (3.18)
\end{aligned}
$$

Collecting some terms in (3.18) one can deduce

$$
\begin{aligned}
&\mathbb{E}\bigg[\int_0^{t\wedge\tau_N}\left\langle \frac{p|X_s-Y_s|^{p-2}(X_s-Y_s)}{\exp(\int_0^s \eta_r\,\mathrm{d}r)}, \int_{\lfloor s\rfloor_N}^s \langle f'(Y_r), b(r)\,\mathrm{d}W_r\rangle\right\rangle\mathrm{d}s\bigg]\\
&\le \mathbb{E}\bigg[\int_0^{t\wedge\tau_N}\left\langle \frac{p|X_{\lfloor s\rfloor_N}-Y_{\lfloor s\rfloor_N}|^{p-2}(X_{\lfloor s\rfloor_N}-Y_{\lfloor s\rfloor_N})}{\exp(\int_0^{\lfloor s\rfloor_N}\eta_r\,\mathrm{d}r)}, \int_{\lfloor s\rfloor_N}^s \langle f'(Y_r), b(r)\,\mathrm{d}W_r\rangle\right\rangle\mathrm{d}s\bigg]
\end{aligned}
$$

$$
\begin{aligned}
&+\mathbb{E}\bigg[\int_0^{t\wedge\tau_N}\int_{\lfloor s\rfloor_N}^{s}\tfrac{p(p-1)|X_r-Y_r|^{p-2}|f(X_r)-a(r)|}{\exp(\int_0^r\eta_\iota\,\mathrm{d}\iota)}\,\mathrm{d}r\bigg|\int_{\lfloor s\rfloor_N}^{s}\langle f'(Y_r),b(r)\,\mathrm{d}W_r\rangle\bigg|\,\mathrm{d}s\bigg]\\
&+\mathbb{E}\bigg[\int_0^{t\wedge\tau_N}\Big\langle\int_{\lfloor s\rfloor_N}^{s}\tfrac{-p|X_r-Y_r|^{p-2}(X_r-Y_r)\eta_r}{\exp(\int_0^r\eta_\iota\,\mathrm{d}\iota)}\,\mathrm{d}r,\int_{\lfloor s\rfloor_N}^{s}\langle f'(Y_r),b(r)\,\mathrm{d}W_r\rangle\Big\rangle\,\mathrm{d}s\bigg]\\
&+\mathbb{E}\bigg[\int_0^{t\wedge\tau_N}\Big\langle\int_{\lfloor s\rfloor_N}^{s}\tfrac{p|X_r-Y_r|^{p-2}(g(X_r)-b(r))}{\exp(\int_0^r\eta_\iota\,\mathrm{d}\iota)}\,\mathrm{d}W_r,\int_{\lfloor s\rfloor_N}^{s}\langle f'(Y_r),b(r)\,\mathrm{d}W_r\rangle\Big\rangle\,\mathrm{d}s\bigg]\\
&+\mathbb{E}\bigg[\int_0^{t\wedge\tau_N}\Big\langle\int_{\lfloor s\rfloor_N}^{s}\tfrac{(X_r-Y_r)|X_r-Y_r|^{p-4}}{\frac{1}{p(p-2)}\exp(\int_0^r\eta_\iota\,\mathrm{d}\iota)}\big\langle X_r-Y_r,\big(g(X_r)-b(r)\big)\,\mathrm{d}W_r\big\rangle,\int_{\lfloor s\rfloor_N}^{s}\langle f'(Y_r),b(r)\,\mathrm{d}W_r\rangle\Big\rangle\,\mathrm{d}s\bigg]\\
&+\mathbb{E}\bigg[\int_0^{t\wedge\tau_N}\int_{\lfloor s\rfloor_N}^{s}\tfrac{p(p-1)(p-2)|X_r-Y_r|^{p-3}}{2\exp(\int_0^r\eta_\iota\,\mathrm{d}\iota)}\big\|g(X_r)-b(r)\big\|^2\,\mathrm{d}r\bigg|\int_{\lfloor s\rfloor_N}^{s}\langle f'(Y_r),b(r)\,\mathrm{d}W_r\rangle\bigg|\,\mathrm{d}s\bigg]\\
&=: B_1+B_2+B_3+B_4+B_5+B_6. \qquad (3.19)
\end{aligned}
$$

Next we estimate $B_i, i=1,2,...,6$ term by term. First, it is trivial to see $B_1=0$ by the condition (a) and (c) in Theorem 3.2. For B_2, using the fact $f\in\mathcal{C}^1_{\mathcal{P}}(\mathbb{R}^d,\mathbb{R}^d)$, Hölder's inequality, Young's inequality and condition (c) in Theorem 3.2, we have

$$
\begin{aligned}
B_2 \le\; & C\int_0^t\int_{\lfloor s\rfloor_N}^{s}\mathbb{E}\bigg[\mathbb{1}_{s\le\tau_N}\tfrac{|X_r-Y_r|^{p-1}(1+|X_r|+|Y_r|)^{c_f}}{\exp(\int_0^r\eta_\iota\,\mathrm{d}\iota)}\bigg|\int_{\lfloor s\rfloor_N}^{s}\big\langle f'(Y_\iota),b(\iota)\,\mathrm{d}W_\iota\big\rangle\bigg|\bigg]\,\mathrm{d}r\,\mathrm{d}s\\
&+C\int_0^t\int_{\lfloor s\rfloor_N}^{s}\mathbb{E}\bigg[\mathbb{1}_{s\le\tau_N}\tfrac{|X_r-Y_r|^{p-2}|f(Y_r)-f(Y_{\lfloor r\rfloor_N})|}{\exp(\int_0^r\eta_\iota\,\mathrm{d}\iota)}\bigg|\int_{\lfloor s\rfloor_N}^{s}\big\langle f'(Y_\iota),b(\iota)\,\mathrm{d}W_\iota\big\rangle\bigg|\bigg]\,\mathrm{d}r\,\mathrm{d}s\\
&+C\int_0^t\int_{\lfloor s\rfloor_N}^{s}\mathbb{E}\bigg[\mathbb{1}_{s\le\tau_N}\tfrac{|X_r-Y_r|^{p-2}|f(Y_{\lfloor r\rfloor_N})-a(r)|}{\exp(\int_0^r\eta_\iota\,\mathrm{d}\iota)}\bigg|\int_{\lfloor s\rfloor_N}^{s}\big\langle f'(Y_\iota),b(\iota)\,\mathrm{d}W_\iota\big\rangle\bigg|\bigg]\,\mathrm{d}r\,\mathrm{d}s\\
\le\; & C\int_0^t\sup_{r\in[\lfloor s\rfloor_N,s]}\mathbb{E}\bigg[\mathbb{1}_{s\le\tau_N}\tfrac{|X_r-Y_r|^{p}}{\exp(\int_0^r\eta_\iota\,\mathrm{d}\iota)}\bigg]\,\mathrm{d}s\\
&+C\int_0^t\bigg(\int_{\lfloor s\rfloor_N}^{s}\bigg\|\mathbb{1}_{s\le\tau_N}\big(1+|X_r|+|Y_r|\big)^{c_f}\bigg|\int_{\lfloor s\rfloor_N}^{s}\big\langle f'(Y_\iota),b(\iota)\,\mathrm{d}W_\iota\big\rangle\bigg|\bigg\|_{L^p(\Omega;\mathbb{R})}\mathrm{d}r\bigg)^p\mathrm{d}s\\
&+C\int_0^t\bigg(\int_{\lfloor s\rfloor_N}^{s}\bigg\|\mathbb{1}_{s\le\tau_N}\big|f(Y_r)-f(Y_{\lfloor r\rfloor_N})\big|\cdot\bigg|\int_{\lfloor s\rfloor_N}^{s}\big\langle f'(Y_\iota),b(\iota)\,\mathrm{d}W_\iota\big\rangle\bigg|\bigg\|_{L^{p/2}(\Omega;\mathbb{R})}\mathrm{d}r\bigg)^{p/2}\mathrm{d}s\\
&+C\int_0^t\bigg(\int_{\lfloor s\rfloor_N}^{s}\bigg\|\mathbb{1}_{s\le\tau_N}\big|f(Y_{\lfloor r\rfloor_N})-a(r)\big|\cdot\bigg|\int_{\lfloor s\rfloor_N}^{s}\big\langle f'(Y_\iota),b(\iota)\,\mathrm{d}W_\iota\big\rangle\bigg|\bigg\|_{L^{p/2}(\Omega;\mathbb{R})}\mathrm{d}r\bigg)^{p/2}\mathrm{d}s\\
\le\; & C\int_0^t\sup_{r\in[0,s]}\mathbb{E}\bigg[\tfrac{|X_{r\wedge\tau_N}-Y_{r\wedge\tau_N}|^p}{\exp(\int_0^{r\wedge\tau_N}\eta_\iota\,\mathrm{d}\iota)}\bigg]\mathrm{d}s+Ch^{\frac{p}{2}-1}\mathbb{E}\bigg[\int_0^t\int_{\lfloor s\rfloor_N}^{s}\mathbb{1}_{r\le\tau_N}|f(Y_{\lfloor r\rfloor_N})-a(r)|^p\,\mathrm{d}r\,\mathrm{d}s\bigg]+Ch^p. \qquad (3.20)
\end{aligned}
$$

With the aid of Hölder's inequality, Young's inequality and condition (c) in Theorem 3.2, we estimate B_3 as follows:

$$\begin{aligned} B_3 &\leq \int_0^t \int_{\lfloor s \rfloor_N}^s \mathbb{E}\Big[\mathbb{1}_{s\leq\tau_N} \frac{p|X_r-Y_r|^{p-1}\eta_r}{\exp(\int_0^r \eta_\iota\,\mathrm{d}\iota)} \Big| \int_{\lfloor s \rfloor_N}^s \big\langle f'(Y_\iota), b(\iota)\,\mathrm{d}W_\iota \big\rangle \Big| \Big] \mathrm{d}r\,\mathrm{d}s \\ &\leq C\int_0^t \sup_{r\in[\lfloor s \rfloor_N, s]} \mathbb{E}\Big[\mathbb{1}_{s\leq\tau_N} \frac{|X_r-Y_r|^p}{\exp(\int_0^r \eta_\iota\,\mathrm{d}\iota)} \Big] \mathrm{d}s \\ &\quad + C\int_0^t \Big(\int_{\lfloor s \rfloor_N}^s \Big\| \mathbb{1}_{s\leq\tau_N} \eta_r \Big| \int_{\lfloor s \rfloor_N}^s \big\langle f'(Y_\iota), b(\iota)\,\mathrm{d}W_\iota \big\rangle \Big| \Big\|_{L^p(\Omega;\mathbb{R})} \mathrm{d}r \Big)^p \mathrm{d}s \\ &\leq C\int_0^t \sup_{r\in[0,s]} \mathbb{E}\Big[\frac{|X_{r\wedge\tau_N}-Y_{r\wedge\tau_N}|^p}{\exp(\int_0^{r\wedge\tau_N} \eta_\iota\,\mathrm{d}\iota)} \Big] \mathrm{d}s + Ch^{3p/2}. \end{aligned} \tag{3.21}$$

Using the property of stochastic integral, Hölder's inequality, Young's inequality and condition (c) in Theorem 3.2 shows the estimate of B_4:

$$\begin{aligned} B_4 &= \int_0^t \int_{\lfloor s \rfloor_N}^s \mathbb{E}\Big[\mathbb{1}_{s\leq\tau_N} \Big\langle \frac{p|X_r-Y_r|^{p-2}(g(X_r)-b(r))}{\exp(\int_0^r \eta_\iota\,\mathrm{d}\iota)}, f'(Y_r)b(r) \Big\rangle_{HS} \Big] \mathrm{d}r\,\mathrm{d}s \\ &\leq \int_0^t \int_{\lfloor s \rfloor_N}^s \mathbb{E}\Big[\mathbb{1}_{s\leq\tau_N} \frac{p|X_r-Y_r|^{p-2}\|g(X_r)-g(Y_r)\|}{\exp(\int_0^r \eta_\iota\,\mathrm{d}\iota)} \big\|f'(Y_r)b(r)\big\| \Big] \mathrm{d}r\,\mathrm{d}s \\ &\quad + \int_0^t \int_{\lfloor s \rfloor_N}^s \mathbb{E}\Big[\mathbb{1}_{s\leq\tau_N} \frac{p|X_r-Y_r|^{p-2}\|g(Y_r)-b(r)\|}{\exp(\int_0^r \eta_\iota\,\mathrm{d}\iota)} \big\|f'(Y_r)b(r)\big\| \Big] \mathrm{d}r\,\mathrm{d}s \\ &\leq C\int_0^t \sup_{r\in[\lfloor s \rfloor_N, s]} \mathbb{E}\Big[\mathbb{1}_{s\leq\tau_N} \frac{|X_r-Y_r|^p}{\exp(\int_0^r \eta_\iota\,\mathrm{d}\iota)} \Big] \mathrm{d}s \\ &\quad + C\int_0^t \Big(\int_{\lfloor s \rfloor_N}^s \Big\| \mathbb{1}_{s\leq\tau_N} \big(1+|X_r|+|Y_r|\big)^{c_g} \big\|f'(Y_r)b(r)\big\| \Big\|_{L^p(\Omega;\mathbb{R})} \mathrm{d}r \Big)^p \mathrm{d}s \\ &\quad + C\int_0^t \Big(\int_{\lfloor s \rfloor_N}^s \Big\| \mathbb{1}_{s\leq\tau_N} \big\|g(Y_r)-b(r)\big\| \big\|f'(Y_r)b(r)\big\| \Big\|_{L^{p/2}(\Omega;\mathbb{R})} \mathrm{d}r \Big)^{p/2} \mathrm{d}s \\ &\leq C\int_0^t \sup_{r\in[0,s]} \mathbb{E}\Big[\frac{|X_{r\wedge\tau_N}-Y_{r\wedge\tau_N}|^p}{\exp(\int_0^{r\wedge\tau_N} \eta_\iota\,\mathrm{d}\iota)} \Big] \mathrm{d}s \\ &\quad + Ch^{\frac{p-2}{2}} \int_0^t \int_{\lfloor s \rfloor_N}^s \Big(\mathbb{E}\Big[\mathbb{1}_{r\leq\tau_N} \big\|g(Y_r)-b(r)\big\|^p \Big] \Big)^{1/2} \mathrm{d}r\,\mathrm{d}s + Ch^p. \end{aligned} \tag{3.22}$$

In a similar way, one can handle B_5 as follows:

$$\begin{aligned} B_5 &\leq \int_0^t \int_{\lfloor s \rfloor_N}^s \mathbb{E}\Big[\mathbb{1}_{s\leq\tau_N} \Big\| \frac{|X_r-Y_r|^{p-2}}{\exp(\int_0^r \eta_\iota\,\mathrm{d}\iota)} \big(g(X_r)-b(r)\big) \Big\| \big\|f'(Y_r)b(r)\big\| \Big] \mathrm{d}r\,\mathrm{d}s \\ &\leq C\int_0^t \sup_{r\in[0,s]} \mathbb{E}\Big[\frac{|X_{r\wedge\tau_N}-Y_{r\wedge\tau_N}|^p}{\exp(\int_0^{r\wedge\tau_N} \eta_\iota\,\mathrm{d}\iota)} \Big] \mathrm{d}s \\ &\quad + Ch^{p/2-1} \int_0^t \int_{\lfloor s \rfloor_N}^s \Big(\mathbb{E}\Big[\mathbb{1}_{r\leq\tau_N} \big\|g(Y_r)-b(r)\big\|^p \Big] \Big)^{1/2} \mathrm{d}r\,\mathrm{d}s + Ch^p. \end{aligned} \tag{3.23}$$

Thanks to the Hölder inequality, the Young inequality and condition (c) in Theorem 3.2, we treat B_6 in the following way:

$$\begin{aligned}
B_6 &\le C\int_0^t\int_{\lfloor s\rfloor_N}^s \mathbb{E}\bigg[\mathbb{1}_{s\le\tau_N}\frac{|X_r-Y_r|^{p-1}}{\exp(\int_0^r\eta_\iota\,\mathrm{d}\iota)}\big(1+|X_r|+|Y_r|\big)^{2c_g}\bigg|\int_{\lfloor s\rfloor_N}^s\big\langle f'(Y_\iota),b(\iota)\,\mathrm{d}W_\iota\big\rangle\bigg|\bigg]\,\mathrm{d}r\,\mathrm{d}s\\
&\quad + C\int_0^t\int_{\lfloor s\rfloor_N}^s \mathbb{E}\bigg[\mathbb{1}_{s\le\tau_N}\frac{|X_r-Y_r|^{p-3}}{\exp(\int_0^r\eta_\iota\,\mathrm{d}\iota)}\big\|g(Y_r)-b(r)\big\|^2\bigg|\int_{\lfloor s\rfloor_N}^s\big\langle f'(Y_\iota),b(\iota)\,\mathrm{d}W_\iota\big\rangle\bigg|\bigg]\,\mathrm{d}r\,\mathrm{d}s\\
&\le C\int_0^t\sup_{r\in[0,s]}\mathbb{E}\bigg[\frac{|X_{r\wedge\tau_N}-Y_{r\wedge\tau_N}|^p}{\exp(\int_0^{r\wedge\tau_N}\eta_\iota\,\mathrm{d}\iota)}\bigg]\,\mathrm{d}s + Ch^{\frac{p-2}{2}}\int_0^t\int_{\lfloor s\rfloor_N}^s\Big(\mathbb{E}\Big[\mathbb{1}_{r\le\tau_N}\big\|g(Y_r)-b(r)\big\|^p\Big]\Big)^{\frac{2}{3}}\,\mathrm{d}r\,\mathrm{d}s + Ch^p.
\end{aligned}\tag{3.24}$$

Gathering (3.17), (3.20), (3.21), (3.22), (3.23) and (3.24) yields

$$\begin{aligned}
\mathbb{S}_2 &\le C\int_0^t\sup_{r\in[0,s]}\mathbb{E}\bigg[\frac{|X_{r\wedge\tau_N}-Y_{r\wedge\tau_N}|^p}{\exp(\int_0^{r\wedge\tau_N}\eta_\iota\,\mathrm{d}\iota)}\bigg]\,\mathrm{d}s + Ch^{\frac{p-2}{2}}\mathbb{E}\bigg[\int_0^t\int_{\lfloor s\rfloor_N}^s\mathbb{1}_{s\le\tau_N}\Big|f(Y_{\lfloor s\rfloor_N})-a(r)\Big|^p\,\mathrm{d}r\,\mathrm{d}s\bigg] + Ch^p\\
&\quad + Ch^{\frac{p-2}{2}}\int_0^t\int_{\lfloor s\rfloor_N}^s\Big(\mathbb{E}\Big[\mathbb{1}_{r\le\tau_N}\big\|g(Y_r)-b(r)\big\|^p\Big]\Big)^{\frac{1}{2}}\,\mathrm{d}r\,\mathrm{d}s + C\mathbb{E}\bigg[\int_0^t\mathbb{1}_{s\le\tau_N}\Big|f(Y_{\lfloor s\rfloor_N})-a(s)\Big|^p\,\mathrm{d}s\bigg].
\end{aligned}\tag{3.25}$$

Then, combining (3.16) with (3.25) and by Gronwall's inequality we arrive at the first assertion (3.12). Now it remains to validate (3.13). With regard to $\mathbb{T}_1$, by Lemma 2.1, the inner product inequality, the Hölder inequality and the elementary inequality, one deduces

$$\begin{aligned}
\mathbb{T}_1 &\le C_p\bigg(\int_0^u\sum_{i=1}^m\bigg\|\mathbb{1}_{s\le\tau_N}\frac{\langle X_s-Y_s,g^{(i)}(X_s)-b^{(i)}(s)\rangle}{\exp(\int_0^s\eta_r\,\mathrm{d}r)}\bigg\|^2_{L^{p/2}(\Omega;\mathbb{R})}\,\mathrm{d}s\bigg)^{1/2}\\
&\le C_p\sum_{i=1}^m\bigg(\int_0^u\bigg\|\mathbb{1}_{s\le\tau_N}\frac{|X_s-Y_s|}{\exp(\int_0^s\frac{1}{2}\eta_r\,\mathrm{d}r)}\bigg\|^2_{L^p(\Omega;\mathbb{R})}\cdot\bigg\|\mathbb{1}_{s\le\tau_N}\frac{|g^{(i)}(X_s)-b^{(i)}(s)|}{\exp(\int_0^s\frac{1}{2}\eta_r\,\mathrm{d}r)}\bigg\|^2_{L^p(\Omega;\mathbb{R})}\,\mathrm{d}s\bigg)^{1/2}\\
&\le \sup_{t\in[0,u]}\bigg\|\frac{|X_{t\wedge\tau_N}-Y_{t\wedge\tau_N}|}{\exp(\int_0^{t\wedge\tau_N}\frac{1}{2}\eta_r\,\mathrm{d}r)}\bigg\|_{L^p(\Omega;\mathbb{R})}\cdot C_p\sum_{i=1}^m\bigg(\int_0^u\bigg\|\mathbb{1}_{s\le\tau_N}\frac{|g^{(i)}(X_s)-b^{(i)}(s)|}{\exp(\int_0^s\frac{1}{2}\eta_r\,\mathrm{d}r)}\bigg\|^2_{L^p(\Omega;\mathbb{R})}\,\mathrm{d}s\bigg)^{1/2}\\
&\le \frac{1}{4}\sup_{t\in[0,u]}\bigg\|\frac{|X_{t\wedge\tau_N}-Y_{t\wedge\tau_N}|}{\exp(\int_0^{t\wedge\tau_N}\frac{1}{2}\eta_r\,\mathrm{d}r)}\bigg\|^2_{L^p(\Omega;\mathbb{R})}\\
&\quad + C_p\sum_{i=1}^m\bigg(\int_0^u\bigg\|\mathbb{1}_{s\le\tau_N}\frac{|g^{(i)}(X_s)-g^{(i)}(Y_s)+g^{(i)}(Y_s)-b^{(i)}(s)|}{\exp(\int_0^s\frac{1}{2}\eta_r\,\mathrm{d}r)}\bigg\|^2_{L^p(\Omega;\mathbb{R})}\,\mathrm{d}s\bigg),
\end{aligned}\tag{3.26}$$

where

$$\int_0^u \Big\| \mathbb{1}_{s\le\tau_N} \tfrac{|g^{(i)}(X_s)-g^{(i)}(Y_s)+g^{(i)}(Y_s)-b^{(i)}(s)|}{\exp(\int_0^s \frac{1}{2}\eta_r\,\mathrm{d}r)} \Big\|^2_{L^p(\Omega;\mathbb{R})} \mathrm{d}s$$
$$\le 2C \int_0^u \Big\| \tfrac{|X_{s\wedge\tau_N}-Y_{s\wedge\tau_N}|}{\exp(\int_0^{s\wedge\tau_N} \frac{1}{2}\eta_r\,\mathrm{d}r)} \Big\|^2_{L^p(\Omega;\mathbb{R})} \mathrm{d}s + 2\int_0^u \Big\| \mathbb{1}_{s\le\tau_N} \tfrac{|g^{(i)}(Y_s)-b^{(i)}(s)|}{\exp(\int_0^s \frac{1}{2}\eta_r\,\mathrm{d}r)} \Big\|^2_{L^p(\Omega;\mathbb{R})} \mathrm{d}s. \tag{3.27}$$

Combining (3.26) with (3.27) yields

$$\mathbb{T}_1 \le \frac{1}{4} \sup_{t\in[0,u]} \Big\| \tfrac{|X_{t\wedge\tau_N}-Y_{t\wedge\tau_N}|}{\exp(\int_0^{t\wedge\tau_N} \frac{1}{2}\eta_r\,\mathrm{d}r)} \Big\|^2_{L^p(\Omega;\mathbb{R})}$$
$$+ C\int_0^u \sup_{s\in[0,t]} \Big\| \tfrac{|X_{s\wedge\tau_N}-Y_{s\wedge\tau_N}|}{\exp(\int_0^{s\wedge\tau_N} \frac{1}{2}\eta_r\,\mathrm{d}r)} \Big\|^2_{L^p(\Omega;\mathbb{R})} \mathrm{d}t + C\int_0^u \Big\| \mathbb{1}_{s\le\tau_N}\big(g(Y_s)-b(s)\big)\Big\|^2_{L^p(\Omega;\mathbb{R}^{d\times m})} \mathrm{d}s. \tag{3.28}$$

When it comes to $\mathbb{T}_3$, one can easily show

$$\mathbb{T}_3 \le C\int_0^u \Big\| \mathbb{1}_{s\le\tau_N}\big(g(Y_s)-b(s)\big)\Big\|^2_{L^p(\Omega;\mathbb{R}^{d\times m})} \mathrm{d}s. \tag{3.29}$$

Note that the term $\mathbb{T}_2$ needs to be treated carefully. First, using the same arguments as $\mathbb{S}_2$ shows

$$\mathbb{T}_2 \le \Big\| \sup_{t\in[0,u]} \int_0^{t\wedge\tau_N} \Big\langle \tfrac{2(X_s-Y_s)}{\exp(\int_0^s \eta_r\,\mathrm{d}r)}, \int_{\lfloor s\rfloor_N}^s \langle f'(Y_r), a(r)\rangle + \tfrac{1}{2}\,\mathrm{trace}\,\big(b(r)^*\,\mathrm{Hess}_x(f(Y_r))b(r)\big)\,\mathrm{d}r\Big\rangle \mathrm{d}s \Big\|_{L^{\frac{p}{2}}(\Omega;\mathbb{R})}$$
$$+ \Big\| \sup_{t\in[0,u]} \int_0^{t\wedge\tau_N} 2\Big\langle \tfrac{X_s-Y_s}{\exp(\int_0^s \eta_r\,\mathrm{d}r)}, \int_{\lfloor s\rfloor_N}^s \langle f'(Y_r), b(r)\,\mathrm{d}W_r\rangle \Big\rangle \mathrm{d}s \Big\|_{L^{p/2}(\Omega;\mathbb{R})}$$
$$+ \Big\| \sup_{t\in[0,u]} \int_0^{t\wedge\tau_N} 2\Big\langle \tfrac{X_s-Y_s}{\exp(\int_0^s \eta_r\,\mathrm{d}r)}, f(Y_{\lfloor s\rfloor_N}) - a(s) \Big\rangle \mathrm{d}s \Big\|_{L^{p/2}(\Omega;\mathbb{R})}$$
$$\le C\int_0^u \sup_{s\in[0,t]} \Big\| \tfrac{|X_{s\wedge\tau_N}-Y_{s\wedge\tau_N}|}{\exp(\int_0^{s\wedge\tau_N} \frac{1}{2}\eta_r\,\mathrm{d}r)} \Big\|^2_{L^p(\Omega;\mathbb{R})} \mathrm{d}t + \int_0^u \Big\| \mathbb{1}_{s\le\tau_N} \big|f(Y_{\lfloor s\rfloor_N}) - a(s)\big| \Big\|^2_{L^p(\Omega;\mathbb{R})} \mathrm{d}s$$
$$+ \Big\| \sup_{t\in[0,u]} \int_0^{t\wedge\tau_N} 2\Big\langle \tfrac{X_s-Y_s}{\exp(\int_0^s \eta_r\,\mathrm{d}r)}, \int_{\lfloor s\rfloor_N}^s \langle f'(Y_r), b(r)\,\mathrm{d}W_r\rangle \Big\rangle \mathrm{d}s \Big\|_{L^{p/2}(\Omega;\mathbb{R})} + Ch^2. \tag{3.30}$$

To estimate the last but one term for $p \ge 4$, we expand the left item in the inner product by Itô's formula and Itô's product rule to acquire

$$\tfrac{X_s-Y_s}{\exp(\int_0^s \eta_r\,\mathrm{d}r)} = \tfrac{X_{\lfloor s\rfloor_N}-Y_{\lfloor s\rfloor_N}}{\exp(\int_0^{\lfloor s\rfloor_N} \eta_r\,\mathrm{d}r)} + \int_{\lfloor s\rfloor_N}^s \tfrac{f(X_r)-a(r)}{\exp(\int_0^r \eta_\iota\,\mathrm{d}\iota)}\,\mathrm{d}r$$
$$+ \int_{\lfloor s\rfloor_N}^s \tfrac{g(X_r)-b(r)}{\exp(\int_0^r \eta_\iota\,\mathrm{d}\iota)}\,\mathrm{d}W_r + \int_{\lfloor s\rfloor_N}^s \tfrac{(X_r-Y_r)(-\eta_r)}{\exp(\int_0^r \eta_\iota\,\mathrm{d}\iota)}\,\mathrm{d}r. \tag{3.31}$$

As a consequence,

$$
\begin{aligned}
&\Big\| \sup_{t\in[0,u]} \int_0^{t\wedge\tau_N} 2\Big\langle \tfrac{X_s-Y_s}{\exp(\int_0^s \eta_r\,\mathrm{d}r)}, \int_{\lfloor s\rfloor_N}^s \big\langle f'(Y_r), b(r)\,\mathrm{d}W_r\big\rangle\Big\rangle \mathrm{d}s\Big\|_{L^{p/2}(\Omega;\mathbb{R})}\\
&\le \Big\| \sup_{t\in[0,u]} \int_0^{t\wedge\tau_N} 2\Big\langle \tfrac{X_{\lfloor s\rfloor_N}-Y_{\lfloor s\rfloor_N}}{\exp(\int_0^{\lfloor s\rfloor_N} \eta_r\,\mathrm{d}r)}, \int_{\lfloor s\rfloor_N}^s \big\langle f'(Y_r), b(r)\,\mathrm{d}W_r\big\rangle\Big\rangle \mathrm{d}s\Big\|_{L^{p/2}(\Omega;\mathbb{R})}\\
&\quad+\Big\| \sup_{t\in[0,u]} \int_0^{t\wedge\tau_N} 2\Big\langle \int_{\lfloor s\rfloor_N}^s \tfrac{f(X_r)-a(r)}{\exp(\int_0^r \eta_\iota\,\mathrm{d}\iota)}\,\mathrm{d}r, \int_{\lfloor s\rfloor_N}^s \big\langle f'(Y_r), b(r)\,\mathrm{d}W_r\big\rangle\Big\rangle \mathrm{d}s\Big\|_{L^{p/2}(\Omega;\mathbb{R})}\\
&\quad+\Big\| \sup_{t\in[0,u]} \int_0^{t\wedge\tau_N} 2\Big\langle \int_{\lfloor s\rfloor_N}^s \tfrac{g(X_r)-b(r)}{\exp(\int_0^r \eta_\iota\,\mathrm{d}\iota)}\,\mathrm{d}W_r, \int_{\lfloor s\rfloor_N}^s \big\langle f'(Y_r), b(r)\,\mathrm{d}W_r\big\rangle\Big\rangle \mathrm{d}s\Big\|_{L^{p/2}(\Omega;\mathbb{R})}\\
&\quad+\Big\| \sup_{t\in[0,u]} \int_0^{t\wedge\tau_N} 2\Big\langle \int_{\lfloor s\rfloor_N}^s \tfrac{(X_r-Y_r)(-\eta_r)}{\exp(\int_0^r \eta_\iota\,\mathrm{d}\iota)}\,\mathrm{d}r, \int_{\lfloor s\rfloor_N}^s \big\langle f'(Y_r), b(r)\,\mathrm{d}W_r\big\rangle\Big\rangle \mathrm{d}s\Big\|_{L^{p/2}(\Omega;\mathbb{R})}\\
&=: \widetilde{B}_1+\widetilde{B}_2+\widetilde{B}_3+\widetilde{B}_4.
\end{aligned}
\tag{3.32}
$$

Let us estimate these four items in (3.32) separately. We first split $\widetilde{B}_1$ into two parts:

$$
\begin{aligned}
\widetilde{B}_1 &\le \Big\| \sup_{t\in[0,u]} \Big| \sum_{k=0}^{n_t-1} \int_{t_k}^{t_{k+1}} 2\mathbb{1}_{s\le\tau_N} \Big\langle \tfrac{X_{\lfloor s\rfloor_N}-Y_{\lfloor s\rfloor_N}}{\exp(\int_0^{\lfloor s\rfloor_N} \eta_r\,\mathrm{d}r)}, \int_{\lfloor s\rfloor_N}^s \big\langle f'(Y_r), b(r)\,\mathrm{d}W_r\big\rangle\Big\rangle \mathrm{d}s\Big|\Big\|_{L^{p/2}(\Omega;\mathbb{R})}\\
&\quad+\Big\| \sup_{t\in[0,u]} \int_{\lfloor t\rfloor_N}^{t} 2\mathbb{1}_{s\le\tau_N} \Big\langle \tfrac{X_{\lfloor s\rfloor_N}-Y_{\lfloor s\rfloor_N}}{\exp(\int_0^{\lfloor s\rfloor_N} \eta_r\,\mathrm{d}r)}, \int_{\lfloor s\rfloor_N}^s \big\langle f'(Y_r), b(r)\,\mathrm{d}W_r\big\rangle\Big\rangle \mathrm{d}s\Big\|_{L^{p/2}(\Omega;\mathbb{R})}\\
&=: \widetilde{B}_{11}+\widetilde{B}_{12},
\end{aligned}
\tag{3.33}
$$

where we denote $n_t := \lfloor t\rfloor_N/h$. By the condition (c) in Theorem 3.2, it follows that

$$
\zeta_n := \sum_{k=0}^{n-1} \int_{t_k}^{t_{k+1}} 2\mathbb{1}_{s\le\tau_N} \Big\langle \tfrac{X_{\lfloor s\rfloor_N}-Y_{\lfloor s\rfloor_N}}{\exp(\int_0^{\lfloor s\rfloor_N} \eta_r\,\mathrm{d}r)}, \int_{\lfloor s\rfloor_N}^s \big\langle f'(Y_r), b(r)\,\mathrm{d}W_r\big\rangle\Big\rangle \mathrm{d}s
$$

is a discrete martingale. The Doob discrete martingale inequality, Lemma 2.2, Hölder's inequality and the condition (c) in Theorem 3.2 imply that

$$
\begin{aligned}
\widetilde{B}_{11} &\le C_p \Big\| \sum_{k=0}^{n_u-1} \int_{t_k}^{t_{k+1}} 2\mathbb{1}_{s\le\tau_N} \Big\langle \tfrac{X_{\lfloor s\rfloor_N}-Y_{\lfloor s\rfloor_N}}{\exp(\int_0^{\lfloor s\rfloor_N} \eta_r\,\mathrm{d}r)}, \int_{\lfloor s\rfloor_N}^s \big\langle f'(Y_r), b(r)\,\mathrm{d}W_r\big\rangle\Big\rangle \mathrm{d}s\Big\|_{L^{p/2}(\Omega;\mathbb{R})}\\
&\le C_p \left(\sum_{k=0}^{n_u-1} \Big\| \int_{t_k}^{t_{k+1}} 2\mathbb{1}_{s\le\tau_N} \Big\langle \tfrac{X_{\lfloor s\rfloor_N}-Y_{\lfloor s\rfloor_N}}{\exp(\int_0^{\lfloor s\rfloor_N} \eta_r\,\mathrm{d}r)}, \int_{\lfloor s\rfloor_N}^s \big\langle f'(Y_r), b(r)\,\mathrm{d}W_r\big\rangle\Big\rangle \mathrm{d}s\Big\|^2_{L^{p/2}(\Omega;\mathbb{R})} \right)^{1/2}\\
&\le C_p \left(h \int_0^{\lfloor u\rfloor_N} \Big\| \tfrac{|X_{\lfloor s\rfloor_N\wedge\tau_N}-Y_{\lfloor s\rfloor_N\wedge\tau_N}|}{\exp(\int_0^{\lfloor s\rfloor_N\wedge\tau_N} \frac12\eta_r\,\mathrm{d}r)}\Big\|^2_{L^p(\Omega;\mathbb{R})} \Big\| \Big| \int_{\lfloor s\rfloor_N}^s \big\langle f'(Y_r), b(r)\,\mathrm{d}W_r\big\rangle \Big| \Big\|^2_{L^p(\Omega;\mathbb{R})} \mathrm{d}s \right)^{1/2}\\
&\le \sup_{s\in[0,u]} \Big\| \tfrac{|X_{s\wedge\tau_N}-Y_{s\wedge\tau_N}|}{\exp(\int_0^{s\wedge\tau_N} \frac12\eta_r\,\mathrm{d}r)}\Big\|_{L^p(\Omega;\mathbb{R})} C_p \left(h \int_0^{\lfloor u\rfloor_N} \Big\| \Big| \int_{\lfloor s\rfloor_N}^s \big\langle f'(Y_r), b(r)\,\mathrm{d}W_r\big\rangle \Big| \Big\|^2_{L^p(\Omega;\mathbb{R})} \mathrm{d}s \right)^{1/2}\\
&\le \tfrac18 \sup_{s\in[0,u]} \Big\| \tfrac{|X_{s\wedge\tau_N}-Y_{s\wedge\tau_N}|}{\exp(\int_0^{s\wedge\tau_N} \frac12\eta_r\,\mathrm{d}r)}\Big\|^2_{L^p(\Omega;\mathbb{R})} + Ch^2.
\end{aligned}
\tag{3.34}
$$

With the help of Hölder's inequality, an elementary inequality and the condition (c) in Theorem 3.2, for $p \geq 4$ one can estimate $\widetilde{B}_{12}$ as

$$\begin{aligned}
\widetilde{B}_{12} &\leq \Bigg(\mathbb{E}\Bigg[\sup_{t\in[0,u]}\Bigg|\int_{\lfloor t\rfloor_N}^{t} 2\mathbb{1}_{s\leq\tau_N}\Big\langle \tfrac{X_{\lfloor s\rfloor_N}-Y_{\lfloor s\rfloor_N}}{\exp(\int_0^{\lfloor s\rfloor_N}\eta_r\,\mathrm{d}r)}, \int_{\lfloor s\rfloor_N}^{s}\big\langle f'(Y_r), b(r)\,\mathrm{d}W_r\big\rangle\Big\rangle\,\mathrm{d}s\Bigg|^{p/2}\Bigg]\Bigg)^{2/p} \\
&\leq Ch^{1-2/p}\Bigg(\mathbb{E}\Bigg[\int_0^u \mathbb{1}_{s\leq\tau_N}\Bigg|\Big\langle \tfrac{X_{\lfloor s\rfloor_N}-Y_{\lfloor s\rfloor_N}}{\exp(\int_0^{\lfloor s\rfloor_N}\eta_r\,\mathrm{d}r)}, \int_{\lfloor s\rfloor_N}^{s}\big\langle f'(Y_r), b(r)\,\mathrm{d}W_r\big\rangle\Big\rangle\Bigg|^{p/2}\mathrm{d}s\Bigg]\Bigg)^{2/p} \\
&\leq Ch^{1-2/p}\Bigg(\int_0^u \Bigg\|\tfrac{|X_{\lfloor s\rfloor_N\wedge\tau_N}-Y_{\lfloor s\rfloor_N\wedge\tau_N}|}{\exp(\int_0^{\lfloor s\rfloor_N\wedge\tau_N}\frac12\eta_r\,\mathrm{d}r)}\Bigg\|_{L^p(\Omega;\mathbb{R})}^{p/2}\Bigg\|\Bigg|\int_{\lfloor s\rfloor_N}^{s}\big\langle f'(Y_r), b(r)\,\mathrm{d}W_r\big\rangle\Bigg|\Bigg\|_{L^p(\Omega;\mathbb{R})}^{p/2}\mathrm{d}s\Bigg)^{2/p} \\
&\leq \tfrac18\sup_{s\in[0,u]}\Bigg\|\tfrac{|X_{s\wedge\tau_N}-Y_{s\wedge\tau_N}|}{\exp(\int_0^{s\wedge\tau_N}\frac12\eta_r\,\mathrm{d}r)}\Bigg\|_{L^p(\Omega;\mathbb{R})}^{2} + Ch^{2-4/p}\Bigg(\int_0^u\Bigg\|\Bigg|\int_{\lfloor s\rfloor_N}^{s}\big\langle f'(Y_r), b(r)\,\mathrm{d}W_r\big\rangle\Bigg|\Bigg\|_{L^p(\Omega;\mathbb{R})}^{p/2}\mathrm{d}s\Bigg)^{4/p} \\
&\leq \tfrac18\sup_{s\in[0,u]}\Bigg\|\tfrac{|X_{s\wedge\tau_N}-Y_{s\wedge\tau_N}|}{\exp(\int_0^{s\wedge\tau_N}\frac12\eta_r\,\mathrm{d}r)}\Bigg\|_{L^p(\Omega;\mathbb{R})}^{2} + Ch^2.
\end{aligned} \tag{3.35}$$

Hence, one concludes that

$$\widetilde{B}_1 \leq \tfrac14\sup_{s\in[0,u]}\Bigg\|\tfrac{|X_{s\wedge\tau_N}-Y_{s\wedge\tau_N}|}{\exp(\int_0^{s\wedge\tau_N}\frac12\eta_r\,\mathrm{d}r)}\Bigg\|_{L^p(\Omega;\mathbb{R})}^{2} + Ch^2. \tag{3.36}$$

Similar to the estimate of B_2, one treat $\widetilde{B}_2$ as follows:

$$\begin{aligned}
\widetilde{B}_2 \leq{}& C\int_0^u\int_{\lfloor s\rfloor_N}^{s}\Bigg\|\mathbb{1}_{s\leq\tau_N}\tfrac{|X_r-Y_r|(1+|X_r|+|Y_r|)^{c_f}}{\exp(\int_0^r\eta_\iota\,\mathrm{d}\iota)}\Bigg|\int_{\lfloor s\rfloor_N}^{s}\big\langle f'(Y_\iota), b(\iota)\,\mathrm{d}W_\iota\big\rangle\Bigg|\Bigg\|_{L^{p/2}(\Omega;\mathbb{R})}\mathrm{d}r\,\mathrm{d}s \\
&+ C\int_0^u\int_{\lfloor s\rfloor_N}^{s}\Bigg\|\mathbb{1}_{s\leq\tau_N}\tfrac{|f(Y_r)-f(Y_{\lfloor r\rfloor_N})|}{\exp(\int_0^r\eta_\iota\,\mathrm{d}\iota)}\Bigg|\int_{\lfloor s\rfloor_N}^{s}\big\langle f'(Y_\iota), b(\iota)\,\mathrm{d}W_\iota\big\rangle\Bigg|\Bigg\|_{L^{p/2}(\Omega;\mathbb{R})}\mathrm{d}r\,\mathrm{d}s \\
&+ C\int_0^u\int_{\lfloor s\rfloor_N}^{s}\Bigg\|\mathbb{1}_{s\leq\tau_N}\tfrac{|f(Y_{\lfloor r\rfloor_N})-a(r)|}{\exp(\int_0^r\eta_\iota\,\mathrm{d}\iota)}\Bigg|\int_{\lfloor s\rfloor_N}^{s}\big\langle f'(Y_\iota), b(\iota)\,\mathrm{d}W_\iota\big\rangle\Bigg|\Bigg\|_{L^{p/2}(\Omega;\mathbb{R})}\mathrm{d}r\,\mathrm{d}s \\
\leq{}& C\int_0^u\sup_{r\in[0,s]}\Bigg\|\tfrac{|X_{r\wedge\tau_N}-Y_{r\wedge\tau_N}|}{\exp(\int_0^{r\wedge\tau_N}\frac12\eta_\iota\,\mathrm{d}\iota)}\Bigg\|_{L^p(\Omega;\mathbb{R})}^{2}\mathrm{d}s + Ch^2 \\
&+ Ch^{1/2}\int_0^u\int_{\lfloor s\rfloor_N}^{s}\Big\|\mathbb{1}_{s\leq\tau_N}|f(Y_{\lfloor r\rfloor_N})-a(r)|\Big\|_{L^p(\Omega;\mathbb{R})}\mathrm{d}r\,\mathrm{d}s.
\end{aligned} \tag{3.37}$$

With regard to $\widetilde{B}_3$, we employ Hölder's inequality, Young's inequality and Lemma 2.1 to derive

$$
\begin{aligned}
\widetilde{B}_3 \leq{}& C\int_0^u \Big\| \int_{\lfloor s\rfloor_N}^s \mathbb{1}_{s\leq\tau_N} \tfrac{g(X_r)-g(Y_r)}{\exp(\int_0^r \eta_\iota\,\mathrm{d}\iota)}\,\mathrm{d}W_r \Big\|_{L^p(\Omega;\mathbb{R}^d)} \Big\| \int_{\lfloor s\rfloor_N}^s \big\langle f'(Y_r), b(r)\,\mathrm{d}W_r\big\rangle \Big\|_{L^p(\Omega;\mathbb{R}^d)} \mathrm{d}s \\
&+ C\int_0^u \Big\| \int_{\lfloor s\rfloor_N}^s \mathbb{1}_{s\leq\tau_N} \tfrac{g(Y_r)-b(r)}{\exp(\int_0^r \eta_\iota\,\mathrm{d}\iota)}\,\mathrm{d}W_r \Big\|_{L^p(\Omega;\mathbb{R}^d)} \Big\| \int_{\lfloor s\rfloor_N}^s \big\langle f'(Y_r), b(r)\,\mathrm{d}W_r\big\rangle \Big\|_{L^p(\Omega;\mathbb{R}^d)} \mathrm{d}s \\
\leq{}& C\int_0^u \Big\| \int_{\lfloor s\rfloor_N}^s \mathbb{1}_{s\leq\tau_N} \tfrac{g(X_r)-g(Y_r)}{\exp(\int_0^r \eta_\iota\,\mathrm{d}\iota)}\,\mathrm{d}W_r \Big\|_{L^p(\Omega;\mathbb{R}^d)}^{4/3} \mathrm{d}s + Ch^2 \\
&+ Ch^{1/2}\int_0^u \Big\| \int_{\lfloor s\rfloor_N}^s \mathbb{1}_{s\leq\tau_N} \tfrac{g(Y_r)-b(r)}{\exp(\int_0^r \eta_\iota\,\mathrm{d}\iota)}\,\mathrm{d}W_r \Big\|_{L^p(\Omega;\mathbb{R}^d)} \mathrm{d}s \\
\leq{}& C\int_0^u \Big(\int_{\lfloor s\rfloor_N}^s \Big\| \mathbb{1}_{s\leq\tau_N} \tfrac{|X_r-Y_r|}{\exp(\int_0^r \eta_\iota\,\mathrm{d}\iota)} \Big\|_{L^p(\Omega;\mathbb{R})}^2 \mathrm{d}r \Big)^{2/3} \mathrm{d}s + Ch^2 \\
&+ Ch^{1/2}\int_0^u \Big(\int_{\lfloor s\rfloor_N}^s \Big\| \mathbb{1}_{s\leq\tau_N} \big(g(Y_r)-b(r)\big) \Big\|_{L^p(\Omega;\mathbb{R}^{d\times m})}^2 \mathrm{d}r \Big)^{1/2} \mathrm{d}s \\
\leq{}& C\int_0^u \sup_{r\in[0,s]} \Big\| \tfrac{|X_{r\wedge\tau_N}-Y_{r\wedge\tau_N}|}{\exp(\int_0^{r\wedge\tau_N} \frac{1}{2}\eta_\iota\,\mathrm{d}\iota)} \Big\|_{L^p(\Omega;\mathbb{R})}^2 \mathrm{d}s + Ch^2 \\
&+ Ch^{1/2}\int_0^u \Big(\int_{\lfloor s\rfloor_N}^s \Big\| \mathbb{1}_{s\leq\tau_N} \big(g(Y_r)-b(r)\big) \Big\|_{L^p(\Omega;\mathbb{R}^{d\times m})}^2 \mathrm{d}r \Big)^{1/2} \mathrm{d}s.
\end{aligned}
\tag{3.38}
$$

Similar to the estimate of B_3, we bound $\widetilde{B}_4$ in the following way:

$$
\begin{aligned}
\widetilde{B}_4 \leq{}& C\Big\| \int_0^u \int_{\lfloor s\rfloor_N}^s \mathbb{1}_{s\leq\tau_N} \tfrac{|X_r-Y_r|\eta_r}{\exp(\int_0^r \eta_\iota\,\mathrm{d}\iota)} \Big| \int_{\lfloor s\rfloor_N}^s \big\langle f'(Y_\iota), b(\iota)\,\mathrm{d}W_\iota\big\rangle \Big| \mathrm{d}r\,\mathrm{d}s \Big\|_{L^{p/2}(\Omega;\mathbb{R})} \\
\leq{}& C\int_0^u \int_{\lfloor s\rfloor_N}^s \Big\| \mathbb{1}_{s\leq\tau_N} \tfrac{|X_r-Y_r|}{\exp(\int_0^r \eta_\iota\,\mathrm{d}\iota)} \Big\|_{L^p(\Omega;\mathbb{R})} \Big\| \eta_r \Big| \int_{\lfloor s\rfloor_N}^s \big\langle f'(Y_\iota), b(\iota)\,\mathrm{d}W_\iota\big\rangle \Big| \Big\|_{L^p(\Omega;\mathbb{R})} \mathrm{d}r\,\mathrm{d}s \\
\leq{}& C\int_0^u \sup_{r\in[0,s]} \Big\| \tfrac{|X_{s\wedge\tau_N}-Y_{s\wedge\tau_N}|}{\exp(\int_0^{s\wedge\tau_N} \frac{1}{2}\eta_\iota\,\mathrm{d}\iota)} \Big\|_{L^p(\Omega;\mathbb{R})}^2 \mathrm{d}s + Ch^2.
\end{aligned}
\tag{3.39}
$$

Putting (3.30), (3.36), (3.37), (3.38) and (3.39) together yields

$$
\begin{aligned}
\mathbb{T}_2 \leq{}& C\int_0^u \sup_{s\in[0,t]} \Big\| \tfrac{|X_{s\wedge\tau_N}-Y_{s\wedge\tau_N}|}{\exp(\int_0^{s\wedge\tau_N} \frac{1}{2}\eta_r\,\mathrm{d}r)} \Big\|_{L^p(\Omega;\mathbb{R})}^2 \mathrm{d}t + \int_0^u \Big\| \mathbb{1}_{s\leq\tau_N} \big|f(Y_{\lfloor s\rfloor_N}) - a(s)\big| \Big\|_{L^p(\Omega;\mathbb{R})}^2 \mathrm{d}s \\
&+ \frac{1}{4} \sup_{s\in[0,u]} \Big\| \tfrac{|X_{s\wedge\tau_N}-Y_{s\wedge\tau_N}|}{\exp(\int_0^{s\wedge\tau_N} \frac{1}{2}\eta_r\,\mathrm{d}r)} \Big\|_{L^p(\Omega;\mathbb{R})}^2 + Ch^{1/2}\int_0^u \int_{\lfloor s\rfloor_N}^s \Big\| \mathbb{1}_{r\leq\tau_N} \big|f(Y_{\lfloor r\rfloor_N}) - a(r)\big| \Big\|_{L^p(\Omega;\mathbb{R})} \mathrm{d}r\,\mathrm{d}s \\
&+ Ch^{1/2}\int_0^u \Big(\int_{\lfloor s\rfloor_N}^s \Big\| \mathbb{1}_{r\leq\tau_N} \big(g(Y_r)-b(r)\big) \Big\|_{L^p(\Omega;\mathbb{R}^{d\times m})}^2 \mathrm{d}r \Big)^{1/2} \mathrm{d}s + Ch^2.
\end{aligned}
\tag{3.40}
$$

Then the proof is thus completed by combining (3.28), (3.29), (3.40) and Gronwall's inequality. □

In the rest of this article, we concentrate on applications of the previously obtained perturbation estimates to identify the order-one strong convergence of numerical methods for SDEs with non-globally monotone coefficients.

4. Order-one pathwise uniformly strong convergence of the SITEM scheme

In order to numerically solve SDEs (3.1) on a uniform grid $\{t_n = nh\}_{0\le n\le N}$ with stepsize $h = \frac{T}{N}$, a class of stopped increment-tamed EM (SITEM) method was proposed in Hutzenthaler *et al.* (2018):

$$Y_t = Y_{t_n} + \mathbb{1}_{|Y_{t_n}|<\exp(|\ln(h)|^{1/2})}\left[\frac{f(Y_{t_n})(t-t_n)+g(Y_{t_n})(W_t-W_{t_n})}{1+|f(Y_{t_n})(t-t_n)+g(Y_{t_n})(W_t-W_{t_n})|^{\delta}}\right],\ Y_0 = X_0,\ \delta \ge 2,\ t \in [t_n, t_{n+1}], \quad (4.1)$$

which was shown to inherit exponential integrability properties of original SDEs. Combining the perturbation theory obtained in Hutzenthaler & Jentzen (2020) with exponential integrability properties of both numerical solution and exact solution, the authors of Hutzenthaler & Jentzen (2020) successfully identified the order $\frac{1}{2}$ strong convergence of the SITEM method. An interesting question arises as to whether higher convergence rate than order $\frac{1}{2}$ can be obtained when high-order (e.g., Milstein-type) schemes are used. This is also expected by Hutzenthaler & Jentzen (2020) (see Remark 3.1 therein). Unfortunately, following (Hutzenthaler & Jentzen, 2020, Theorem 1.2), the convergence rates of any schemes would not exceed order $\frac{1}{2}$, which is nothing but the order of the Hölder regularity of the approximation process. In the present section, we aim to fill this gap and reveal order-one strong convergence of the SITEM method for some particular SDEs with non-globally monotone coefficients, for which the Euler type method coincides with the Milstein method and thus the order-one convergence is expected. To begin with, we define a stopping time $\tau_N^e : \Omega \to \{t_0, t_1, ..., t_N\}$ as

$$\tau_N^e := \inf\Big\{\{T\} \cup \{t \in t_0, ..., t_N : |Y_t| \ge \exp(|\ln(h)|^{1/2})\}\Big\}. \quad (4.2)$$

Equipped with the stopping time, we can introduce the continuous version of (4.1) as

$$Y_t = X_0 + \int_0^t \mathbb{1}_{s<\tau_N^e} a(s)\,\mathrm{d}s + \int_0^t \mathbb{1}_{s<\tau_N^e} b(s)\,\mathrm{d}W_s. \quad (4.3)$$

Here, for $s \in [t_k, t_{k+1})$, $a(s)$ and $b(s)$ are given by

$$a(s) := \psi^{[1]}(Z_s)f(Y_{t_k}) + \frac{1}{2}\sum_{j=1}^{m}\psi^{[2]}(Z_s)\Big(g(Y_{t_k})e_j, g(Y_{t_k})e_j\Big), \quad b(s) := \psi^{[1]}(Z_s)g(Y_{t_k}), \quad (4.4)$$

where $e_1 = (1, ..., 0)^*, ..., e_m = (0, ..., 1)^*$ are the Euclidean orthonormal basis of $\mathbb{R}^m$,

$$Z_s := f(Y_{t_k})(s - t_k) + g(Y_{t_k})(W_s - W_{t_k}) \quad (4.5)$$

and for a fixed $\delta \ge 2$,

$$\psi(x) := x(1 + |x|^{\delta})^{-1}, x \in \mathbb{R}^d. \quad (4.6)$$

By (Hutzenthaler *et al.*, 2018, Theorem 2.9) and in the notation of (2.5), for any $z, u \in \mathbb{R}^d$ we have

$$\psi^{[1]}(z)u = \begin{cases} u & : z = 0, \\ \frac{u}{1+|z|^\delta} - \frac{\delta z|z|^{(\delta-2)}\langle z,u\rangle}{(1+|z|^\delta)^2} & : z \neq 0, \end{cases} \tag{4.7}$$

and

$$\psi^{[2]}(z)(u,u) = \begin{cases} 0 & : z = 0, \\ \frac{2\delta^2|z|^{2(\delta-2)}z|\langle z,u\rangle|^2}{(1+|z|^\delta)^3} - \frac{\delta|z|^{(\delta-2)}\left[2u\langle z,u\rangle+z|u|^2\right]}{(1+|z|^\delta)^2} & \\ -\frac{\delta(\delta-2)|z|^{(\delta-4)}z|\langle z,u\rangle|^2}{(1+|z|^\delta)^2} & : z \neq 0. \end{cases} \tag{4.8}$$

Moreover, one can show the following properties of $\psi^{[1]}, \psi^{[2]}$, which are needed in the error analysis.

LEMMA 4.1. Let $\delta \geq 2$ and let ψ be defined by (4.6). Then for all $x \in \mathbb{R}^d$,

$$\begin{aligned} &\|\psi^{[1]}(x)\|_{L(\mathbb{R}^d,\mathbb{R}^d)} \leq 1 + \tfrac{\delta}{4}, \; \|\psi^{[1]}(x) - I\|_{L(\mathbb{R}^d,\mathbb{R}^d)} \leq \left[(1 + \tfrac{\delta}{4}) \wedge (\delta + 1)|x|^\delta\right], \\ &\sup_{u\in\mathbb{R}^d,|u|\leq 1} |\psi^{[2]}(x)(u,u)| \leq \left[(3\delta^2 + \delta) \wedge (3\delta^2 + \delta)|x|^{\delta-1}\right]. \end{aligned} \tag{4.9}$$

Proof. By (4.7) and (4.8), it is clear that

$$\begin{aligned} \|\psi^{[1]}(x)\|_{L(\mathbb{R}^d,\mathbb{R}^d)} &\leq 1 \vee \left(\tfrac{1}{1+|x|^\delta} + \tfrac{\delta|x|^\delta}{(1+|x|^\delta)^2}\right) \leq 1 + \tfrac{\delta}{4}, \\ \|\psi^{[1]}(x) - I\|_{L(\mathbb{R}^d,\mathbb{R}^d)} &\leq \left(\tfrac{|x|^\delta}{1+|x|^\delta} + \tfrac{\delta|x|^\delta}{(1+|x|^\delta)^2}\right) \leq \left[(1 + \tfrac{\delta}{4}) \wedge \big(\delta + 1\big)|x|^\delta\right], \\ \sup_{u\in\mathbb{R}^d,|u|\leq 1} |\psi^{[2]}(x)(u,u)| &\leq \tfrac{2\delta^2|x|^{2\delta-1}}{(1+|x|^\delta)^3} + \tfrac{(\delta^2+\delta)|x|^{\delta-1}}{(1+|x|^\delta)^2} \leq \left[(3\delta^2 + \delta) \wedge (3\delta^2 + \delta)|x|^{\delta-1}\right]. \end{aligned} \tag{4.10}$$

Now we are ready to state the main convergence result of this section.

THEOREM 4.2. Let $f : \mathbb{R}^d \to \mathbb{R}^d, g : \mathbb{R}^d \to \mathbb{R}^{d\times m}$ be measurable functions and let $f \in C^2(\mathbb{R}^d, \mathbb{R}^d)$. Let $f \in \mathcal{C}^1_{\mathcal{P}}(\mathbb{R}^d, \mathbb{R}^d)$ and let $g \in C^2(\mathbb{R}^d, \mathbb{R}^{d\times m})$ be Lipschitz satisfying, for all $k_1, k_2 \in \{1, ..., d\}, j_1, j_2 \in \{1, ..., m\}$,

$$\frac{\partial g^{(k_2,j_2)}}{\partial x_{k_1}} g^{(k_1,j_1)} = 0. \tag{4.11}$$

Let $U_0 \in \mathcal{C}^3_{\mathcal{D}}\big(\mathbb{R}^d, [0,\infty)\big)$ and $U_1 \in \mathcal{C}^1_{\mathcal{P}}\big(\mathbb{R}^d, [0,\infty)\big)$. Let a class of SITEM methods be defined by (4.3) with $\delta \geq 3$ and let $c, v, T \in (0,\infty), q, q_1, q_2 \in (0,\infty], \alpha \in [0,\infty), p \geq 4$. For all $x, y \in \mathbb{R}^d$, assume additionally that

(1) there exist constants $L, \kappa \geq 0$ such that for any $i = 1, ..., d, j = 1, ..., m$,

$$\left\| \mathrm{Hess}_x(f^{(i)}(x)) \right\| \bigvee \left\| \mathrm{Hess}_x(g^{(ij)}(x)) \right\| \leq L(1 + |x|)^\kappa;$$

(2) $|x|^{1/c} \leq c(1 + U_0(x))$ and $\mathbb{E}\left[e^{U_0(X_0)}\right] < \infty$;

(3) $(\mathcal{A}_{f,g}U_0)(x) + \frac{1}{2}|g(x)^*(\nabla U_0(x))|^2 + U_1(x) \leq c + \alpha U_0(x)$;

(4) $\langle x - y, f(x) - f(y)\rangle \leq \left[c + \frac{U_0(x)+U_0(y)}{2q_1 Te^{\alpha T}} + \frac{U_1(x)+U_1(y)}{2q_2 e^{\alpha T}}\right]|x - y|^2$.

Then for $\frac{1}{q} = \frac{1}{q_1} + \frac{1}{q_2}, \frac{1}{\nu} = \frac{1}{p} + \frac{1}{q}$ the approximation (4.3) used to solve (3.1) admits

$$\left\| \sup_{t\in[0,T]} |X_t - Y_t| \right\|_{L^\nu(\Omega;\mathbb{R})} \leq Ch,\ h \to 0. \tag{4.12}$$

If the condition (4) in Theorem 4.2 is replaced by the following one:

(4') for any $\eta > 0$, there exists a constant K_η such that

$$\langle x - y, f(x) - f(y)\rangle \leq \left[K_\eta + \eta\big(U_0(x) + U_0(y) + U_1(x) + U_1(y)\big)\right]|x - y|^2, \tag{4.13}$$

then for any $\nu > 0$ we have

$$\left\| \sup_{t\in[0,T]} |X_t - Y_t| \right\|_{L^\nu(\Omega;\mathbb{R})} \leq Ch,\ h \to 0. \tag{4.14}$$

Before we come to the proof, let us give some comments on functions U_0, U_1 and some parameters used in the above theorem. We first mention that, similar conditions have been used in Hutzenthaler & Jentzen (2020). Here the non-negative function U_0 plays a role of Lyapunov function for (stochastic) differential equations (see conditions (2), (3)). From conditions (4) or (4'), one observes that U_0 is also used to control the growth of the derivative of the drift f. For some models such as the Brownian dynamics, stochastic van der Pol oscillator and stochastic Duffing–van der Pol oscillator, U_0 alone is, however, not able to control the growth and an additional non-negative function U_1 is introduced in conditions (3), (4). As shown later, different models require different choices of functions U_0, U_1 such that conditions (2), (3), (4) or (4') in Theorem 4.2 are all satisfied. In view of Lemma 4.1, we require the method parameter $\delta \geq 3$ to guarantee the convergence order $1 \wedge \frac{\delta-1}{2} \geq 1$. In addition, the parameters q, q_1, q_2 and ν, p, q are two sets of conjugate numbers for the use of Hölder's inequality. Now we start the proof.

Proof of Theorem 4.2. The proof relies on the use of Theorem 3.2 and in what follows we check all the conditions there. First, let $\tau_N = \tau_N^e$ and it is obvious that for $s \in [0, T]$

$$\{s \leq \tau_N^e\} = \begin{cases} \{\lfloor s\rfloor_N \leq \tau_N^e\}, \lfloor s\rfloor_N = s; \\ \{\lfloor s\rfloor_N < \tau_N^e\}, \lfloor s\rfloor_N < s, \end{cases} \tag{4.15}$$

which implies $\{s \leq \tau_N^e\} \in \mathcal{F}_{\lfloor s \rfloor_N}$. By the virtue of the condition (4), Hölder's inequality, Jensen's inequality (see (Cox *et al.*, 2024, Lemma 2.22) and the fact that $U_0(x), U_1(x) \geq 0$ one derives that

$$
\begin{aligned}
&\left\| \exp\left(\int_0^{\tau_N^e} \left[\frac{\langle X_s - Y_s, f(X_s) - f(Y_s)\rangle + \frac{1+\varepsilon}{2}\|g(X_s) - g(Y_s)\|^2}{|X_s - Y_s|^2} \right]^+ \mathrm{d}s \right) \right\|_{L^q(\Omega;\mathbb{R})} \\
&\leq C_{\varepsilon,\alpha,c,q_1,q_2,T} \left\| \exp\left(\int_0^{\tau_N^e} \left[\frac{U_0(X_s)+U_0(Y_s)}{2q_1 T e^{\alpha T}} + \frac{U_1(X_s)+U_1(Y_s)}{2q_2 e^{\alpha T}} \right] \mathrm{d}s \right) \right\|_{L^q(\Omega;\mathbb{R})} \\
&\leq C \left(\mathbb{E}\left[\exp\left(\int_0^{\tau_N^e} \frac{U_0(X_s)+U_0(Y_s)}{2Te^{\alpha T}} \mathrm{d}s \right) \right] \right)^{1/q_1} \cdot \left(\mathbb{E}\left[\exp\left(\int_0^{\tau_N^e} \frac{U_1(X_s)+U_1(Y_s)}{2e^{\alpha T}} \mathrm{d}s \right) \right] \right)^{1/q_2} \\
&\leq C \sup_{s\in[0,T]} \left(\mathbb{E}\left[\exp\left(\frac{U_0(X_s)}{e^{\alpha s}} \right) \right] \right)^{\frac{1}{2q_1}} \cdot \sup_{s\in[0,T]} \left(\mathbb{E}\left[\exp\left(\frac{U_0(Y_s)}{e^{\alpha s}} \right) \right] \right)^{\frac{1}{2q_1}} \\
&\quad \cdot \sup_{s\in[0,T]} \left(\mathbb{E}\left[\exp\left(\int_0^{s\wedge\tau_N^e} \frac{U_1(X_u)}{e^{\alpha u}} \mathrm{d}u \right) \right] \right)^{\frac{1}{2q_2}} \cdot \sup_{s\in[0,T]} \left(\mathbb{E}\left[\exp\left(\int_0^{s\wedge\tau_N^e} \frac{U_1(Y_u)}{e^{\alpha u}} \mathrm{d}u \right) \right] \right)^{\frac{1}{2q_2}} \\
&\leq C \sup_{s\in[0,T]} \left(\mathbb{E}\left[\exp\left(\frac{U_0(X_s)}{e^{\alpha s}} + \int_0^{s} \frac{U_1(X_u)}{e^{\alpha u}} \mathrm{d}u \right) \right] \right)^{\frac{1}{2q_1}} \cdot \sup_{s\in[0,T]} \left(\mathbb{E}\left[\exp\left(\frac{U_0(Y_s)}{e^{\alpha s}} + \int_0^{s\wedge\tau_N^e} \frac{U_1(Y_u)}{e^{\alpha u}} \mathrm{d}u \right) \right] \right)^{\frac{1}{2q_1}} \\
&\quad \cdot \sup_{s\in[0,T]} \left(\mathbb{E}\left[\exp\left(\frac{U_0(X_s)}{e^{\alpha s}} + \int_0^{s} \frac{U_1(X_u)}{e^{\alpha u}} \mathrm{d}u \right) \right] \right)^{\frac{1}{2q_2}} \cdot \sup_{s\in[0,T]} \left(\mathbb{E}\left[\exp\left(\frac{U_0(Y_s)}{e^{\alpha s}} + \int_0^{s\wedge\tau_N^e} \frac{U_1(Y_u)}{e^{\alpha u}} \mathrm{d}u \right) \right] \right)^{\frac{1}{2q_2}} \\
&\leq C \sup_{s\in[0,T]} \left(\mathbb{E}\left[\exp\left(\frac{U_0(X_s)}{e^{\alpha s}} + \int_0^{s} \frac{U_1(X_u)}{e^{\alpha u}} \mathrm{d}u \right) \right] \right)^{\frac{1}{2q}} \cdot \sup_{s\in[0,T]} \left(\mathbb{E}\left[\exp\left(\frac{U_0(Y_s)}{e^{\alpha s}} + \int_0^{s\wedge\tau_N^e} \frac{U_1(Y_u)}{e^{\alpha u}} \mathrm{d}u \right) \right] \right)^{\frac{1}{2q}} \\
&< \infty.
\end{aligned}
\tag{4.16}
$$

Here the last inequality stands due to the exponential integrability property for both exact solution $\{X_s\}_{s\in[0,T]}$ and numerical solution $\{Y_s\}_{s\in[0,T]}$ (see (Cox *et al.*, 2024, Corollary 2.4) and (Hutzenthaler *et al.*, 2018, Corollary 2.10)). For any $p \geq 4$, by (4.16) and noting $U_1 \in \mathcal{C}^1_{\mathcal{P}}(\mathbb{R}^d, [0,\infty))$ and $|x|^{1/c} \leq c(1 + U_0(x))$, we get

$$
\begin{aligned}
&\sup_{s\in[0,T]} \left\| \mathbb{1}_{s\leq\tau_N^e} \left[\frac{\langle X_s - Y_s, f(X_s) - f(Y_s)\rangle + \frac{1+\varepsilon}{2}\|g(X_s) - g(Y_s)\|^2}{|X_s - Y_s|^2} \right]^+ \right\|_{L^{3p}(\Omega;\mathbb{R})} \\
&\leq \sup_{s\in[0,T]} \left\| \left[C + \frac{U_0(X_s)+U_0(Y_s)}{2q_1 T e^{\alpha T}} + \frac{U_1(X_s)+U_1(Y_s)}{2q_2 e^{\alpha T}} \right] \right\|_{L^{3p}(\Omega;\mathbb{R})} \\
&\leq C + C \Big[\sup_{s\in[0,T]} \|U_0(X_s)\|_{L^{3p}(\Omega;\mathbb{R})} + \sup_{s\in[0,T]} \|U_1(X_s)\|_{L^{3p}(\Omega;\mathbb{R})} \\
&\quad + \sup_{s\in[0,T]} \|U_0(Y_s)\|_{L^{3p}(\Omega;\mathbb{R})} + \sup_{s\in[0,T]} \|U_1(Y_s)\|_{L^{3p}(\Omega;\mathbb{R})} \Big] < \infty,
\end{aligned}
\tag{4.17}
$$

which confirms condition (b) in Theorem 3.2. By (4.16) and condition (1) in Theorem 4.2,

$$\sup_{s\in[0,T]} \|X_s\|_{L^{6pc_g\vee 3pc_f\vee 3p}(\Omega;\mathbb{R}^d)} \bigvee \sup_{s\in[0,T]} \|Y_s\|_{L^{6pc_g\vee 3pc_f}(\Omega;\mathbb{R}^d)} < \infty \tag{4.18}$$

and for any $i=1,...d$, $\sup_{s\in[0,T]} \|\operatorname{Hess}_x(f^{(i)}(Y_s))\|_{L^{3p}(\Omega;\mathbb{R}^{d\times d})} < \infty$. By Lemma 4.1,

$$\sup_{s\in[0,T]} \|\mathbb{1}_{s<\tau_N^e} a(s)\|_{L^{3p}(\Omega;\mathbb{R}^d)} \bigvee \sup_{s\in[0,T]} \|\mathbb{1}_{s<\tau_N^e} b(s)\|_{L^{3p}(\Omega;\mathbb{R}^{d\times m})} < \infty. \tag{4.19}$$

This verifies condition (c) in Theorem 3.2, which in turn implies

$$\begin{aligned}
\Big\| \sup_{t\in[0,T]} |X_{t\wedge\tau_N^e} - Y_{t\wedge\tau_N^e}| \Big\|_{L^\nu(\Omega;\mathbb{R})} &\le C\Big[h^2 + h^{\frac{1}{2}} \int_0^T \int_{\lfloor s\rfloor_N}^s \Big\| \mathbb{1}_{r\le\tau_N^e} |f(Y_{\lfloor r\rfloor_N}) - \mathbb{1}_{r<\tau_N^e} a(r)| \Big\|_{L^p(\Omega;\mathbb{R})} \mathrm{d}r\, \mathrm{d}s \\
&+ \int_0^T \Big\| \mathbb{1}_{s\le\tau_N^e} \|g(Y_s) - \mathbb{1}_{s<\tau_N^e} b(s)\| \Big\|^2_{L^p(\Omega;\mathbb{R})} \mathrm{d}s + \int_0^T \Big\| \mathbb{1}_{s\le\tau_N^e} \big|f(Y_{\lfloor s\rfloor_N}) - \mathbb{1}_{s<\tau_N^e} a(s)\big| \Big\|^2_{L^p(\Omega;\mathbb{R})} \mathrm{d}s \\
&+ h^{\frac{1}{2}} \int_0^T \Big(\int_{\lfloor s\rfloor_N}^s \Big\| \mathbb{1}_{r\le\tau_N^e} \|g(Y_r) - \mathbb{1}_{r<\tau_N^e} b(r)\| \Big\|^2_{L^p(\Omega;\mathbb{R})} \mathrm{d}r \Big)^{\frac{1}{2}} \mathrm{d}s \Big]^{\frac{1}{2}}.
\end{aligned} \tag{4.20}$$

By the property of Lebesgue integral and Lemma 4.1, one can show

$$\begin{aligned}
&\int_0^T \Big\| \mathbb{1}_{s\le\tau_N^e} \big|f(Y_{\lfloor s\rfloor_N}) - \mathbb{1}_{s<\tau_N^e} a(s)\big| \Big\|^2_{L^p(\Omega;\mathbb{R})} \mathrm{d}s \\
&= \int_0^T \Big\| \mathbb{1}_{s<\tau_N^e} \big|f(Y_{\lfloor s\rfloor_N}) - a(s)\big| \Big\|^2_{L^p(\Omega;\mathbb{R})} \mathrm{d}s \\
&\le \int_0^T \Big\| \Big|f(Y_{\lfloor s\rfloor_N}) - \psi^{[1]}(Z_s) f(Y_{\lfloor s\rfloor_N}) - \tfrac{1}{2} \sum_{j=1}^m \psi^{[2]}(Z_s)\Big(g(Y_{\lfloor s\rfloor_N})e_j, g(Y_{\lfloor s\rfloor_N})e_j\Big) \Big| \Big\|^2_{L^p(\Omega;\mathbb{R})} \mathrm{d}s \\
&\le Ch^{\delta-1}.
\end{aligned} \tag{4.21}$$

Moreover,

$$\begin{aligned}
&\int_0^T \Big\| \mathbb{1}_{s\le\tau_N^e} \Big\| g(Y_s) - \mathbb{1}_{s<\tau_N^e} b(s) \Big\| \Big\|^2_{L^p(\Omega;\mathbb{R})} \mathrm{d}s \\
&\le C \int_0^T \Big\| \mathbb{1}_{s<\tau_N^e} \Big\| g(Y_s) - g(Y_{\lfloor s\rfloor_N}) \Big\| \Big\|^2_{L^p(\Omega;\mathbb{R})} + \Big\| \mathbb{1}_{s<\tau_N^e} \Big\| g(Y_{\lfloor s\rfloor_N}) - \psi^{[1]}(Z_s) g(Y_{\lfloor s\rfloor_N}) \Big\| \Big\|^2_{L^p(\Omega;\mathbb{R})} \mathrm{d}s \\
&\le C \int_0^T \Big\| \Big\| g(Y_s) - g(Y_{\lfloor s\rfloor_N}) \Big\| \Big\|^2_{L^p(\Omega;\mathbb{R})} \mathrm{d}s + Ch^3.
\end{aligned} \tag{4.22}$$

For $i = 1, ..., d, j = 1, ..., m$, by the Itô formula and recalling $\frac{\partial g^{(k_2,j_2)}}{\partial x_{k_1}} g^{(k_1,j_1)} = 0$, we arrive at

$$\begin{aligned}
&\int_0^T \left\| g^{(ij)}(Y_s) - g^{(ij)}(Y_{\lfloor s \rfloor_N}) \right\|^2_{L^p(\Omega;\mathbb{R})} \mathrm{d}s \\
&\leq C \int_0^T \left\| \int_{\lfloor s \rfloor_N}^s \left\langle g^{(ij)\prime}(Y_r), \mathbb{1}_{r<\tau_N^e} a(r) \right\rangle + \tfrac{1}{2} \operatorname{trace}\left(\mathbb{1}_{r<\tau_N^e} b(r)^* \operatorname{Hess}_x(g^{(ij)}(Y_r)) b(r) \right) \mathrm{d}r \right\|^2_{L^p(\Omega;\mathbb{R})} \mathrm{d}s \\
&\quad + C \int_0^T \left\| \int_{\lfloor s \rfloor_N}^s \left\langle g^{(ij)\prime}(Y_r), \mathbb{1}_{r<\tau_N^e} \psi^{[1]}(Z_r) g(Y_{\lfloor r \rfloor_N}) \,\mathrm{d}W_r \right\rangle \right\|^2_{L^p(\Omega;\mathbb{R})} \mathrm{d}s \\
&\leq C \int_0^T \left\| \int_{\lfloor s \rfloor_N}^s \left\langle g^{(ij)\prime}(Y_r) - g^{(ij)\prime}(Y_{\lfloor r \rfloor_N}), \mathbb{1}_{r<\tau_N^e} \psi^{[1]}(Z_r) g(Y_{\lfloor r \rfloor_N}) \,\mathrm{d}W_r \right\rangle \right\|^2_{L^p(\Omega;\mathbb{R})} \mathrm{d}s \\
&\quad + C \int_0^T \left\| \int_{\lfloor s \rfloor_N}^s \left\langle g^{(ij)\prime}(Y_{\lfloor r \rfloor_N}), \mathbb{1}_{r<\tau_N^e} \psi^{[1]}(Z_r) g(Y_{\lfloor r \rfloor_N}) \,\mathrm{d}W_r \right\rangle \right\|^2_{L^p(\Omega;\mathbb{R})} \mathrm{d}s + Ch^2 \\
&\leq C \int_0^T \left\| \int_{\lfloor s \rfloor_N}^s \left(g^{(ij)\prime}(Y_{\lfloor r \rfloor_N})^* \mathbb{1}_{r<\tau_N^e} \psi^{[1]}(Z_r) g(Y_{\lfloor r \rfloor_N}) - g^{(ij)\prime}(Y_{\lfloor r \rfloor_N})^* \mathbb{1}_{r<\tau_N^e} g(Y_{\lfloor r \rfloor_N}) \right) \mathrm{d}W_r \right\|^2_{L^p(\Omega;\mathbb{R})} \mathrm{d}s \\
&\quad + Ch^2 \\
&\leq Ch^2.
\end{aligned} \tag{4.23}$$

Therefore, one obtains

$$\int_0^T \left\| \mathbb{1}_{s \leq \tau_N^e} \left\| g(Y_s) - \mathbb{1}_{s<\tau_N^e} b(s) \right\| \right\|^2_{L^p(\Omega;\mathbb{R})} \mathrm{d}s \leq Ch^2. \tag{4.24}$$

The same arguments used in (4.21) and (4.24) can be applied to estimate the second and fifth terms on the right-hand side of (4.20). Hence, we deduce that for any $\frac{1}{v} = \frac{1}{p} + \frac{1}{q}$ and $\delta \geq 3$

$$\left\| \sup_{t\in[0,T]} |X_{t\wedge\tau_N^e} - Y_{t\wedge\tau_N^e}| \right\|_{L^v(\Omega;\mathbb{R})} \leq Ch. \tag{4.25}$$

Observe that

$$\begin{aligned}
&\left\| \sup_{t\in[0,T]} |X_t - Y_t| \right\|_{L^v(\Omega;\mathbb{R})} \\
&\leq \left\| \sup_{t\in[0,T]} \mathbb{1}_{\tau_N^e<t} |X_t - Y_t| \right\|_{L^v(\Omega;\mathbb{R})} + \left\| \sup_{t\in[0,T]} \mathbb{1}_{\tau_N^e \geq t} |X_t - Y_t| \right\|_{L^v(\Omega;\mathbb{R})} \\
&\leq \left\| \mathbb{1}_{\tau_N^e<T} \right\|_{L^{2v}(\Omega;\mathbb{R})} \left\| \sup_{t\in[0,T]} |X_t - Y_t| \right\|_{L^{2v}(\Omega;\mathbb{R})} + \left\| \sup_{t\in[0,T]} |X_{t\wedge\tau_N^e} - Y_{t\wedge\tau_N^e}| \right\|_{L^v(\Omega;\mathbb{R})}.
\end{aligned} \tag{4.26}$$

Using Lemma 2.1 ensures

$$
\begin{aligned}
&\Big\| \sup_{t\in[0,T]} |X_t - Y_t| \Big\|_{L^{2\nu}(\Omega;\mathbb{R})} \\
&\leq \Big\| \sup_{t\in[0,T]} \Big| \int_0^t \big(f(X_s) - \mathbb{1}_{s<\tau_N^e} a(s)\big)\,\mathrm{d}s \Big| + \sup_{t\in[0,T]} \Big| \int_0^t \big(g(X_s) - \mathbb{1}_{s<\tau_N^e} b(s)\big)\,\mathrm{d}W_s \Big| \Big\|_{L^{2\nu}(\Omega;\mathbb{R})} \\
&\leq C,
\end{aligned}
\tag{4.27}
$$

and by the condition (2) in Theorem 4.2, the Markov inequality and $-\frac{1}{4!}x^4 \geq -e^x,\ x \geq 0$ one infers

$$
\begin{aligned}
\mathbb{P}[\tau_N^e < T] &\leq \mathbb{P}\big[|Y_T| \geq \exp(|\ln(h)|^{1/2})\big] \\
&\leq \mathbb{P}\Big[\tfrac{1+U_0(Y_T)}{e^{\alpha T}} \geq \tfrac{1}{ce^{\alpha T}} \exp(\tfrac{1}{c}|\ln(h)|^{1/2})\Big] \\
&\leq \mathbb{E}\Big[\exp\big(\tfrac{1+U_0(Y_T)}{e^{\alpha T}}\big)\Big] \exp\big(-\tfrac{1}{ce^{\alpha T}} \exp(\tfrac{1}{c}|\ln(h)|^{1/2})\big) \\
&\leq C_1 \exp\big(-\tfrac{|\ln(h)|^2}{24c^5 e^{\alpha T}}\big).
\end{aligned}
\tag{4.28}
$$

For any $C_2 > 0$ and $h < 1$ being small enough, one knows $|\ln(h)|^2 \geq -\frac{1}{C_2}(2\nu \ln(h))$ and hence one gets for small $h < 1$,

$$
\exp\big(-\tfrac{|\ln(h)|^2}{24c^5 e^{\alpha T}}\big) \leq h^{2\nu},
$$

which validates (4.12). Finally, note that if (4.13) holds, then for any $\gamma > 0$,

$$
\Big\| \exp\Big(\int_0^{\tau_N^e} \Big[\tfrac{\langle X_s - Y_s, f(X_s) - f(Y_s)\rangle + \frac{1+\varepsilon}{2}\|g(X_s)-g(Y_s)\|^2}{|X_s - Y_s|^2} \Big]^+ \mathrm{d}s \Big) \Big\|_{L^\gamma(\Omega;\mathbb{R})} < \infty.
\tag{4.29}
$$

The assertion (4.14) can be acquired by repeating the above arguments, which finishes the proof. □

In what follows we employ Theorem 4.2 to obtain the first-order strong convergence of the time-stepping scheme (4.1) for SDE models without globally monotone coefficients, taken from Hutzenthaler *et al.* (2018); Hutzenthaler & Jentzen (2020). In the recent publication Hutzenthaler & Jentzen (2020), the authors derived only a convergence rate of order $\frac{1}{2}$ of the same scheme (4.1), even for the following additive noise driven SDE models and multiplicative noise driven second-order SDE models. Since the conditions of Theorem 4.2 are the same as those in (Hutzenthaler & Jentzen, 2020, Proposition 3.3), we just give the convergence results here and do not repeat the verification of the conditions. Indeed, one can refer to (Hutzenthaler & Jentzen, 2020, 3.1.2, 3.1.6, 3.1.7, 3.1.3, 3.1.4) and (Cox *et al.*, 2024, Chapter 4) for details on the verification of the conditions for the following different models. The initial value X_0 of the following models is assumed to be deterministic for simplicity.

Stochastic Lorenz equation with additive noise. Let $d = m = 3$ and $\alpha_1, \alpha_2, \alpha_3 \in [0, \infty)$. For $x = (x_1, x_2, x_3)^* \in \mathbb{R}^3$, we let

$$f(x) = (\alpha_1(x_2 - x_1), \alpha_2 x_1 - x_2 - x_1 x_3, x_1 x_2 - \alpha_3 x_3) \tag{4.30}$$

and let $g(\cdot)$ be a constant matrix. Moreover, we take $U_0(x) = |x|^2$ and $U_1(x) = 0$. Then all conditions in Theorem 4.2 are fulfilled. Therefore, using the SITEM method (4.3) with $\delta \geq 3$ to solve the above stochastic Lorenz equation with additive noise yields that for any $r > 0$, there exists a constant $C_r > 0$ such that

$$\Big\| \sup_{t \in [0,T]} |X_t - Y_t| \Big\|_{L^r(\Omega;\mathbb{R})} \leq C_r h. \tag{4.31}$$

Brownian dynamics. Let $d = m \geq 1, c, \beta \in (0, \infty)$ and $\theta \in [0, \frac{2}{\beta})$. Assume that $V \in \mathcal{C}^3_{\mathcal{D}}\big(\mathbb{R}^d, [0, \infty)\big) \cap C^3\big(\mathbb{R}^d, [0, \infty)\big), V', \mathrm{Hess}_x\left(V^{(i)}\right) \in \mathcal{C}^1_{\mathcal{P}}\big(\mathbb{R}^d, \mathbb{R}^{d\times d}\big), i = 1, ..., d$ and $\limsup_{r \searrow 0} \sup_{z \in \mathbb{R}^d} \frac{|z|^r}{1+V(z)} < \infty$. For $x \in \mathbb{R}^d$, we set

$$f(x) = -(\nabla V)(x), g(x) = \sqrt{\beta} I_{\mathbb{R}^d \times \mathbb{R}^d}. \tag{4.32}$$

This equation is also termed as the overdamped Langevin dynamics in literature. In addition, we suppose that $(\Delta V)(x) \leq c + cV(x) + \theta \|(\nabla V)(x)\|^2$ and for any $\eta > 0$

$$\sup_{x,y \in \mathbb{R}^d, x \neq y} \Big[\tfrac{\langle x-y, (\nabla V)(y) - (\nabla V)(x) \rangle}{|x-y|^2} - \eta\big(V(x) + V(y) + |(\nabla V)(x)|^2 + |(\nabla V)(y)|^2\big) \Big] < \infty.$$

Let $\nu \in (0, \frac{2}{\beta} - \theta)$, $U_0(x) = \nu V(x)$ and $U_1(x) = \nu(1 - \frac{\beta}{2}(\theta + \nu))|(\nabla V)(x)|^2$. Then all conditions in Theorem 4.2 are fulfilled. Therefore, applying the SITEM method (4.3) ($\delta \geq 3$) to the above Brownian dynamics yields that for any $r > 0$, there exists a constant $C_r > 0$ such that

$$\Big\| \sup_{t \in [0,T]} |X_t - Y_t| \Big\|_{L^r(\Omega;\mathbb{R})} \leq C_r h. \tag{4.33}$$

Langevin dynamics. Let $d = 2m \geq 1, \gamma \in (0, \infty)$ and $\beta \in (0, \infty)$. Assume that $V \in \mathcal{C}^3_{\mathcal{D}}\big(\mathbb{R}^m, [0, \infty)\big) \cap C^3\big(\mathbb{R}^m, [0, \infty)\big), V', \mathrm{Hess}_x\left(V^{(i)}\right) \in \mathcal{C}^1_{\mathcal{P}}\big(\mathbb{R}^m, \mathbb{R}^{m\times m}\big), i = 1, ..., m$ and $\limsup_{r \searrow 0} \sup_{z \in \mathbb{R}^m} \frac{|z|^r}{1+V(z)} < \infty$. For $x = (x_1, x_2)^* \in \mathbb{R}^{2m}, u \in \mathbb{R}^m$, we let

$$f(x) = (x_2, -(\nabla V)(x_1) - \gamma x_2), g(x)u = (0, \sqrt{\beta} u). \tag{4.34}$$

This equation is also termed as the underdamped Langevin dynamics in literature. In addition, we suppose that for any $\eta > 0$

$$\sup_{x,y \in \mathbb{R}^m, x \neq y} \Big[\tfrac{|(\nabla V)(x) - (\nabla V)(y)|}{|x-y|} - \eta\big(V(x) + V(y) + |x|^2 + |y|^2\big) \Big] < \infty.$$

Let $v \in (0,\infty)$, and for $x = (x_1, x_2)^* \in \mathbb{R}^{2m}$, $U_0(x) = \frac{v}{2}(|x_1|^2 + |x_2|^2) + vV(x_1)$, $U_1(x) = 0$. Then all conditions in Theorem 4.2 are fulfilled. Therefore, using the SITEM method (4.3)($\delta \geq 3$) for the Langevin dynamics yields that for any $r > 0$, there exists a constant $C_r > 0$ such that

$$\Big\| \sup_{t\in[0,T]} |X_t - Y_t| \Big\|_{L^r(\Omega;\mathbb{R})} \leq C_r h. \tag{4.35}$$

In Cui *et al.* (2022), the authors proposed a splitting averaged vector field (AVF) scheme for the Langevin dynamics (4.34). Equipped with the exponential integrability properties of the implicit approximations $\{Y_n\}_{0\leq n\leq N}$, the authors of Cui *et al.* (2022) spent a lot of efforts to analyze the pointwise strong error $\big(\sup_{0\leq n\leq N} \mathbb{E}\big[\|X_{t_n} - Y_{t_n}\|^p\big]\big)^{1/p}, p \geq 2$. As the first step, the pointwise strong convergence rate of order $\frac{1}{2}$ was obtained, which was later lifted to be order one by using technical arguments in the Malliavin calculus. Instead, we analyze the pathwise uniformly strong error of an explicit time-stepping scheme directly. By simply relying on the newly developed perturbation estimates in Section 3, we show a pathwise uniformly strong convergence rate of exact order one given by (4.35). It is worthwhile to mention that the authors of Cui *et al.* (2022) also proved the existence of the density function of the numerical solution produced by the splitting AVF scheme and provided the convergence rate of density functions for the scheme. Despite the same convergence rate, the splitting AVF scheme proposed by Cui *et al.* (2022), as an implicit one, is expected to be more numerically stable than the explicit SITEM scheme, particularly for large step-sizes.

Stochastic van der Pol oscillator. Let $d = 2, m \geq 1, c, \alpha \in (0,\infty)$ and $\gamma, \beta \in [0,\infty)$. For $x = (x_1, x_2)^* \in \mathbb{R}^2, u \in \mathbb{R}^m$, we let

$$f(x) = (x_2, (\gamma - \alpha x_1^2)x_2 - \beta x_1)^*, \; g(x)u = (0, \phi(x_1)u)^*, \tag{4.36}$$

where $\phi \in C^2(\mathbb{R}, \mathbb{R}^{1\times m})$ is a globally Lipschitz function. Let $v \in (0, \frac{\alpha}{2c})$, $U_0(x) = \frac{v}{2}|x|^2$ and $U_1(x) = v(\alpha - 2cv)(x_1x_2)^2$. Then all conditions in Theorem 4.2 are fulfilled. Therefore, applying the SITEM method (4.3) ($\delta \geq 3$) to the stochastic van der Pol oscillator yields that, for any $r > 0$, there exists a constant $C_r > 0$ such that

$$\Big\| \sup_{t\in[0,T]} |X_t - Y_t| \Big\|_{L^r(\Omega;\mathbb{R})} \leq C_r h. \tag{4.37}$$

Stochastic Duffing–van der Pol oscillator. Let $d = 2, m \geq 1, \alpha_1, \alpha_2 \in \mathbb{R}$ and $\alpha_3, c \in (0,\infty)$. For $x = (x_1, x_2)^* \in \mathbb{R}^2, u \in \mathbb{R}^m$, we let

$$f(x) = (x_2, \alpha_2 x_2 - \alpha_1 x_1 - \alpha_3 x_1^2 x_2 - x_1^3)^*, \; g(x)u = (0, \phi(x_1)u)^*, \tag{4.38}$$

where $\phi \in C^2(\mathbb{R}; \mathbb{R}^{1\times m})$ is a globally Lipschitz function. Let $v \in (0, \frac{\alpha_3}{c})$, $U_0(x) = \frac{v}{2}(\frac{x_1^4}{2} + x_2^2)$ and $U_1(x) = v(\alpha_3 - cv)(x_1x_2)^2$. Then all conditions in Theorem 4.2 are fulfilled. Therefore, applying the SITEM method (4.3) ($\delta \geq 3$) to the stochastic Duffing–van der Pol oscillator yields that, for any $r > 0$,

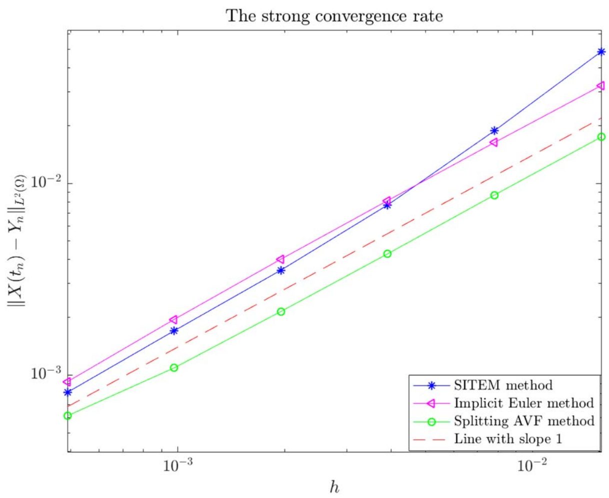


FIG. 1. A comparison of strong convergence rates for Langevin dynamics.

there exists a constant $C_r > 0$ such that

$$\left\| \sup_{t\in[0,T]} |X_t - Y_t| \right\|_{L^r(\Omega;\mathbb{R})} \leq C_r h. \tag{4.39}$$

Numerical experiments. Now let us present some numerical experiments to test, not only the strong convergence rate, but also the dynamic properties of the proposed method. We take the Langevin dynamics and the stochastic van der Pol oscillator as test examples. Let $T = 1, N = 2^k, k = 6, 7, ..., 11$ and regard the fine approximations with $h_{\text{exact}} = 2^{-14}$ as the 'true' solution. Also, we take $M = 5000$ Monte Carlo sample paths to approximate the expectation.

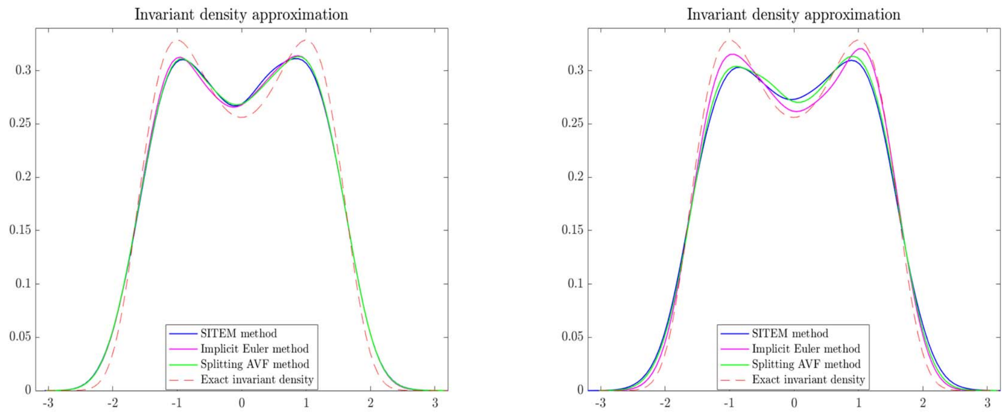
For the Langevin dynamics (4.34), we assign

$$m = 1, \nabla V(x) = x^3 - x, \gamma = 1, \beta = 2, X_0 = (1, 1)^*$$

and $\delta = 3$ for the SITEM method. Such a type of potential $V(x_1) = \frac{1}{4}x_1^4 - \frac{1}{2}x_1^2$ is called double-well potential. Fig. 1 displays the mean-square approximation errors of the SITEM method, the implicit splitting AVF method in Cui *et al.* (2022) and the implicit Euler method in Talay (2002). Numerical results show that, the three methods all have a strong convergence rate of order one and the splitting AVF method is slightly better in terms of computational error.

Furthermore, it is known that (see, e.g., Mattingly *et al.* (2002)), the Langevin dynamics admits a unique invariant distribution

$$\mathbf{p}(x_1, x_2) = \Gamma_{\mathbf{p}} \exp(-V(x_1)) \exp(\tfrac{1}{2}|x_2|^2), \tag{4.40}$$



(a) (b)

FIG. 2. Invariant density approximation: (a) $h = 2^{-7}$; (b) $h = 2^{-4}$.

where $\Gamma_{\mathbf{p}}$ is a normalization constant. Therefore, this equation is always used to sample from a target probability distribution $\pi(x_1) \propto e^{-V(x_1)}$. Next we test the ability of the SITEM method to sample from the distribution. We take a large time endpoint $T = 500$ and test the SITEM in this paper, the implicit splitting AVF method in Cui *et al.* (2022) and the implicit Euler method in Talay (2002). Numerical results are depicted in Fig. 2, using two different stepsizes with $h = 2^{-7}, 2^{-4}$. There one can observe that, these three methods all perform very well in the case of small stepsize $h = 2^{-7}$. As the stepsize increases to $h = 2^{-4}$, the implicit Euler method produces better approximations than the other two methods. The SITEM method and the implicit splitting AVF method perform similarly and give acceptable approximations. It should be noted that both the splitting AVF method and the implicit Euler method are implicit time-stepping schemes and their computational costs are more expensive than the SITEM method in the high-dimensional setting $m > 1$.

We next turn to the stochastic van der Pol oscillator (4.36) with coefficients

$$m = 1, \gamma = \alpha = 0.2, \beta = 1, X_0 = (0.5, 1.5)^*, \delta = 3 \tag{4.41}$$

in the case of the additive noise $\phi_1(x) \equiv \vartheta = \sqrt{0.1}$ and the multiplicative noise $\phi_2(x) = 0.8x$. The mean-square approximation errors are presented in Fig. 3, where one can observe order-one convergence rate for both additive and multiplicative cases.

Now let us focus on the case of additive noise $\phi(x) \equiv \vartheta = \sqrt{0.1}$. According to (To, 2000, page 137), one knows that, in our setting, the stochastic van der Pol oscillator model (4.36) admits a stationary joint probability density

$$\mathbf{p}(x_1, x_2) = \Gamma_{\mathbf{p}} \exp\Big(-\tfrac{\alpha}{8\vartheta^2}\big((x_1^2 + x_2^2)^2 - 8(x_1^2 + x_2^2)\big)\Big), \tag{4.42}$$

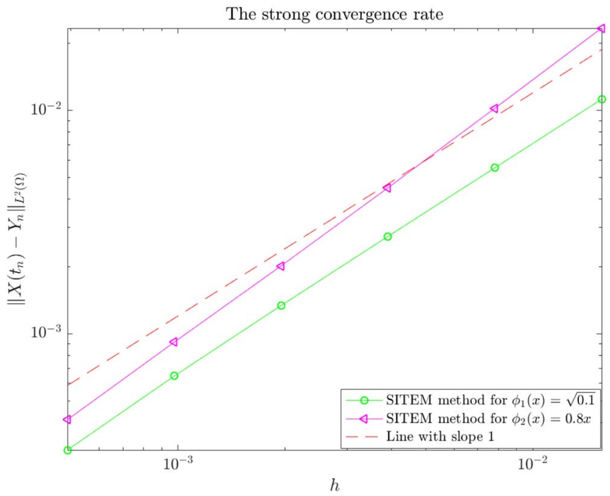


FIG. 3. Strong convergence rate for stochastic van der Pol oscillator.

where $\Gamma_{\mathbf{p}}$ is a normalization constant. Let $\mathbf{p}_1$ and $\mathbf{p}_2$ be marginal distribution of $\mathbf{p}$ in (4.42), defined by

$$\mathbf{p}_1(x_1) := \int_{\mathbb{R}} \mathbf{p}(x_1, x_2)\ \mathrm{d}x_2, \quad \mathbf{p}_2(x_2) := \int_{\mathbb{R}} \mathbf{p}(x_1, x_2)\ \mathrm{d}x_1.$$

It is obvious to observe that $\mathbf{p}(x_1, x_2)$ is symmetric with respect to variables x_1, x_2. As a consequence, $\mathbf{p}_1 \equiv \mathbf{p}_2$ and for polynomial functions $\Phi : \mathbb{R} \to \mathbb{R}$,

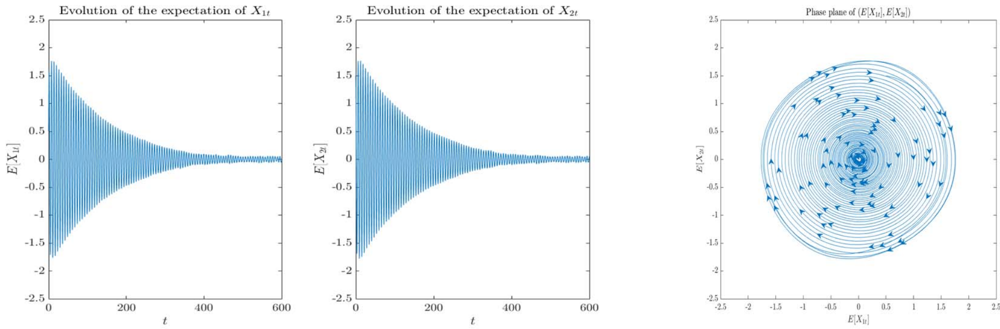
$$\mathbb{E}[\Phi(\xi_1)] = \mathbb{E}[\Phi(\xi_2)], \quad \xi_1 \sim \mathbf{p}_1(x_1),\ \xi_2 \sim \mathbf{p}_2(x_2).$$

Moreover, for a fixed variable, $\mathbf{p}$ is an even function with respect to the other. Therefore, by particularly taking $\Phi = I$ (identity map) we obtain

$$\mathbb{E}[\xi_1] = \mathbb{E}[\xi_2] = 0, \quad \xi_1 \sim \mathbf{p}_1(x_1),\ \xi_2 \sim \mathbf{p}_2(x_2).$$

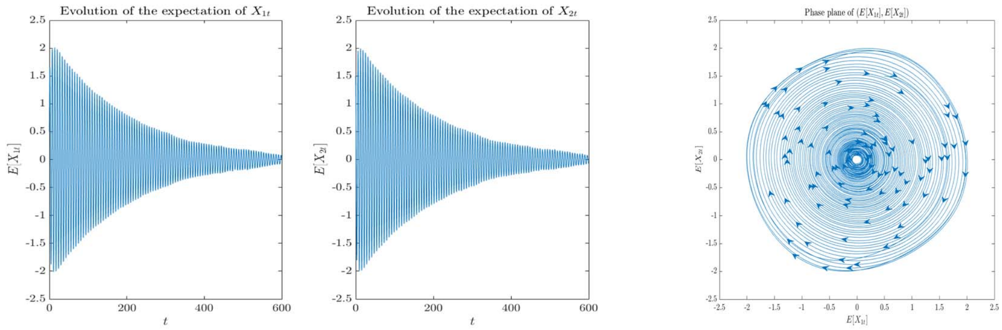
This implies that, in the phase plane, $(\mathbb{E}[X_{1t}], \mathbb{E}[X_{2t}])$ will gradually tend to the trivial steady state $(0, 0)$, as $t \to \infty$. Unlike the deterministic case, the average oscillation period and limit cycle do not exist for the stochastic van der Pol oscillator, due to the presence of the noise.

In what follows we test the dynamics of numerical approximations produced by the SITEM method. Over the time interval $[0, 600]$, Figures 4, 5 show the sample average trajectory and phase plane of the SITEM method for the stochastic van der Pol oscillator using two stepsizes $h = 2^{-7}, 2^{-4}$. From these figures, one can clearly see that, numerical approximations of $(\mathbb{E}[X_{1t}], \mathbb{E}[X_{2t}])$ produced by the SITEM method, even for a relatively large stepsize $h = 2^{-4}$, tend to the trivial steady state $(0, 0)$, as $t \to \infty$, reproducing the dynamics of the original model. Moreover, Figures 6, 7 demonstrate numerical approximations of $\mathbb{E}[|X_{1t}|^2]$ and $\mathbb{E}[|X_{2t}|^2]$ with two stepsizes $h = 2^{-7}, 2^{-4}$, where one can observe that they all tend to some non-zero steady states, as $t \to \infty$. The above numerical experiments indicate that,



(a) (b)

FIG. 4. Sample average trajectory and phase plane with $h = 2^{-7}$.



(a) (b)

FIG. 5. Sample average trajectory and phase plane with $h = 2^{-4}$.

using stepsizes of moderate size, the SITEM method is able to reproduce the dynamic property of the stochastic van der Pol oscillator.

It is very interesting to mention that, for the deterministic van der Pol oscillator, i.e., $\phi(x) = 0$, the exact solution is periodic and has a limit cycle. We use the SITEM method to numerically discretize it with various stepsizes. Our numerical results indicate that, both the period and limit cycle of the deterministic model are well reproduced by the SITEM method, even using a large stepsize $h = 2^{-2}$.

5. A positive preserving Milstein type scheme for the stochastic LV competition model with order-one pathwise uniformly strong convergence

In this section, we look at strong approximations of the following d-dimensional stochastic Lotka–Volterra (LV) competition model for interacting multi-species in ecology Bahar & Mao (2004);

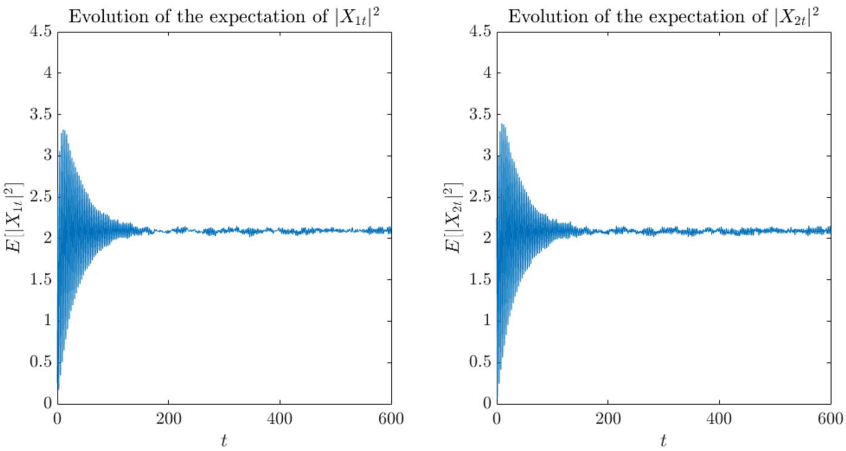


FIG. 6. Numerical approximations of $\mathbb{E}[|X_{1t}|^2]$ and $\mathbb{E}[|X_{2t}|^2]$ with $h = 2^{-7}$.

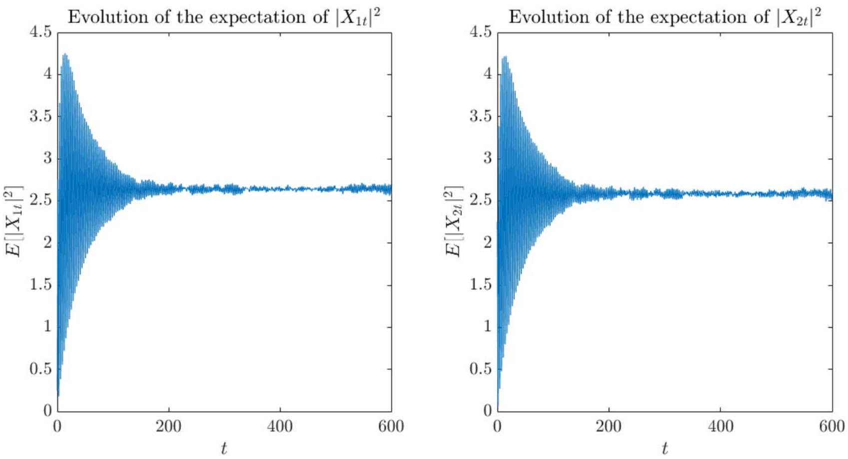


FIG. 7. Numerical approximations of $\mathbb{E}[|X_{1t}|^2]$ and $\mathbb{E}[|X_{2t}|^2]$ with $h = 2^{-4}$.

Mao (2007); Li & Cao (2023):

$$\begin{cases} \mathrm{d}X_t = \mathrm{diag}(X_t)(b - AX_t)\,\mathrm{d}t + \mathrm{diag}(X_t)\sigma\,\mathrm{d}W_t,\ t \in [0,T], \\ X_0 = x_0 \in (\mathbb{R}^d)^+, \end{cases} \tag{5.1}$$

where $T \in (0,\infty)$, $b := (b^{(i)})_{i=1,\dots,d} \in \mathbb{R}^d$, $A := (a^{(ij)})_{i,j=1,\dots,d} \in \mathbb{R}^{d\times d}$, $\sigma := (\sigma^{(ij)})_{i=1,\dots,d,j=1,\dots,m} \in \mathbb{R}^{d\times m}$ and $(\mathbb{R}^d)^+ := \{x \in \mathbb{R}^d : x^{(1)} > 0, \dots, x^{(d)} > 0\}$. Here the model is driven by multi-dimensional noise and $\{W_t\}_{t\in[0,T]}$ stands for a m-dimensional standard Brownian motion defined on

$(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$. For any vector $x \in \mathbb{R}^d$, we use diag(x) to denote a $d \times d$ diagonal matrix whose principal diagonal is x. In order to show the well-posedness of the underlying model in $(\mathbb{R}^d)^+$, some assumptions are put on the elements of the matrix A.

ASSUMPTION 5.1. Every element of A is non-negative and $\min_{1\leq i\leq d}\{a^{(ii)}\} > 0$.

Under the above assumption, the model (5.1) has a unique global strong solution in $(\mathbb{R}^d)^+$.

LEMMA 5.2. Under Assumption 5.1, there exists a unique strong solution $\{X_t\}_{t\geq 0}$ for the equation (5.1) staying in $(\mathbb{R}^d)^+$.

Proof. The well-posedness of the model in $(\mathbb{R}^d)^+$ has been established in Bahar & Mao (2004); Mao (2007) for the scalar noise case. For the present model driven by multi-dimensional noise, one can similarly prove it without any difficulty. □

Despite the existence and uniqueness of the positive solution to the stochastic LV model, the closed-form solution is not explicitly known and efficient numerical approximations become an important tool in applications. Recently, several researchers proposed and analyzed positivity-preserving numerical schemes for such a typical multi-dimensional SDE model with highly nonlinear and positive solution (see, e.g., Li *et al.* (2019); Mao *et al.* (2021); Hong *et al.* (2022); Cai *et al.* (2023)). It is worthwhile to point out that, under certain assumptions specified later, the highly nonlinear drift coefficient $f(x) := \text{diag}(x)(b - Ax)$ and the linear diffusion coefficient $g(x) := \text{diag}(x)\sigma$ obey the Razumikhin-type growth condition

$$\langle x, f(x)\rangle + c\|g(x)\|^2 \leq K(1 + |x|^2),$$

but violate the global monotonicity condition

$$\langle x - y, f(x) - f(y)\rangle + c\|g(x) - g(y)\|^2 \leq K|x - y|^2, \tag{5.2}$$

where $x, y \in (\mathbb{R}^d)^+$. As already mentioned in the introduction part, the lack of the global monotonicity condition causes an essential difficulty in obtaining convergence rates of numerical approximations. Usually, one needs to resort to exponential integrability of both the analytical and numerical solutions. Very recently, the authors of Li *et al.* (2019) constructed a Lamperti transformed EM method for (5.1) and used exponential integrability of both the analytical and numerical solutions to obtain pointwise strong convergence rate of order $\frac{1}{2}$ under some restrictive conditions.

In this work, we aim to propose a novel positivity preserving explicit Milstein-type method for the stochastic LV competition model and recover exactly order-one pathwise uniformly strong convergence of the new method under much relaxed conditions on the coefficients and stepsize, by relying on the use of previous perturbation estimates in Section 3 (see Theorem 5.6). To introduce the novel scheme, we regard the system (5.1) as an interacting particle system of d particles evolving on the line. For any $i \in \{1, ..., d\}$, we consider a single particle of (5.1) as follows:

$$\mathrm{d}X_t^{(i)} = \left(X_t^{(i)}b^{(i)} - X_t^{(i)}\sum_{j=1}^{d} a^{(ij)}X_t^{(j)}\right)\mathrm{d}t + X_t^{(i)}\sum_{j=1}^{m}\sigma^{(ij)}\,\mathrm{d}W_t^{(j)}. \tag{5.3}$$

In order to numerically approximate (5.3) on a uniform grid $\{t_n = nh\}_{0\le n\le N}$ with stepsize $h = \frac{T}{N}$, we propose the following linear-implicit (explicit) Milstein method starting from $Y_0 = X_0$:

$$Y^{(i)}_{t_{n+1}} = Y^{(i)}_{t_n} + Y^{(i)}_{t_{n+1}}\left(b^{(i)} - \sum_{j=1}^{d} a^{(ij)}Y^{(j)}_{t_n}\right)h + Y^{(i)}_{t_n}\sum_{j=1}^{m}\sigma^{(ij)}\Delta W^{(j)}_{t_n}$$
$$+\tfrac{1}{2}Y^{(i)}_{t_n}\sum_{j_1=1}^{m}\sum_{j_2=1}^{m}\sigma^{(ij_1)}\sigma^{(ij_2)}\Delta W^{(j_1)}_{t_n}\Delta W^{(j_2)}_{t_n} - \tfrac{1}{2}Y^{(i)}_{t_{n+1}}\sum_{j=1}^{m}(\sigma^{(ij)})^2 h, \tag{5.4}$$

where for $n = 0,...,N-1, j = 1,...,m, \Delta W^{(j)}_{t_n} := \Delta W^{(j)}_{t_{n+1}} - \Delta W^{(j)}_{t_n}$. Further, some elementary rearrangements turn (5.4) into

$$\left(1 - b^{(i)}h + h\sum_{j=1}^{d} a^{(ij)}Y^{(j)}_{t_n} + \tfrac{1}{2}\sum_{j=1}^{m}(\sigma^{(ij)})^2 h\right)Y^{(i)}_{t_{n+1}}$$
$$= Y^{(i)}_{t_n}\left(1 + \sum_{j=1}^{m}\sigma^{(ij)}\Delta W^{(j)}_{t_n} + \tfrac{1}{2}\sum_{j_1=1}^{m}\sum_{j_2=1}^{m}\sigma^{(ij_1)}\sigma^{(ij_2)}\Delta W^{(j_1)}_{t_n}\Delta W^{(j_2)}_{t_n}\right). \tag{5.5}$$

Under mild assumptions on the stepsize $h > 0$, one can readily check that the proposed scheme is well-posed and positivity preserving.

Proposition 5.3 (Positivity preserving). Let Assumption 5.1 be satisfied and let $X_0 = x_0 \in (\mathbb{R}^d)^+$. For some $\gamma > 1$ and $0 < h \le \min_{1\le i\le d}\left\{\frac{1}{\gamma(b_i - \frac{1}{2}\sum_{j=1}^{m}(\sigma^{(ij)})^2)\vee 0} \wedge T\right\}$ $(\frac{1}{0} := \infty)$, the proposed scheme (5.4) is well-defined and has a unique positive solution in $(\mathbb{R}^d)^+$.

Proof. Given $Y_{t_n} \in (\mathbb{R}^d)^+$, by Assumption 5.1 and the range of h , it is clear to see

$$0 < \left(1 - b^{(i)}h + h\sum_{j=1}^{d} a^{(ij)}Y^{(j)}_{t_n} + \tfrac{1}{2}\sum_{j=1}^{m}(\sigma^{(ij)})^2 h\right)^{-1} \le \tfrac{\gamma}{\gamma-1},$$

which implies equation (5.4) is well-defined and for any $1 \le i \le d$,

$$Y^{(i)}_{t_{n+1}} = Y^{(i)}_{t_n}\left(\tfrac{1}{2}\left(1 + \sum_{j=1}^{m}\sigma^{(ij)}\Delta W^{(j)}_{t_n}\right)^2 + \tfrac{1}{2}\right)\left(1 - b^{(i)}h + h\sum_{j=1}^{d} a^{(ij)}Y^{(j)}_{t_n} + \tfrac{1}{2}\sum_{j=1}^{m}(\sigma^{(ij)})^2 h\right)^{-1} > 0. \tag{5.6}$$

Since $Y_0 \in (\mathbb{R}^d)^+$, one infers that $Y_{t_n} \in (\mathbb{R}^d)^+$ for any $n = 0,...,N-1$, which ensures that the linear-implicit Milstein scheme is positivity-preserving and the proof is thus completed. □

Before coming to the convergence analysis of the scheme, we would like to mention an interesting observation for the particular stochastic LV model. Although the global monotonicity condition (5.2) does not hold, the drift coefficients of the model obey a special form of locally monotonicity condition (5.7), where the control terms on the right-hand side (cf. $U_0(x), U_1(x)$) depend only on x and does not depend on y. This interesting finding helps us to derive the convergence rate more easily, without requiring the exponential integrability properties of the numerical approximations. As a direct consequence of Theorem 3.2, we formulate the following proposition for approximations of SDEs fulfilling the special case of locally monotonicity condition (5.7).

Proposition 5.4. Let $f : \mathbb{R}^d \to \mathbb{R}^d, g : \mathbb{R}^d \to \mathbb{R}^{d\times m}$ be measurable functions and $f \in C^2(\mathbb{R}^d, \mathbb{R}^d)$. Further, let $f \in \mathcal{C}^1_{\mathcal{P}}(\mathbb{R}^d, \mathbb{R}^d)$ with constants K_f, c_f and g be Lipschitz. For a set $D_X \subset \mathbb{R}^d$, we assume that $X\colon \Omega \times [0,T] \to D_X$ and $Y\colon \Omega \times [0,T] \to D_X$ be defined by (3.1) and (3.2) with continuous sample paths, respectively, satisfying $\xi_Y = \xi_X = X_0$. Let $U_0 \in C^2\big(\mathbb{R}^d, [0,\infty)\big), U_1 \in C\big(\mathbb{R}^d, [0,\infty)\big)$, and let $c, v, q, T \in (0,\infty), \alpha \in [0,\infty), p \geq 4$. Besides, suppose that for all $x, y \in \mathbb{R}^d$,

(1) there exist constants $L, \kappa \geq 0$ such that for any $i = 1, ..., d$,

$$|U_0(x)| \bigvee \left\| \mathrm{Hess}_x(f^{(i)}(x)) \right\| \leq L(1+|x|)^{\kappa};$$

(2) $|x|^{1/c} \leq c(1+U_0(x))$ and $\mathbb{E}\left[e^{U_0(X_0)}\right] < \infty$;

(3) $(\mathcal{A}_{f,g}U_0)(x) + \frac{1}{2}|g(x)^*(\nabla U_0(x))|^2 + U_1(x) \leq c + \alpha U_0(x)$;

(4) for any $\eta > 0$, there exists a constant K_η such that

$$\langle x-y, f(x)-f(y)\rangle \leq \big[K_\eta + \eta(U_0(x)+U_1(x))\big]|x-y|^2,\ x, y \in D_X; \tag{5.7}$$

(5) for any $\theta \geq 1$,

$$\sup_{s\in[0,T]} \|a(s)\|_{L^\theta(\Omega;\mathbb{R}^d)} \bigvee \sup_{s\in[0,T]} \|b(s)\|_{L^\theta(\Omega;\mathbb{R}^{d\times m})} < K_\theta, \tag{5.8}$$

where $K_\theta > 0$ is independent of h.

Then for $\frac{1}{v} = \frac{1}{p} + \frac{1}{q}$, the approximation (3.2) of (3.1) admits

$$\begin{aligned}\left\| \sup_{t\in[0,T]} |X_t - Y_t| \right\|_{L^v(\Omega;\mathbb{R})} &\leq C\Bigg[h^2 + \int_0^T \left\| \mathbb{1}_{s\leq\tau_N} \|g(Y_s)-b(s)\| \right\|^2_{L^p(\Omega;\mathbb{R})} \mathrm{d}s \\ &+ h^{\frac{1}{2}} \int_0^T \int_{\lfloor s\rfloor_N}^s \left\| \mathbb{1}_{r\leq\tau_N} |f(Y_{\lfloor r\rfloor_N}) - a(r)| \right\|_{L^p(\Omega;\mathbb{R})} \mathrm{d}r\,\mathrm{d}s + \int_0^T \left\| \mathbb{1}_{s\leq\tau_N} \big|f(Y_{\lfloor s\rfloor_N}) - a(s)\big| \right\|^2_{L^p(\Omega;\mathbb{R})} \mathrm{d}s \\ &+ h^{\frac{1}{2}} \int_0^T \left(\int_{\lfloor s\rfloor_N}^s \left\| \mathbb{1}_{r\leq\tau_N} \|g(Y_r)-b(r)\| \right\|^2_{L^p(\Omega;\mathbb{R})} \mathrm{d}r \right)^{\frac{1}{2}} \mathrm{d}s \Bigg]^{\frac{1}{2}},\end{aligned} \tag{5.9}$$

where C is independent of h.

Proof. Proposition 5.4 can be directly obtained by using Theorem 3.2. To see this, setting $\tau_N = T$ in Theorem 3.2 and utilizing the same arguments as used in (4.16) one deduces that

$$\left\| \exp\left(\int_0^T \left[\frac{\langle X_s - Y_s, f(X_s) - f(Y_s)\rangle + \frac{1+\varepsilon}{2}\|g(X_s)-g(Y_s)\|^2}{|X_s - Y_s|^2} \right]^+ \mathrm{d}s \right) \right\|_{L^q(\Omega;\mathbb{R})} < \infty. \tag{5.10}$$

In view of (5.10) and conditions (1), (2), (4) in Proposition 5.4, we arrive at

$$\sup_{s\in[0,T]} \left\| \mathbb{1}_{s\le\tau_N} \left[\frac{\langle X_s - Y_s, f(X_s) - f(Y_s)\rangle + \frac{1+\varepsilon}{2}\|g(X_s)-g(Y_s)\|^2}{|X_s - Y_s|^2} \right]^+ \right\|_{L^{3p}(\Omega;\mathbb{R})} < \infty \tag{5.11}$$

and

$$\sup_{s\in[0,T]} \|X_s\|_{L^{6pc_g \vee 3pc_f \vee 3p}(\Omega;\mathbb{R}^d)} < \infty. \tag{5.12}$$

Combining Theorem 3.2 with conditions (1), (5) in Proposition 5.4 then completes the proof. □

It is worth noting that the restrictions on U_0 and U_1 in Proposition 5.4 are more relaxed than those in Theorem 4.2 due to the particular condition (4) (see (Cox *et al.*, 2024, Corollary 2.4)). In what follows, we utilize Proposition 5.4 to prove the strong convergence rate of the newly developed Milstein type method. For simplicity of presentation, we denote

$$\underline{a} := \min_{1\le i\le d}\{a^{(ii)}\},\ \bar{b} := \max_{1\le i\le d}\{|b^{(i)}|\};\ \bar{\sigma} := \max_{1\le i\le d, 1\le j\le m}\{|\sigma^{(ij)}|\}$$

and

$$Q_{t_n}^{(i)} := 1 - b^{(i)}h + h\sum_{j=1}^{d} a^{(ij)} Y_{t_n}^{(j)} + \tfrac{1}{2}\sum_{j=1}^{m} (\sigma^{(ij)})^2 h,\ i \in \{1, ..., d\}. \tag{5.13}$$

Under conditions in Proposition 5.3, one knows

$$0 < (Q_{t_n}^{(i)})^{-1} \le \min\left\{ \tfrac{\gamma}{\gamma-1}, 1 + \tfrac{\gamma}{\gamma-1}\left| b_i - \tfrac{1}{2}\sum_{j=1}^{m} (\sigma^{(ij)})^2 \right| h \right\}. \tag{5.14}$$

In the notation of $Q^{(i)}_{t_n}$, one can rewrite the scheme (5.5) as

$$
\begin{aligned}
Y^{(i)}_{t_{n+1}} &= Y^{(i)}_{t_n}\left(1+\sum_{j=1}^{m}\sigma^{(ij)}\Delta W^{(j)}_{t_n}+\frac{1}{2}\sum_{j_1=1}^{m}\sum_{j_2=1}^{m}\sigma^{(ij_1)}\sigma^{(ij_2)}\Delta W^{(j_1)}_{t_n}\Delta W^{(j_2)}_{t_n}\right)(Q^{(i)}_{t_n})^{-1}\\
&= Y^{(i)}_{t_n}+Y^{(i)}_{t_n}\big((Q^{(i)}_{t_n})^{-1}-1\big)\\
&\quad + Y^{(i)}_{t_n}\left(\sum_{j=1}^{m}\sigma^{(ij)}\Delta W^{(j)}_{t_n}+\frac{1}{2}\sum_{j_1=1}^{m}\sum_{j_2=1}^{m}\sigma^{(ij_1)}\sigma^{(ij_2)}\Delta W^{(j_1)}_{t_n}\Delta W^{(j_2)}_{t_n}\right)(Q^{(i)}_{t_n})^{-1}\\
&= Y^{(i)}_{t_n}+Y^{(i)}_{t_n}(Q^{(i)}_{t_n})^{-1}\left(b^{(i)}h-h\sum_{j=1}^{d}a^{(ij)}Y^{(j)}_{t_n}\right)\\
&\quad + Y^{(i)}_{t_n}\left(\sum_{j=1}^{m}\sigma^{(ij)}\Delta W^{(j)}_{t_n}+\frac{1}{2}\sum_{j_1=1}^{m}\sum_{j_2=1}^{m}\sigma^{(ij_1)}\sigma^{(ij_2)}\Delta W^{(j_1)}_{t_n}\Delta W^{(j_2)}_{t_n}-\frac{1}{2}\sum_{j=1}^{m}(\sigma^{(ij)})^2h\right)(Q^{(i)}_{t_n})^{-1}\\
&= Y^{(i)}_{t_n}+\int_{t_n}^{t_{n+1}}\frac{b^{(i)}-\sum_{j=1}^{d}a^{(ij)}Y^{(j)}_{t_n}}{Q^{(i)}_{t_n}}Y^{(i)}_{t_n}\,\mathrm{d}s+\sum_{j_1=1}^{m}\int_{t_n}^{t_{n+1}}\left(\frac{\sigma^{(ij_1)}+\sigma^{(ij_1)}\sum_{j_2=1}^{m}\sigma^{(ij_2)}\big(W^{(j_2)}_s-W^{(j_2)}_{\lfloor s\rfloor_N}\big)}{Q^{(i)}_{t_n}}Y^{(i)}_{t_n}\right)\mathrm{d}W^{(j_1)}_s.
\end{aligned}
\tag{5.15}
$$

For any $t\in[0,T]$, one can thus define a continuous version of (5.15) as follows:

$$
Y^{(i)}_t = Y^{(i)}_0+\int_0^t\frac{b^{(i)}-\sum_{j=1}^{d}a^{(ij)}Y^{(j)}_{\lfloor s\rfloor_N}}{Q^{(i)}_{\lfloor s\rfloor_N}}Y^{(i)}_{\lfloor s\rfloor_N}\,\mathrm{d}s+\sum_{j_1=1}^{m}\int_0^t\left(\frac{\sigma^{(ij_1)}+\sigma^{(ij_1)}\sum_{j_2=1}^{m}\sigma^{(ij_2)}\big(W^{(j_2)}_s-W^{(j_2)}_{\lfloor s\rfloor_N}\big)}{Q^{(i)}_{\lfloor s\rfloor_N}}Y^{(i)}_{\lfloor s\rfloor_N}\right)\mathrm{d}W^{(j_1)}_s.
\tag{5.16}
$$

To prove the strong convergence rate of the scheme, the following lemma is also indispensable.

LEMMA 5.5 (Bounded moments). Let all conditions in Proposition 5.3 hold. Then for any $p\geq 1$, there exists a positive constant C_p independent of h, such that the numerical approximations produced by (5.15) obey

$$
\sup_{1\leq n\leq N}\mathbb{E}[|Y_{t_n}|^p]\leq C_p.
\tag{5.17}
$$

Proof. For any $p \in \mathbb{Z}^+$, by (5.15) and binomial expansion we have that for $i = 1, ..., d,\ j, j_1, j_2 = 1, ..., m$,

$$
\begin{aligned}
&\mathbb{E}[|Y_{t_{n+1}}^{(i)}|^p] \\
&= \mathbb{E}\big[|Y_{t_n}^{(i)}|^p|(Q_{t_n}^{(i)})^{-1}|^p\big] \times \mathbb{E}\left[\left(1 + \sum_{j=1}^m \sigma^{(ij)} \Delta W_{t_n}^{(j)} + \frac{1}{2}\sum_{j_1=1}^m \sum_{j_2=1}^m \sigma^{(ij_1)}\sigma^{(ij_2)} \Delta W_{t_n}^{(j_1)} \Delta W_{t_n}^{(j_2)}\right)^p\right] \\
&\leq (1 + Ch)^p \mathbb{E}[|Y_{t_n}^{(i)}|^p] \sum_{k=0}^p \binom{p}{k} \mathbb{E}\left[\left(\sum_{j=1}^m \sigma^{(ij)} \Delta W_{t_n}^{(j)} + \frac{1}{2}\sum_{j_1=1}^m \sum_{j_2=1}^m \sigma^{(ij_1)}\sigma^{(ij_2)} \Delta W_{t_n}^{(j_1)} \Delta W_{t_n}^{(j_2)}\right)^k\right] \\
&\leq (1 + Ch)^p \mathbb{E}[|Y_{t_n}^{(i)}|^p] \\
&\quad \times \left(1 + C_p \sum_{k=1}^p \mathbb{E}\left[\left(\sum_{j=1}^m \sigma^{(ij)} \Delta W_{t_n}^{(j)} + \frac{1}{2}\sum_{j_1=1}^m \sum_{j_2=1}^m \sigma^{(ij_1)}\sigma^{(ij_2)} \Delta W_{t_n}^{(j_1)} \Delta W_{t_n}^{(j_2)}\right)^k\right]\right),
\end{aligned}
\tag{5.18}
$$

where $\left(pk := \frac{p!}{k!(p-k)!}\right)$ are the coefficients of binomial expansion. By iteration one concludes that for any $n = 0, 1, ..., N-1$,

$$
\mathbb{E}[|Y_{t_{n+1}}^{(i)}|^p] \leq (1 + Ch)^{p+1} \mathbb{E}[|Y_{t_n}^{(i)}|^p] \leq (1 + Ch)^{(p+1)(n+1)} \mathbb{E}[|Y_0^{(i)}|^p] \leq e^{2(p+1)CT} |x_0^{(i)}|^p, \tag{5.19}
$$

which finishes the proof for positive integer p. Thanks to the Hölder inequality, the inequality (5.17) also holds true for any non-integer $p \geq 1$. □

Now we are well-prepared to show the order-one pathwise uniformly strong convergence of the proposed scheme.

Theorem 5.6 (Order-one pathwise uniformly strong convergence). Let Assumption 5.1 be satisfied and assume $0 < h \leq \min_{1 \leq i \leq d} \left\{ \frac{1}{\gamma(b_i - \frac{1}{2}\sum_{j=1}^m (\sigma^{(ij)})^2) \vee 0} \wedge T \right\}$ with $\gamma > 1$. Let X_s and Y_s be the exact solution and numerical solution defined by (5.3) and (5.16), respectively. Then for any $r > 0$,

$$
\mathbb{E}\left[\sup_{t \in [0,T]} |X_t - Y_t|^r\right] \leq Ch^r. \tag{5.20}
$$

Proof. Define

$$
U_0(x) := v(1 + |x|^2)^{1/2}, \quad U_1(x) := v(1 + |x|^2)^{-\frac{1}{2}}\left(\underline{a}d^{-\frac{1}{2}} - \frac{vm\bar{\sigma}^2}{2}\right)|x|^3 + v(1 + |x|^2)^{1/2}.
$$

Here $v > 0$ is chosen to be some small constant such that $\left(\underline{a}d^{-\frac{1}{2}} - \frac{vm\bar{\sigma}^2}{2}\right) > 0$, which in turn ensures $U_1(x)$ is positive and there exists an $\epsilon > 0$ satisfying $\epsilon|x|^2 \leq U_1(x)$. It is easy to check that conditions

(1) and (2) in Proposition 5.4 are fulfilled. To validate condition (3), we derive

$$
\begin{aligned}
&U_0'(x)f(x) + \tfrac{1}{2}\operatorname{trace}\Big(g(x)g(x)^*\operatorname{Hess}_x(U_0(x))\Big) + \tfrac{1}{2}\left|g(x)^*(\nabla U_0)(x)\right|^2 + U_1(x) - \alpha U_0(x)\\
&= \sum_{i=1}^{d} v(1+|x|^2)^{-\frac{1}{2}}(x^{(i)})^2(b^{(i)} - \sum_{j=1}^{d} a^{(ij)}x^{(j)}) + \tfrac{1}{2}\sum_{i=1}^{d}\sum_{k=1}^{m} v(|1+|x|^2)^{-\frac{1}{2}}(x^{(i)})^2(\sigma^{(ik)})^2\\
&\quad - \tfrac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d}\sum_{k=1}^{m} v(|1+|x|^2)^{-\frac{3}{2}}(x^{(i)})^2(x^{(j)})^2\sigma^{(ik)}\sigma^{(jk)} + U_1(x)\\
&\quad + \tfrac{1}{2}\sum_{k=1}^{m}\Big(\sum_{i=1}^{d} v(1+|x|^2)^{-\frac{1}{2}}(x^{(i)})^2\sigma^{(ik)}\Big)^2 - \alpha v(1+|x|^2)^{\frac{1}{2}}\\
&\le \sum_{i=1}^{d} v(1+|x|^2)^{-\frac{1}{2}}(x^{(i)})^2\bar{b} + \tfrac{1}{2}\sum_{i=1}^{d}\sum_{k=1}^{m} v(1+|x|^2)^{-\frac{1}{2}}(x^{(i)})^2\bar{\sigma}^2\\
&\quad + \tfrac{1}{2}\sum_{i=1}^{d}\sum_{j=1}^{d}\sum_{k=1}^{m} v(1+|x|^2)^{-\frac{3}{2}}(x^{(i)})^2(x^{(j)})^2\bar{\sigma}^2 - \alpha v(1+|x|^2)^{\frac{1}{2}}\\
&\quad + \tfrac{m}{2}v^2\bar{\sigma}^2|x|^4(1+|x|^2)^{-1} - v(1+|x|^2)^{-\frac{1}{2}}\underline{a}\sum_{i=1}^{d}(x^{(i)})^3 + U_1(x)\\
&\le v\bar{b}(1+|x|^2)^{\frac{1}{2}} + mv\bar{\sigma}^2(1+|x|^2)^{\frac{1}{2}} - \alpha v(1+|x|^2)^{\frac{1}{2}} + v(1+|x|^2)^{-\frac{1}{2}}\big(\tfrac{mv}{2}\bar{\sigma}^2 - \underline{a}d^{-\frac{1}{2}}\big)|x|^3 + U_1(x)\\
&= (\bar{b} + m\bar{\sigma}^2 - \alpha + 1)v(1+|x|^2)^{\frac{1}{2}} \le 0, \qquad (5.21)
\end{aligned}
$$

where $\alpha > 0$ is chosen to be large enough so that $\bar{b} + m\bar{\sigma}^2 - \alpha + 1 \le 0$ and the condition (3) in Proposition 5.4 is hence validated. Furthermore, for any $x, y \in (\mathbb{R}^d)^+$, it holds that

$$
\begin{aligned}
\langle x-y, f(x)-f(y)\rangle &= \langle x-y, \operatorname{diag}(x)b - \operatorname{diag}(x)Ax - \operatorname{diag}(y)b + \operatorname{diag}(y)Ay\rangle\\
&= \langle x-y, \operatorname{diag}(x-y)b\rangle - \langle x-y, \operatorname{diag}(x)A(x-y)\rangle - \langle x-y, \operatorname{diag}(x-y)Ay\rangle\\
&\le |b||x-y|^2 + |x-y|^2\|\operatorname{diag}(x)A\| - \sum_{i=1}^{d}(x^{(i)}-y^{(i)})^2\sum_{j=1}^{d}a^{(ij)}y^{(j)}\\
&\le (|b| + \|A\||x|)|x-y|^2. \qquad (5.22)
\end{aligned}
$$

This implies condition (4) in Proposition 5.4 with $K_\eta \ge \frac{1}{4\eta\epsilon}$. By observing that for $i = 1,...,d$, $j = 1,...,m$,

$$
a^{(i)}(s) = \frac{b^{(i)} - \sum_{j=1}^{d} a^{(ij)}Y^{(j)}_{\lfloor s\rfloor_N}}{Q^{(i)}_{\lfloor s\rfloor_N}}Y^{(i)}_{\lfloor s\rfloor_N}, \quad b^{(ij)}(s) = \frac{\sigma^{(ij)} + \sigma^{(ij)}\sum_{j_1=1}^{m}\sigma^{(ij_1)}\Big(W^{(j_1)}_s - W^{(j_1)}_{\lfloor s\rfloor_N}\Big)}{Q^{(i)}_{\lfloor s\rfloor_N}}Y^{(i)}_{\lfloor s\rfloor_N}, \qquad (5.23)
$$

the condition (5) in Proposition 5.4 is therefore satisfied due to Lemma 5.5 and (5.14). Now all conditions of Proposition 5.4 have been confirmed. As a consequence, we arrive at the assertion (5.9). Following the notation of (5.23) and recalling $f(x) := \mathrm{diag}(x)(b - Ax)$, $g(x) := \mathrm{diag}(x)\sigma$, we use Lemma 5.5 and the Hölder inequality to show that, for any $p \geq 4$,

$$\begin{aligned}
\left\| f^{(i)}(Y_{\lfloor s \rfloor_N}) - a^{(i)}(s) \right\|_{L^p(\Omega;\mathbb{R})} &= \left\| \left(1 - \tfrac{1}{Q^{(i)}_{\lfloor s \rfloor_N}}\right) \left(b^{(i)} - \sum_{j=1}^{d} a^{(ij)} Y^{(j)}_{\lfloor s \rfloor_N}\right) Y^{(i)}_{\lfloor s \rfloor_N} \right\|_{L^p(\Omega;\mathbb{R})} \\
&\leq \left\| \left(1 - \tfrac{1}{Q^{(i)}_{\lfloor s \rfloor_N}}\right) \right\|_{L^{2p}(\Omega;\mathbb{R})} \left\| \left(b^{(i)} - \sum_{j=1}^{d} a^{(ij)} Y^{(j)}_{\lfloor s \rfloor_N}\right) Y^{(i)}_{\lfloor s \rfloor_N} \right\|_{L^{2p}(\Omega;\mathbb{R})} \\
&\leq Ch.
\end{aligned} \tag{5.24}$$

Meanwhile, expanding $Y^{(i)}(s)$ at $s = \lfloor s \rfloor_N$ by Itô's formula gives

$$\begin{aligned}
&\left\| g^{(ij)}(Y_s) - b^{(ij)}(s) \right\|_{L^p(\Omega;\mathbb{R})} \\
&= \Bigg\| Y^{(i)}_{\lfloor s \rfloor_N} \sigma^{(ij)} + \frac{b^{(i)} - \sum_{k=1}^{d} a^{(ik)} Y^{(k)}_{\lfloor s \rfloor_N}}{Q^{(i)}_{\lfloor s \rfloor_N}} Y^{(i)}_{\lfloor s \rfloor_N} (s - \lfloor s \rfloor_N) \sigma^{(ij)} + \frac{Y^{(i)}_{\lfloor s \rfloor_N} \sigma^{(ij)}}{Q^{(i)}_{\lfloor s \rfloor_N}} \sum_{j_1=1}^{m} \Big(\sigma^{(ij_1)} (W^{(j_1)}_s - W^{(j_1)}_{\lfloor s \rfloor_N}) \\
&\quad + \tfrac{1}{2} \sum_{j_2=1}^{m} \sigma^{(ij_1)} \sigma^{(ij_2)} (W^{(j_1)}_s - W^{(j_1)}_{\lfloor s \rfloor_N})(W^{(j_2)}_s - W^{(j_2)}_{\lfloor s \rfloor_N}) - \tfrac{1}{2} (\sigma^{(ij_1)})^2 (s - \lfloor s \rfloor_N) \Big) \\
&\quad - \frac{Y^{(i)}_{\lfloor s \rfloor_N}}{Q^{(i)}_{\lfloor s \rfloor_N}} \sigma^{(ij)} \Big(1 + \sum_{j_1=1}^{m} \sigma^{(ij_1)} (W^{(j_1)}_s - W^{(j_1)}_{\lfloor s \rfloor_N}) \Big) \Bigg\|_{L^p(\Omega;\mathbb{R})} \\
&= \Bigg\| Y^{(i)}_{\lfloor s \rfloor_N} \sigma^{(ij)} \frac{Q^{(i)}_{\lfloor s \rfloor_N} - 1}{Q^{(i)}_{\lfloor s \rfloor_N}} + \frac{b^{(i)} - \sum_{k=1}^{d} a^{(ik)} Y^{(k)}_{\lfloor s \rfloor_N}}{Q^{(i)}_{\lfloor s \rfloor_N}} Y^{(i)}_{\lfloor s \rfloor_N} (s - \lfloor s \rfloor_N) \sigma^{(ij)} \\
&\quad + \frac{Y^{(i)}_{\lfloor s \rfloor_N} \sigma^{(ij)}}{2 Q^{(i)}_{\lfloor s \rfloor_N}} \sum_{j_1=1}^{m} \Big(\sum_{j_2=1}^{m} \sigma^{(ij_1)} \sigma^{(ij_2)} (W^{(j_1)}_s - W^{(j_1)}_{\lfloor s \rfloor_N})(W^{(j_2)}_s - W^{(j_2)}_{\lfloor s \rfloor_N}) - (\sigma^{(ij_1)})^2 (s - \lfloor s \rfloor_N) \Big) \Bigg\|_{L^p(\Omega;\mathbb{R})} \\
&\leq \left\| Y^{(i)}_{\lfloor s \rfloor_N} \sigma^{(ij)} \frac{Q^{(i)}_{\lfloor s \rfloor_N} - 1}{Q^{(i)}_{\lfloor s \rfloor_N}} \right\|_{L^p(\Omega;\mathbb{R})} + \left\| \frac{b^{(i)} - \sum_{k=1}^{d} a^{(ik)} Y^{(k)}_{\lfloor s \rfloor_N}}{Q^{(i)}_{\lfloor s \rfloor_N}} Y^{(i)}_{\lfloor s \rfloor_N} (s - \lfloor s \rfloor_N) \sigma^{(ij)} \right\|_{L^p(\Omega;\mathbb{R})} \\
&\quad + \left\| \frac{Y^{(i)}_{\lfloor s \rfloor_N} \sigma^{(ij)}}{2 Q^{(i)}_{\lfloor s \rfloor_N}} \sum_{j_1=1}^{m} \left(\sum_{j_2=1}^{m} \sigma^{(ij_1)} \sigma^{(ij_2)} \left(W^{(j_1)}_s - W^{(j_1)}_{\lfloor s \rfloor_N}\right) \left(W^{(j_2)}_s - W^{(j_2)}_{\lfloor s \rfloor_N}\right) - (\sigma^{(ij_1)})^2 \left(s - \lfloor s \rfloor_N\right) \right) \right\|_{L^p(\Omega;\mathbb{R})} \\
&\leq Ch.
\end{aligned} \tag{5.25}$$

Combining (5.24)–(5.25) with (5.9) finally yields the desired assertion (5.20). □

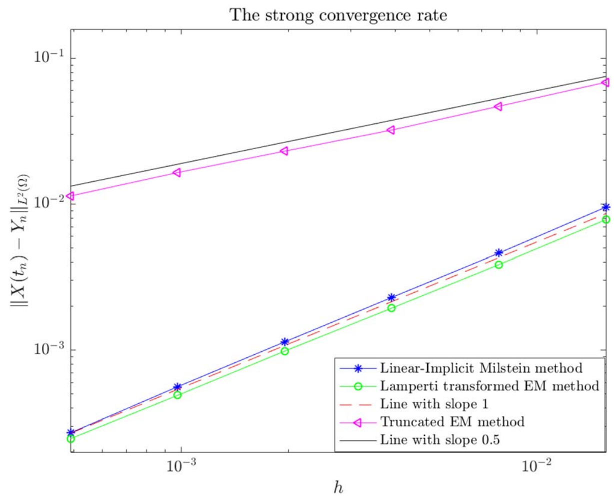


FIG. 8. A comparison of strong convergence rates for LV competition model.

Numerical experiments. Now we present some numerical experiments to confirm the theoretical results. Let $d = m = 2$ and consider the stochastic LV competition model with coefficients

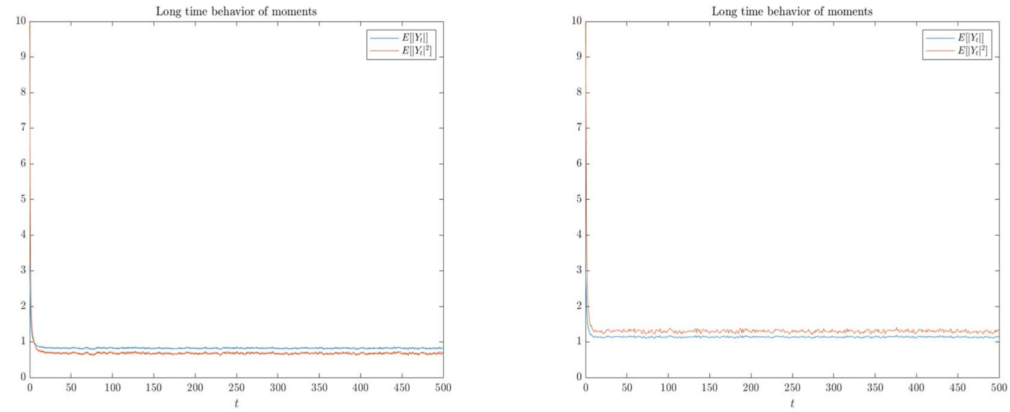
$$b = \begin{pmatrix} 1 \\ 0.5 \end{pmatrix}, A = \begin{pmatrix} 1 & 0.5 \\ 0 & 0.5 \end{pmatrix}, \sigma = \begin{pmatrix} 1 & 0 \\ 0 & 0.75 \end{pmatrix}, Y(0) = \begin{pmatrix} 1 \\ 3 \end{pmatrix}. \tag{5.26}$$

Let $T = 1$, $N = 2^k, k = 6, 7, ..., 11$ and regard the fine approximations with $h_{\text{exact}} = 2^{-14}$ as the 'true' solution. We consider the mean-square approximation errors and take $M = 5000$ Monte Carlo sample paths to approximate the expectation. A comparison of strong convergence rates for our method, the Lamperti transformed EM method in Li *et al.* (2019) and the truncated EM method in Mao *et al.* (2021) is presented in Fig. 8. One can clearly observe order-one convergence of our method and the Lamperti transformed EM method, and order $\frac{1}{2}$ convergence of the truncated EM method.

In addition to the strong convergence rate, we would also like to investigate the dynamic preservation of the proposed method. As shown by Bahar & Mao (2004); Mao (2007), under Assumption 5.1, the exact solution $\{X_t\}_{t\geq 0}$ of (5.1) admits an ultimate boundedness property, i.e., there exist two positive constants C_1, C_2 independent of X_0 such that

$$\limsup_{t\to\infty} \mathbb{E}[|X_t|] \leq C_1 \quad \text{and} \quad \limsup_{t\to\infty} \frac{1}{t}\int_0^t \mathbb{E}[|X_s|^2]\, \mathrm{d}s \leq C_2. \tag{5.27}$$

Fig. 9 displays moments over long-time interval $[0, T], T = 500$ of the linear-implicit Milstein method using small stepsize $h = 2^{-7}$ and large stepsize $h = 1$. It is observed that, the numerical approximations produced by the linear-implicit Milstein method remain bounded after a long time, even for a large stepsize $h = 1$. This recovers the property (5.27) of the exact solution, which can be theoretically explained as follows. For any $p \geq 1$ and $0 < h < \min_{1\leq i\leq d}\left\{\frac{1}{(b_i - \frac{1}{2}\sum_{j=1}^{m}(\sigma^{(ij)})^2)\vee 0}\right\}$, recall (5.13) and



(a) (b)

FIG. 9. Evolution of numerical moments: (a) $h = 2^{-7}$; (b) $h = 1$.

(5.15), it holds that

$$\mathbb{E}\big[|Y^{(i)}_{t_{n+1}}|^p\big] = \mathbb{E}\left[\left|\frac{Y^{(i)}_{t_n}}{Q^{(i)}_{t_n}}\right|^p\right]\mathbb{E}\left[\left(\frac{1}{2}\left(1+\sum_{j=1}^{m}\sigma^{(ij)}\Delta W^{(j)}_{t_n}\right)^2+\frac{1}{2}\right)^p\right] \le C_{p,h}\left(\frac{1}{ha^{(ii)}}\right)^p < +\infty,$$

for any $n \in \mathbb{N}$.

Another significant dynamic of the model (5.1) is permanence and extinction. Permanence means that the species will persist and extinction means that the species will eventually become extinct (see (Li & Cao, 2023, Definition 3.6, Section 4) for precise definitions). According to (Li & Cao, 2023, Theorem 3.7), in our setting (5.26) the species will be permanent. In Fig. 10, we draw $M = 500$ sample paths of the proposed linear implicit Milstein scheme over time interval $[0, 500]$ with $h = 2^{-3}$. Evidently, the numerical approximations reproduce the dynamic of the permanence of the original model. When the noise intensity increases from σ to 2σ, in view of (Li & Cao, 2023, Corollary 2), the species will eventually become extinct. Fig. 11 presents $M = 500$ sample paths of the proposed approximation method over time interval $[0, 500]$ with $h = 2^{-3}$. There one can see that the numerical approximations tend to zero, reproducing the dynamic of the extinct of the original model. For both cases, all paths stay in $(\mathbb{R}^2)^+$, confirming the positivity preserving of the proposed method.

6. Conclusion

In this paper, we successfully reveal order-one strong convergence of two kinds of numerical methods for several SDEs without globally monotone coefficients, which fills the gap left by Hutzenthaler & Jentzen (2020). This is accomplished by developing some new perturbation estimates and some more careful estimates. Numerical experiments are also provided to support the theoretical findings. As an ongoing

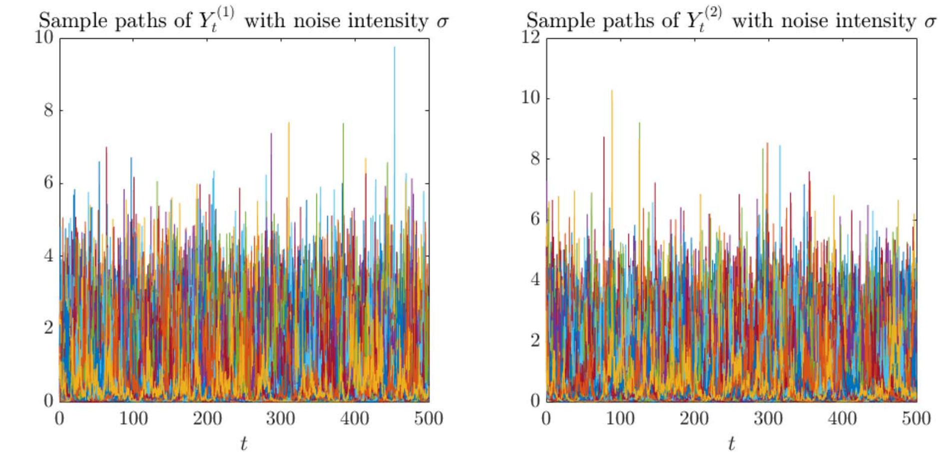


FIG. 10. Permanent performance of linear-implicit Milstein method with $h = 2^{-3}$.

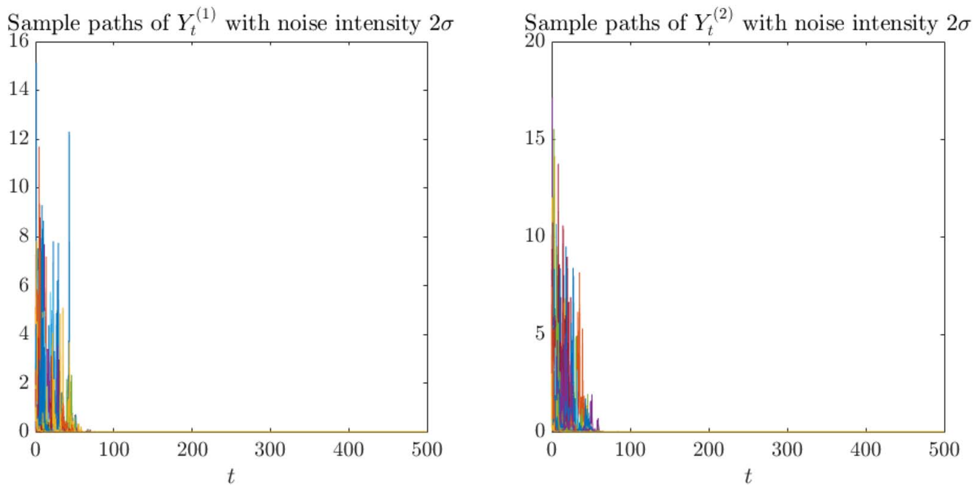


FIG. 11. Extinct performance of linear-implicit Milstein method with $h = 2^{-3}$.

project, we propose and analyze new higher order (strong order 1 and 1.5) time-stepping schemes for general SDEs without globally monotone coefficients.

Funding

Natural Science Foundation of China (nos: 12471394, 12071488, 12371417); Postgraduate Scientific Research Innovation Project of Hunan Province (no CX20230349) and Fundamental Research Funds for the Central Universities of Central South University (no: 2021zzts0480). The authors would like to thank Prof. Arnulf Jentzen for his helpful comments to improve the manuscript, when he was invited by Xiaojie Wang to visit Central South University in June of 2024.

References

Alfonsi, A. (2013) Strong order one convergence of a drift implicit Euler scheme: application to the CIR process. *Statist. Probab. Lett.*, **83**, 602–607.

Andersson, A. & Kruse, R. (2017) Mean-square convergence of the BDF2-Maruyama and backward Euler schemes for SDE satisfying a global monotonicity condition. *BIT Numer. Math.*, **57**, 21–53.

Bahar, A. & Mao, X. (2004) Stochastic delay population dynamics. *Int. J. Pure Appl. Math.*, **11**, 377–400.

Beyn, W.-J., Isaak, E. & Kruse, R. (2017) Stochastic C-stability and B-consistency of explicit and implicit Milstein-type schemes. *J. Sci. Comput.*, **70**, 1042–1077.

Brehier, C.-E. (2023) Approximation of the invariant distribution for a class of ergodic SDEs with one-sided Lipschitz continuous drift coefficient using an explicit tamed Euler scheme. *ESAIM: Prob. Stat.*, **27**, 841–866.

Cai, Y., Guo, Q. & Mao, X. (2023) Positivity preserving truncated scheme for the stochastic Lotka–Volterra model with small moment convergence. *Calcolo*, **60**, 24.

Cox, S., Hutzenthaler, M. & Jentzen, A. (2024) *Local Lipschitz Continuity in the Initial Value and Strong Completeness for Nonlinear Stochastic Differential Equations*, vol. **296**. American Mathematical Society. AQ7

Cui, J., Hong, J. & Sheng, D. (2022) Density function of numerical solution of splitting AVF scheme for stochastic Langevin equation. *Math. Comp.*, **91**, 2283–2333.

Fang, W. & Giles, M. B. (2020) Adaptive Euler–Maruyama method for SDEs with nonglobally Lipschitz drift. *Ann. Appl. Probab.*, **30**, 526–560.

Higham, D. J. (2000) Mean-square and asymptotic stability of the stochastic theta method. *SIAM J. Numer. Anal.*, **38**, 753–769.

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

Hong, J., Ji, L., Wang, X. & Zhang, J. (2022) Positivity-preserving symplectic methods for the stochastic Lotka–Volterra predator-prey model. *BIT*, **62**, 493–520.

Hutzenthaler, M. & Jentzen, A. (2011) Convergence of the stochastic Euler scheme for locally Lipschitz coefficients. *Found. Comput. Math.*, **11**, 657–706.

Hutzenthaler, M. & Jentzen, A. (2015) *Numerical Approximations of Stochastic Differential Equations with Non-globally Lipschitz Continuous Coefficients*, vol. **236**. American Mathematical Society. AQ8

Hutzenthaler, M. & Jentzen, A. (2020) On a perturbation theory and on strong convergence rates for stochastic ordinary and partial differential equations with nonglobally monotone coefficients. *Ann. Probab.*, **48**, 53–93.

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

Hutzenthaler, M., Jentzen, A. & Kloeden, P. E. (2012) Strong convergence of an explicit numerical method for SDEs with nonglobally Lipschitz continuous coefficients. *Ann. Appl. Probab.*, **22**, 1611–1641.

Hutzenthaler, M., Jentzen, A. & Wang, X. (2018) Exponential integrability properties of numerical approximation processes for nonlinear stochastic differential equations. *Math. Comp.*, **87**, 1353–1413.

Kelly, C., Lord, G. & Sun, F. (2023) Strong convergence of an adaptive time-stepping Milstein method for SDEs with monotone coefficients. *BIT Numer. Math.*, **63**, 33.

Kloeden, P. E. & Platen, E. (1992) *Numerical Solution of Stochastic Differential Equations*. Berlin: Springer Berlin Heidelberg.

Kumar, C. & Sabanis, S. (2019) On Milstein approximations with varying coefficients: the case of super-linear diffusion coefficients. *BIT Numer. Math.*, **59**, 929–968.

Li, X. & Mao, X. (2009) Population dynamical behavior of non-autonomous Lotka-Volterra competitive system with random perturbation. *Discrete Continuous Dyn. Syst.-Ser. A*, **24**, 523–545.

Li, X., Mao, X. & Yin, G. (2019) Explicit numerical approximations for stochastic differential equations in finite and infinite horizons: truncation methods, convergence in pth moment, and stability. *IMA J. Numer. Anal.*, **39**, 847–892.

Li, Y. & Cao, W. (2023) A positivity preserving Lamperti transformed Euler–Maruyama method for solving the stochastic Lotka–Volterra competition model. *Commun. Nonlinear Sci. Numer. Simul.*, **122**, 107260.

Liu, W. & Mao, X. (2013) Strong convergence of the stopped Euler–Maruyama method for nonlinear stochastic differential equations. *Appl. Math. Comput.*, **223**, 389–400.

AQ9 Mao, X. (2007) *Stochastic Differential Equations and Applications*. Woodhead Publishing Limited.

Mao, X. (2015) The truncated Euler–Maruyama method for stochastic differential equations. *J. Comput. Appl. Math.*, **290**, 370–384.

Mao, X. & Szpruch, L. (2013) Strong convergence rates for backward Euler–Maruyama method for non-linear dissipative-type stochastic differential equations with super-linear diffusion coefficients. *Stochastics*, **85**, 144–171.

Mao, X., Wei, F. & Wiriyakraikul, T. (2021) Positivity preserving truncated Euler–Maruyama method for stochastic Lotka–Volterra competition model. *J. Comput. Appl. Math.*, **394**, 113566. AQ10

Mattingly, J. C., Stuart, A. M. & Higham, D. J. (2002) Ergodicity for SDEs and approximations: locally Lipschitz vector fields and degenerate noise. *Stochastic Process. Appl.*, **101**, 185–232.

Milstein, G. N. & Tretyakov, M. V. (2004) *Stochastic Numerics for Mathematical Physics*. Berlin: Springer.

Neuenkirch, A. & Szpruch, L. (2014) First order strong approximations of scalar SDEs defined in a domain. *Numer. Math.*, **128**, 103–136.

Sabanis, S. (2016) Euler approximations with varying coefficients: the case of superlinearly growing diffusion coefficients. *Ann. Appl. Probab.*, **26**, 2083–2105.

Talay, D. (2002) Stochastic Hamiltonian systems: exponential convergence to the invariant measure, and discretization by the implicit Euler scheme. *Markov Process. Related Fields*, **8**, 163–198.

AQ11 To, C. W. (2000) *Nonlinear Random Vibration: Analytical Techniques and Applications*. Crc Press.

Tretyakov, M. V. & Zhang, Z. (2013) A fundamental mean-square convergence theorem for SDEs with locally Lipschitz coefficients and its applications. *SIAM J. Numer. Anal.*, **51**, 3135–3162.

Wang, X. (2023) Mean-square convergence rates of implicit Milstein type methods for SDEs with non-Lipschitz coefficients. *Adv. Comput. Math.*, **49**, 37.

Wang, X. & Gan, S. (2013) The tamed Milstein method for commutative stochastic differential equations with non-globally Lipschitz continuous coefficients. *J. Differ. Equations Appl.*, **19**, 466–490.

Wang, X., Wu, J. & Dong, B. (2020) Mean-square convergence rates of stochastic theta methods for SDEs under a coupled monotonicity condition. *BIT*, **60**, 759–790.

Zong, X., Wu, F. & Xu, G. (2018) Convergence and stability of two classes of theta-Milstein schemes for stochastic differential equations. *J. Comput. Appl. Math.*, **336**, 8–29.