Unconditionally positivity-preserving approximations of the Aït-Sahalia type model: Explicit Milstein-type schemes

Yingsong Jiang[a], Ruishu Liu[a], Xiaojie Wang[a], Jinghua Zhuo[b] *

[a] School of Mathematics and Statistics, HNP-LAMA, Central South University, Changsha, Hunan, China

[b] Economics Department, University of Warwick, Coventry CV4 7AL, United Kingdom

Abstract

The present article aims to design and analyze efficient first-order strong schemes for a generalized Aït-Sahalia type model arising in mathematical finance and evolving in a positive domain $(0, \infty)$, which possesses a diffusion term with superlinear growth and a highly nonlinear drift that blows up at the origin. Such a complicated structure of the model unavoidably causes essential difficulties in the construction and convergence analysis of time discretizations. By incorporating implicitness in the term $\alpha_{-1}x^{-1}$ and a corrective mapping Φ_h in the recursion, we develop a novel class of explicit and unconditionally positivity-preserving (i.e., for any step-size $h > 0$) Milstein-type schemes for the underlying model. In both non-critical and general critical cases, we introduce a novel approach to analyze mean-square error bounds of the novel schemes, without relying on a priori high-order moment bounds of the numerical approximations. The expected order-one mean-square convergence is attained for the proposed scheme. The above theoretical guarantee can be used to justify the optimal complexity of the Multilevel Monte Carlo method. Numerical experiments are finally provided to verify the theoretical findings.

AMS subject classification: 60H35, 60H15, 65C30.

Keywords: Aït-Sahalia type model; Unconditionally positivity-preserving; Explicit Milstein-type scheme; Order-one mean-square convergence.

1 Introduction

Stochastic differential equations (SDEs) have found extensive applications in various disciplines such as finance, biology, chemistry, physics and engineering. Since analytical solutions for most nonlinear SDEs are typically not accessible, there has been a growing interest in examining their numerical counterparts. The past few decades have witnessed a lot of progress in

*This work was supported by Natural Science Foundation of China (12071488, 12371417, 11971488) and all authors contributed equally to this work.

E-mails: yingsong@csu.edu.cn, chicago@mail.ustc.edu.cn, x.j.wang7@csu.edu.cn, Jinghua.Zhuo@warwick.ac.uk.

this area, where the traditional setting imposed the global Lipschitz condition on the coefficient functions of SDEs [21, 23]. A natural question then arises: what if the restrictive global Lipschitz condition was violated? In 2011, Hutzenthaler, Jentzen and Kloeden [14] gave a negative answer to the question in the sense that the popularly used Euler-Maruyama method produces divergent numerical approximations when used to solve a large class of SDEs with super-linearly growing coefficients. Therefore, it is highly non-trivial to design and analyze numerical SDEs in the absence of the Lipschitz regularity of coefficients. Indeed, most nonlinear SDE models from computational finance not only have non-globally Lipschitz coefficients, but also have positive solutions, which motivates the positivity-preserving numerical approximations. Following this direction, many researchers recently proposed and analyzed various positivity-preserving schemes for nonlinear SDEs with non-globally Lipschitz coefficients [2, 5–13, 15–20, 22, 25–29], to just mention a few.

As one of typical nonlinear SDEs with non-globally Lipschitz coefficients, the Aït-Sahalia interest rate model was widespreadly used in finance and economics, which was initially introduced by Aït Sahalia [1] and later expanded by [24] to a general version, given by

$$dX_t = (\alpha_{-1}X_t^{-1} - \alpha_0 + \alpha_1 X_t - \alpha_2 X_t^r)dt + \sigma X_t^\rho dW_t,\ t > 0, \quad X_0 = x_0 > 0. \tag{1.1}$$

Here $\alpha_{-1}, \alpha_0, \alpha_1, \alpha_2 > 0$, $r, \rho > 1$ such that $r + 1 \geq 2\rho$, and $\{W_t\}_{t\in[0,T]}$ is a standard Brownian motion. As asserted by [24], the considered model (1.1) is well-posed in the domain $(0, \infty)$ and admits a unique positive solution. Clearly, the model has a polynomially growing drift that blows up at the origin and a diffusion term with superlinear growth. These facts cause essential difficulties for the analysis of numerical approximations (see [24] for more comments). Due to the polynomially growing coefficients, the widely used Euler–Maruyama method, also known to be not positivity preserving, is apparently not a good candidate scheme for the model (1.1).

In [24], the authors discretized (1.1) by using the backward Euler (BE) method and obtained positivity-preserving approximations. A strong convergence analysis was conducted there for the BE scheme applied to the model (1.1) with $r + 1 > 2\rho$, but without any convergence rate disclosed. This gap was filled by [26], where the authors achieved the mean-square convergence rate of order 1/2 for stochastic theta methods applied to the Aït-Sahalia type model under the condition $r + 1 \geq 2\rho$, also covering the general critical case $r + 1 = 2\rho$. Recently, a kind of implicit Milstein method was proposed for the Aït-Sahalia type model in [25], where a mean-square convergence rate of order 1 was successfully recovered.

Nevertheless, the implementation of implicit methods is computationally expensive as one needs to solve an implicit algebraic equation in each step. In order to reduce computational costs, some researchers attempt to find explicit positivity-preserving schemes. Based on a Lamperti-type transformation, the authors of [11] proposed an explicit, positivity-preserving scheme with first-order strong convergence, for (1.1) in the special critical case $r = 2$, $\rho = \frac{3}{2}$, which, however, only works for the special case. The recent publication [8] offered a positivity-preserving truncated Euler method for the non-critical case $r + 1 > 2\rho$, with a mean-square convergence rate of nearly 1/4. More recently, the authors of [17] constructed an explicit, unconditionally positivity-preserving Euler-type method, which was proved to achieve a mean-square convergence of order 1/2 for the general case $r + 1 \geq 2\rho$.

In this paper, we aim to introduce a novel class of explicit Milstein-type schemes, which is easily implementable, unconditionally positivity-preserving and strongly convergent with order one. On a uniform mesh within the interval $[0,T]$ with a time step-size $h=\frac{T}{N}, N\in\mathbb{N}$, we propose the following time stepping scheme:

$$\begin{cases} Y_{n+1}=\Phi_h(Y_n)+(\alpha_{-1}Y_{n+1}^{-1}-\alpha_0+\alpha_1\Phi_h(Y_n)+f(\Phi_h(Y_n)))h+g(\Phi_h(Y_n))\Delta W_n \\ \qquad\qquad +\frac{1}{2}(|\Delta W_n|^2-h)\hat{g}(\Phi_h(Y_n)),\quad n\in\{0,1,2,...,N-1\}, \\ Y_0=x_0, \end{cases} \tag{1.2}$$

where we denote $f(x):=-\alpha_2x^r, g(x):=\sigma x^\rho, \hat{g}(x):=g'(x)g(x)=\rho\sigma^2x^{2\rho-1}, x\in(0,+\infty)$ and $\Delta W_n:=W_{t_{n+1}}-W_{t_n}$. Here, the crucial term Φ_h is a kind of corrective mapping depending on the time step-size h and satisfying Assumption 3.1, which is incorporated to tackle the tough issue caused by the polynomially growing coefficients. A typical choice of such operator could be a projection operator $\mathscr{P}_h$ defined by (3.7). In addition, the introduction of the implicit term Y_{n+1}^{-1} is used to preserve the positivity of the original model. Such a partial implicitness is, however, explicitly solved, by finding a positive root of a quadratic equation (see (3.6) below).

It is worthwhile to mention that, identifying a convergence rate of the proposed scheme for the considered model is highly non-trivial, due to a highly nonlinear drift that blows up at the origin, superlinearly growing diffusion coefficients and a mixture of implicitness and explicitness in the drift part of the scheme. Based on the globally monotone condition of f and g in $(0,\infty)$ (Lemma 2.1):

$$(x-y)(f(x)-f(y))+\tfrac{\upsilon-1}{2}|g(x)-g(y)|^2\le L|x-y|^2,\quad \text{for some } \upsilon>2,\quad x,y>0, \tag{1.3}$$

we introduce a novel approach to analyze mean-square error bounds of the new schemes, which does not rely on a priori high-order moment bounds of the numerical approximations. In both non-critical ($r+1>2\rho$) and general critical ($r+1=2\rho$) cases, the expected order-one mean-square convergence is successfully attained for the proposed scheme. More accurately, for the non-critical case $r+1>2\rho$ or the general critical case $r+1=2\rho$ with $\frac{\alpha_2}{\sigma^2}\ge 4r+\frac{1}{2}$, the proposed scheme is proved to achieve first-order convergence in the following sense (see Theorem 4.3):

$$\mathbb{E}\big[|X_{t_n}-Y_n|^2\big]\le Ch^2,\quad \forall\ h=\tfrac{T}{N}>0,\ T>0,\ N\in\mathbb{N}. \tag{1.4}$$

As far as we know, this is the first article to propose and analyze an explicit, unconditionally positivity-preserving method of first-order convergence for the Aït-Sahalia type model (1.1) in both the non-critical and the general critical cases.

The remainder of this article is organized as follows. The next section presents some preliminaries. In Section 3, the numerical scheme and its properties are presented. The mean-square convergence is analyzed in Section 4, with the convergence rate obtained. Numerical experiments are provided to verify the previous theoretical findings in Section 5.

2 Preliminaries

Let $\mathbb{N}$ be the set of all positive integers and C be a positive constant that is independent of the time step-size and may vary at different appearance. Denote the Euclidean norm in $\mathbb{R}$ by $|\cdot|$ and set $T\in(0,\infty)$. Given a filtered probability space $(\Omega,\mathcal{F},\{\mathcal{F}_t\}_{t\in[0,T]},\mathbb{P})$, we use $\mathbb{E}$ to represent the expectation and $L^p(\Omega;\mathbb{R}), p>0$ to represent the space of all $\mathbb{R}$-valued random variables η satisfing $\mathbb{E}[|\eta|^p]<\infty$, equipped with the norm $\|\cdot\|_{L^p(\Omega;\mathbb{R})}$ defined by:

$$\|\eta\|_{L^p(\Omega;\mathbb{R})}:=(\mathbb{E}[|\eta|^p])^{\frac{1}{p}},\quad \forall\ \eta\in L^p(\Omega;\mathbb{R}),\ p>0. \tag{2.1}$$

Let us consider the Aït-Sahalia type model of the following form:

$$dX_t = (\alpha_{-1}X_t^{-1} - \alpha_0 + \alpha_1 X_t + f(X_t))dt + g(X_t)dW_t,\ t > 0, \quad X_0 = x_0 > 0, \tag{2.2}$$

where for short we denote

$$f(x) := -\alpha_2 x^r, \quad g(x) := \sigma x^\rho, \quad x \in D := (0, \infty), \tag{2.3}$$

with $\alpha_{-1}, \alpha_0, \alpha_1, \alpha_2, \sigma > 0$ and $r, \rho > 1$. The monotonicity condition for f and g is presented in the following lemma, whose proof can be found in [25, Lemma 5.9, Lemma 5.12].

Lemma 2.1. *Let f and g be defined by* (2.3)*. If the parameters in the model* (2.2) *satisfy one of the following conditions:*

(1) $r + 1 > 2\rho$,

(2) $r + 1 = 2\rho$, $\frac{\alpha_2}{\sigma^2} > \frac{1}{8}\big(r + 2 + \frac{1}{r}\big)$,

then for all $x, y \in D$, there exist constants $\upsilon > 2$ and $L > 0$ such that

$$(x - y)(f(x) - f(y)) + \tfrac{\upsilon-1}{2}|g(x) - g(y)|^2 \leq L|x - y|^2. \tag{2.4}$$

The well-posedness of the Aït-Sahalia type model (2.2) has been established in [24, Theorem 2.1], also quoted as follows.

Lemma 2.2. *For any given initial data $X_0 = x_0 > 0$, there exists a unique positive $\{\mathcal{F}_t\}_{t\in[0,T]}$-adapted global solution with continuous sample paths $\{X_t\}_{t\geq 0}$ to* (2.2)*.*

Next we revisit some lemmas that give moment bounds for the exact solutions of the Aït-Sahalia type model (2.2). For the non-critical case, the next lemma is quoted from [24, Lemma 2.1].

Lemma 2.3. *Let $r + 1 > 2\rho$. Let $\{X_t\}_{t\in[0,T]}$ be the solution of (2.2), then for any $p_0 \geq 2$ it holds that*

$$\sup_{t\in[0,\infty)} \mathbb{E}[|X_t|^{p_0}] < \infty, \qquad \sup_{t\in[0,\infty)} \mathbb{E}[|X_t|^{-p_0}] < \infty. \tag{2.5}$$

For the general critical case, we quote from [26, Lemma 4.6] the following lemma.

Lemma 2.4. *Let $r + 1 = 2\rho$. Let $\{X_t\}_{t\in[0,T]}$ be the solution of* (2.2)*, then for any $2 \leq p_1 \leq \frac{\sigma^2+2\alpha_2}{\sigma^2}$ and for any $p_2 \geq 2$, it holds that*

$$\sup_{t\in[0,\infty)} \mathbb{E}[|X_t|^{p_1}] < \infty, \qquad \sup_{t\in[0,\infty)} \mathbb{E}[|X_t|^{-p_2}] < \infty. \tag{2.6}$$

Equipped with these moment bounds, the Hölder continuity of the solutions can also be derived, as stated in the following lemmas, quoted directly from [26].

Lemma 2.5. *[26, Lemma 4.4] Let $r+1 > 2\rho$. Then it holds that, for any $p \geq 1$ and $t, s \in [0, T]$,*

$$\|X_t - X_s\|_{L^p(\Omega;\mathbb{R})} \leq C|t - s|^{\frac{1}{2}}, \tag{2.7}$$

$$\|X_t^{-1} - X_s^{-1}\|_{L^p(\Omega;\mathbb{R})} \leq C|t - s|^{\frac{1}{2}}. \tag{2.8}$$

Lemma 2.6. *[26, Lemma 4.7] Let $r+1=2\rho$. Then for any $t,s\in[0,T]$ it holds that*

$$\|X_t - X_s\|_{L^{q_1}(\Omega;\mathbb{R})} \le C|t-s|^{\frac{1}{2}}, \quad 2 \le q_1 \le \tfrac{1}{r}\big(\tfrac{2\alpha_2}{\sigma^2}+1\big), \tag{2.9}$$

$$\|X_t^{-1} - X_s^{-1}\|_{L^{q_2}(\Omega;\mathbb{R})} \le C|t-s|^{\frac{1}{2}}, \quad 2 \le q_2 < \tfrac{1}{r}\big(\tfrac{2\alpha_2}{\sigma^2}+1\big). \tag{2.10}$$

The aforementioned lemmas help deduce the following lemma (cf. [26, (4.21), (4.24)]).

Lemma 2.7. *Let $r+1\ge 2\rho$. If one of the following conditions holds:*

(1) $r+1>2\rho$,

(2) $r+1=2\rho$, $\frac{\alpha_2}{\sigma^2}>2r-\frac{3}{2}$,

then for any $t,s\in[0,T]$ we have

$$\begin{aligned}
&\|f(X_t)-f(X_s)\|_{L^2(\Omega;\mathbb{R})} \le C|t-s|^{\frac{1}{2}},\\
&\|g(X_t)-g(X_s)\|_{L^2(\Omega;\mathbb{R})} \le C|t-s|^{\frac{1}{2}},\\
&\|\hat{g}(X_t)-\hat{g}(X_s)\|_{L^2(\Omega;\mathbb{R})} \le C|t-s|^{\frac{1}{2}},
\end{aligned} \tag{2.11}$$

where f,g are defined as (2.3) *and*

$$\hat{g}(x) := g'(x)g(x) = \rho\sigma^2 x^{2\rho-1}, \quad x\in D. \tag{2.12}$$

3 The proposed explicit Milstein-type scheme

To numerically approximate the model (2.2), we do a temperal discretization on a uniform mesh within the interval $[0,T]$, with a uniform step-size $h=\frac{T}{N}, N\in\mathbb{N}$ and grid points $t_k := kh, k\in\{0,1,...,N\}$. We propose a numerical scheme as follows:

$$\begin{aligned}
Y_{n+1} = \Phi_h(Y_n) + (\alpha_{-1}Y_{n+1}^{-1} - \alpha_0 + \alpha_1\Phi_h(Y_n) + f(\Phi_h(Y_n)))h + g(\Phi_h(Y_n))\Delta W_n\\
+ \tfrac{1}{2}(|\Delta W_n|^2 - h)\hat{g}(\Phi_h(Y_n)), \quad n\in\{0,1,2,...,N-1\},
\end{aligned} \tag{3.1}$$

with $Y_0=x_0$, where $\Delta W_n := W_{t_{n+1}} - W_{t_n}$ is the increment of the Brownian motion and f,g, $\hat{g}$ are defined by (2.3), (2.12). Furthermore, $\Phi_h : D\to D$ is a kind of corrective mapping depending on the time step-size h, required to satisfy the following assumptions.

Assumption 3.1. *Let $f,g,\hat{g}$ be defined as* (2.3), (2.12)*. For all $x,y\in D$ and $h=\frac{T}{N}>0$, there exist constants $L_1,L_2,L_3\ge 0$ such that the operator $\Phi_h: D\to D$ obeys*

$$|\Phi_h(x)| \le |x|, \tag{3.2}$$

$$|x-\Phi_h(x)| \le L_1 h(1+|x|^{2r+1})(1\wedge h(1+|x|^{2r})), \tag{3.3}$$

$$|\Phi_h(x)-\Phi_h(y)| \le (1+L_2h)|x-y|, \tag{3.4}$$

$$|f(\Phi_h(x))-f(\Phi_h(y))|^2 + |\hat{g}(\Phi_h(x))-\hat{g}(\Phi_h(y))|^2 \le L_3 h^{-1}|x-y|^2. \tag{3.5}$$

The well-posedness and preservation of positivity are obvious, by noting that solving (3.1) is nothing but finding a unique positive root of a quadratic equation, explicitly given by

$$\begin{aligned} Y_{n+1} =& \frac{1}{2}\Bigg[\Phi_h(Y_n) + \vartheta_n h + g(\Phi_h(Y_n))\Delta W_n + \tfrac{1}{2}(|\Delta W_n|^2 - h)\hat{g}(\Phi_h(Y_n)) \\ & + \sqrt{\Big(\Phi_h(Y_n) + \vartheta_n h + g(\Phi_h(Y_n))\Delta W_n + \tfrac{1}{2}(|\Delta W_n|^2 - h)\hat{g}(\Phi_h(Y_n))\Big)^2 + 4\alpha_{-1}h}\,\Bigg] > 0, \end{aligned} \tag{3.6}$$

where for short we denote

$$\vartheta_n := -\alpha_0 + \alpha_1 \Phi_h(Y_n) + f(\Phi_h(Y_n)).$$

As a consequence, we have the following lemma.

Lemma 3.2. *For any step-size $h = \frac{T}{N} > 0$, the numerical scheme* (3.1) *is well-defined and positivity preserving, i.e., it admits a unique positive $\{\mathcal{F}_{t_n}\}_{n=0}^N$-adapted solution $\{Y_n\}_{n=0}^N, N \in \mathbb{N}$ for the scheme* (3.1) *given a positive initial data.*

In what follows we provide an exemple operator $\Phi_h : D \to D$ fulfilling Assumption 3.1.

Example 3.3. *Let $f, \hat{g}$ be denoted by* (2.3), (2.12) *and let $r+1 \geq 2\rho$. For any given $q \in [\frac{1}{2r}, \frac{1}{2r-2}]$ we define $\mathscr{P}_h : D \to D$ by*

$$\mathscr{P}_h(x) := \min\{1, h^{-q}|x|^{-1}\}x. \tag{3.7}$$

Such a projection operator was introduced by [3,4] to construct a projection Euler/Milstein schemes for SDEs in non-globally Lipschitz setting. Next we show that $\Phi_h = \mathscr{P}_h$ obeys all conditions in Assumption 3.1. Firstly, one can easily confirm (3.2) and (3.4) with $L_2 = 0$ by observing

$$|\mathscr{P}_h(x)| \leq |x|, \quad |\mathscr{P}_h(x) - \mathscr{P}_h(y)| \leq |x - y|, \quad \forall\, x, y \in D. \tag{3.8}$$

Secondly, for all $x \geq h^{-q}$, one has $\mathscr{P}_h(x) \leq h^{-q} \leq x$. Thus for any $m > 0$ it holds that

$$|x - \mathscr{P}_h(x)| \leq \mathbb{1}_{\{x \geq h^{-q}\}} 2|x| \leq 2h^m |x|^{\frac{m}{q}+1} \leq 2h^m(1 + |x|^{2mr+1}), \quad \forall\, x \in D, \tag{3.9}$$

which confirms (3.3) by taking $L_1 = 2$ and $m = 1$ or $m = 2$. Lastly, since $\mathscr{P}_h(x) \leq h^{-q}$ for all $x \in D$, it is easy to obtain that for all $x, y \in D$,

$$\begin{aligned} \big|f(\mathscr{P}_h(x)) - f(\mathscr{P}_h(y))\big|^2 &= \left| \int_0^1 f'(\theta \mathscr{P}_h(x) + (1-\theta)\mathscr{P}_h(y)) d\theta \cdot (\mathscr{P}_h(x) - \mathscr{P}_h(y)) \right|^2 \\ &\leq \alpha_2^2 r^2 \int_0^1 \big|\theta \mathscr{P}_h(x) + (1-\theta)\mathscr{P}_h(y)\big|^{2r-2} d\theta \cdot \big|\mathscr{P}_h(x) - \mathscr{P}_h(y)\big|^2 \\ &\leq \alpha_2^2 r^2 h^{-q(2r-2)} |x-y|^2 \\ &\leq \alpha_2^2 r^2 h^{1-q(2r-2)} \cdot h^{-1}|x-y|^2 \\ &\leq \alpha_2^2 r^2 T^{1-q(2r-2)} \cdot h^{-1}|x-y|^2, \end{aligned} \tag{3.10}$$

where we used the fact that $1-q(2r-2)\geq 0$. Similarly, we derive that for all $x,y\in D$,

$$\begin{aligned}\big|\hat{g}(\mathscr{P}_h(x))-\hat{g}(\mathscr{P}_h(y))\big|^2 &\leq \int_0^1 |\hat{g}'(\theta\mathscr{P}_h(x)+(1-\theta)\mathscr{P}_h(y))|^2 d\theta\cdot|\mathscr{P}_h(x)-\mathscr{P}_h(y)|^2\\ &\leq \sigma^4\rho^2(2\rho-1)^2\int_0^1 |\theta\mathscr{P}_h(x)+(1-\theta)\mathscr{P}_h(y)|^{4\rho-4}d\theta\cdot|\mathscr{P}_h(x)-\mathscr{P}_h(y)|^2\\ &\leq \sigma^4\rho^2(2\rho-1)^2T^{1-q(2r-2)}\cdot h^{-1}|x-y|^2,\end{aligned}\tag{3.11}$$

where we used the fact that $4\rho-4\leq 2r-2$ in the last inequality. A combination of (3.10) and (3.11) confirms (3.5) with $L_3=\big(\alpha_2^2r^2\vee\sigma^4\rho^2(2\rho-1)^2\big)T^{1-q(2r-2)}$.

In light of the aforementioned evidence, the projection operator $\mathscr{P}_h$ satisfies all condtions in Assumption 3.1.

Armed with the above properties, we can now embark on the error analysis for the proposed scheme in the next section.

4 Order one mean-square convergence

In this section, we focus on the analysis of the mean-square convergence rate of the numerical scheme (3.1). To begin with, we present the subsequent lemma regarding the error caused by the introduce of the corrective mapping Φ_h in f,g and $\hat{g}$.

Lemma 4.1. *Let f,g and $\hat{g}$ be defined by* (2.3) *and* (2.12) *satisfying $r+1\geq 2\rho$. Let Assumption 3.1 hold. For all $x\in D$, there exists a constant C independent of h such that*

$$|f(x)-f(\Phi_h(x))|\vee|g(x)-g(\Phi_h(x))|\vee|\hat{g}(x)-\hat{g}(\Phi_h(x))|\leq Ch(1+|x|^{3r}).\tag{4.1}$$

Proof. By utilizing (3.2) and (3.3), it can be derived that

$$\begin{aligned}|f(x)-f(\Phi_h(x))| &= \left|\int_0^1 f'(\theta x+(1-\theta)\Phi_h(x))d\theta\cdot(x-\Phi_h(x))\right|\\ &\leq \alpha_2 r\int_0^1 |\theta x+(1-\theta)\Phi_h(x)|^{r-1}d\theta\cdot|x-\Phi_h(x)|\\ &\leq \alpha_2 r|x|^{r-1}\cdot|x-\Phi_h(x)|\\ &\leq Ch(1+|x|^{3r}).\end{aligned}\tag{4.2}$$

Noting that $r\geq 2\rho-1>\rho$, one can similarly deduce that

$$|g(x)-g(\Phi_h(x))|\leq\rho\sigma|x|^{\rho-1}\cdot|x-\Phi_h(x)|\leq Ch(1+|x|^{2r+\rho})\leq Ch(1+|x|^{3r}),\tag{4.3}$$

and

$$|\hat{g}(x)-\hat{g}(\Phi_h(x))|\leq\rho(2\rho-1)\sigma^2|x|^{2\rho-2}\cdot|x-\Phi_h(x)|\leq Ch(1+|x|^{3r}).\tag{4.4}$$

The desired assertion can be achieved by combining (4.2), (4.3) with (4.4). □

Noting that

$$\int_{t_n}^{t_{n+1}} \int_{t_n}^{s} dW_l dW_s = \tfrac{1}{2}\big(|\Delta W_n|^2 - h\big), \quad n \in \{0, 1, 2, ..., N-1\},$$

one can rewrite (2.2) as:

$$\begin{aligned} X_{t_{n+1}} = \Phi_h(X_{t_n}) + (\alpha_{-1} X_{t_{n+1}}^{-1} - \alpha_0 + \alpha_1 \Phi_h(X_{t_n}) + f(\Phi_h(X_{t_n})))h + g(\Phi_h(X_{t_n}))\Delta W_n \\ + \tfrac{1}{2}(|\Delta W_n|^2 - h)\hat{g}(\Phi_h(X_{t_n})) + R_{n+1}, \quad \forall\, n \in \{0, 1, ..., N-1\}, \end{aligned} \tag{4.5}$$

where we denote

$$\begin{aligned} R_{n+1} := &\int_{t_n}^{t_{n+1}} [\alpha_{-1} X_s^{-1} - \alpha_{-1} X_{t_{n+1}}^{-1} + \alpha_1 X_s - \alpha_1 \Phi_h(X_{t_n}) + f(X_s) - f(\Phi_h(X_{t_n}))] ds \\ &+ \int_{t_n}^{t_{n+1}} [g(X_s) - g(\Phi_h(X_{t_n})) - \hat{g}(\Phi_h(X_{t_n}))(W_s - W_{t_n})] dW_s + X_{t_n} - \Phi_h(X_{t_n}). \end{aligned} \tag{4.6}$$

We would like to highlight that the introduction of the remainder term R_n plays a crucial role in obtaining the expected convergence rate. First, we need to estimate $\|R_i\|_{L^2(\Omega;\mathbb{R})}$ and $\|\mathbb{E}(R_i|\mathcal{F}_{t_{i-1}})\|_{L^2(\Omega;\mathbb{R})}, i \in \{1, 2, ..., N\}$.

Lemma 4.2. *Let $\{X_t\}_{t\in[0,T]}$ and $\{Y_n\}_{n=0}^N$ be the solutions of* (2.2) *and* (3.1), *respectively. Let Assumption 3.1 hold. If one of the following conditions stands:*

(1) $r + 1 > 2\rho$,

(2) $r + 1 = 2\rho$, $\frac{\alpha_2}{\sigma^2} \geq 4r + \frac{1}{2}$,

then there exists a uniform constant C independent of h, such that for all $n \in \{0, 1, ..., N-1\}$, $N \in \mathbb{N}$,

$$\|R_{n+1}\|_{L^2(\Omega;\mathbb{R})} \leq Ch^{\frac{3}{2}}, \quad \|\mathbb{E}[R_{n+1}|\mathcal{F}_{t_n}]\|_{L^2(\Omega;\mathbb{R})} \leq Ch^2. \tag{4.7}$$

Proof. By the Minkowski inequality, we split $\|R_i\|_{L^2(\Omega;R)}, i \in \{1, 2, ..., N\}$ into three terms as

$$\begin{aligned} \|R_i\|_{L^2(\Omega;\mathbb{R})} \leq &\underbrace{\left\| \int_{t_{i-1}}^{t_i} [\alpha_{-1} X_s^{-1} - \alpha_{-1} X_{t_i}^{-1} + \alpha_1 X_s - \alpha_1 \Phi_h(X_{t_{i-1}}) + f(X_s) - f(\Phi_h(X_{t_{i-1}}))] ds \right\|_{L^2(\Omega;\mathbb{R})}}_{=:I_1} \\ &+ \underbrace{\left\| \int_{t_{i-1}}^{t_i} [g(X_s) - g(\Phi_h(X_{t_{i-1}})) - \hat{g}(\Phi_h(X_{t_{i-1}}))(W_s - W_{t_{i-1}})] dW_s \right\|_{L^2(\Omega;\mathbb{R})}}_{=:I_2} \\ &+ \underbrace{\|X_{t_{i-1}} - \Phi_h(X_{t_{i-1}})\|_{L^2(\Omega;\mathbb{R})}}_{=:I_3}. \end{aligned} \tag{4.8}$$

For the term I_1, utilizing (3.3) and Lemma 4.1 we derive that

$$\begin{aligned} I_1 &\leq \int_{t_{i-1}}^{t_i} \|\alpha_{-1}X_s^{-1} - \alpha_{-1}X_{t_i}^{-1}\|_{L^2(\Omega;\mathbb{R})}ds + \int_{t_{i-1}}^{t_i} \|\alpha_1 X_s - \alpha_1 X_{t_{i-1}}\|_{L^2(\Omega;\mathbb{R})}ds \\ &\quad + h\|\alpha_1 X_{t_{i-1}} - \alpha_1\Phi_h(X_{t_{i-1}})\|_{L^2(\Omega;\mathbb{R})} + \int_{t_{i-1}}^{t_i} \|f(X_s) - f(X_{t_{i-1}})\|_{L^2(\Omega;\mathbb{R})}ds \\ &\quad + h\|f(X_{t_{i-1}}) - f(\Phi_h(X_{t_{i-1}}))\|_{L^2(\Omega;\mathbb{R})} \\ &\leq \alpha_{-1}\int_{t_{i-1}}^{t_i} \|X_s^{-1} - X_{t_i}^{-1}\|_{L^2(\Omega;\mathbb{R})}ds + \alpha_1\int_{t_{i-1}}^{t_i} \|X_s - X_{t_{i-1}}\|_{L^2(\Omega;\mathbb{R})}ds \\ &\quad + \int_{t_{i-1}}^{t_i} \|f(X_s) - f(X_{t_{i-1}})\|_{L^2(\Omega;\mathbb{R})}ds + Ch^2\Big(1 + \sup_{t\in[0,T]} \|X_t\|^{3r}_{L^{6r}(\Omega;\mathbb{R})}\Big) \\ &\leq Ch^{\frac{3}{2}}\Big(1 + h^{\frac{1}{2}} \cdot \sup_{t\in[0,T]} \|X_t\|^{3r}_{L^{6r}(\Omega;\mathbb{R})}\Big), \end{aligned} \tag{4.9}$$

where we used Lemmas 2.5, 2.6 and 2.7 in the last inequality.

For the term I_2, to simplify the denotation, we denote

$$F(x) := \alpha_{-1}x^{-1} - \alpha_0 + \alpha_1 x - \alpha_2 x^r, \quad x \in D. \tag{4.10}$$

Applying the Itô formula to $g(x) = \sigma x^\rho$ and using both the Hölder inequality and the Itô isometry we get

$$\begin{aligned} |I_2|^2 &\leq 2\int_{t_{i-1}}^{t_i} \|g(X_s) - g(X_{t_{i-1}}) - \hat{g}(X_{t_{i-1}})(W_s - W_{t_{i-1}})\|^2_{L^2(\Omega;\mathbb{R})}ds \\ &\quad + 2\int_{t_{i-1}}^{t_i} \|g(X_{t_{i-1}}) - g(\Phi_h(X_{t_{i-1}})) + (\hat{g}(X_{t_{i-1}}) - \hat{g}(\Phi_h(X_{t_{i-1}})))(W_s - W_{t_{i-1}})\|^2_{L^2(\Omega;\mathbb{R})}ds \\ &= 2\int_{t_{i-1}}^{t_i} \Big\| \int_{t_{i-1}}^{s} \Big(g'(X_l)F(X_l) + \frac{1}{2}g''(X_l)g^2(X_l)\Big)dl + \int_{t_{i-1}}^{s} \Big(\hat{g}(X_l) - \hat{g}(X_{t_{i-1}})\Big)dW_l \Big\|^2_{L^2(\Omega;\mathbb{R})} ds \\ &\quad + 2\int_{t_{i-1}}^{t_i} \|g(X_{t_{i-1}}) - g(\Phi_h(X_{t_{i-1}})) + (\hat{g}(X_{t_{i-1}}) - \hat{g}(\Phi_h(X_{t_{i-1}})))(W_s - W_{t_{i-1}})\|^2_{L^2(\Omega;\mathbb{R})}ds \\ &\leq 4h\int_{t_{i-1}}^{t_i}\int_{t_{i-1}}^{s} \|g'(X_l)F(X_l) + \frac{1}{2}g''(X_l)g^2(X_l)\|^2_{L^2(\Omega;\mathbb{R})}dlds \\ &\quad + 4\int_{t_{i-1}}^{t_i}\int_{t_{i-1}}^{s} \|\hat{g}(X_l) - \hat{g}(X_{t_{i-1}})\|^2_{L^2(\Omega;\mathbb{R})}dlds \\ &\quad + 4h\|g(X_{t_{i-1}}) - g(\Phi_h(X_{t_{i-1}}))\|^2_{L^2(\Omega;\mathbb{R})} + 4h^2\|\hat{g}(X_{t_{i-1}}) - \hat{g}(\Phi_h(X_{t_{i-1}}))\|^2_{L^2(\Omega;\mathbb{R})} \\ &\leq Ch^3\Big(1 + \mathbb{1}_{\{\rho<2\}}\sup_{t\in[0,T]} \|X_t^{-1}\|^{4-2\rho}_{L^{4-2\rho}(\Omega;\mathbb{R})} + \sup_{t\in[0,T]} \|X_t\|^{2(r+\rho-1)}_{L^{2(r+\rho-1)}(\Omega;\mathbb{R})} + \sup_{t\in[0,T]} \|X_t\|^{6r}_{L^{6r}(\Omega;\mathbb{R})}\Big), \end{aligned} \tag{4.11}$$

where we used Lemmas 2.7, 4.1 in the last inequality. Observing that $4-2\rho < 2$ and $2(r+\rho-1) < 6r$ and using the Lyapunov inequality imply that

$$\mathbb{1}_{\{\rho<2\}}\|X_t^{-1}\|_{L^{4-2\rho}(\Omega;\mathbb{R})} \leq \|X_t^{-1}\|_{L^2(\Omega;\mathbb{R})}, \quad \|X_t\|_{L^{2(r+\rho-1)}(\Omega;\mathbb{R})} \leq \|X_t\|_{L^{6r}(\Omega;\mathbb{R})}, \quad \forall\ t \in [0,T].$$

As a result, one can deduce from (4.11) that

$$|I_2|^2 \le Ch^3\Big(1 + \|X_t^{-1}\|_{L^2(\Omega;\mathbb{R})} + \sup_{t\in[0,T]} \|X_t\|^{6r}_{L^{6r}(\Omega;\mathbb{R})}\Big). \tag{4.12}$$

In view of (3.3), one infers

$$I_3 \le Ch^2\Big(1 + \sup_{t\in[0,T]} \|X_t\|^{4r+1}_{L^{8r+2}(\Omega;\mathbb{R})}\Big). \tag{4.13}$$

The Lyapunov inequality together with the fact that $6r < 8r+2$ helps us obtain

$$\|X_t\|_{L^{6r}(\Omega;\mathbb{R})} \le \|X_t\|_{L^{8r+2}(\Omega;\mathbb{R})}, \quad \forall\, t \in [0,T].$$

A combination of (4.9), (4.12) with (4.13) yields

$$\|R_i\|_{L^2(\Omega;\mathbb{R})} \le Ch^{\frac{3}{2}}\Big(1 + \|X_t^{-1}\|_{L^2(\Omega;\mathbb{R})} + \sup_{t\in[0,T]} \|X_t\|^{4r+1}_{L^{8r+2}(\Omega;\mathbb{R})}\Big). \tag{4.14}$$

Since $\frac{\alpha_2}{\sigma^2} \ge 4r + \frac{1}{2}$ implies $8r+2 \le \frac{\sigma^2+2\alpha_2}{\sigma^2}$, by Lemma 2.3 and Lemma 2.4 we arrive at

$$\|R_{n+1}\|_{L^2(\Omega;\mathbb{R})} \le Ch^{\frac{3}{2}}, \quad \forall\, n \in \{0,1,...,N-1\}, N \in \mathbb{N}, \tag{4.15}$$

for the non-critical case $r+1 > 2\rho$ and the general critical case $r+1 = 2\rho$ with $\frac{\alpha_2}{\sigma^2} \ge 4r + \frac{1}{2}$.

Now we turn to the estimations for $\|\mathbb{E}(R_i|\mathcal{F}_{i-1})\|_{L^2(\Omega;\mathbb{R})}, i \in \{1,2,...,N\}$. Using basic properties of the conditional expectation, one has

$$\mathbb{E}\bigg(\int_{t_{i-1}}^{t_i} \big[g(X_s) - g(\Phi_h(X_{t_{i-1}})) - \hat{g}(\Phi_h(X_{t_{i-1}}))(W_s - W_{t_{i-1}})\big]dW_s\Big|\mathcal{F}_{t_{i-1}}\bigg) = 0. \tag{4.16}$$

Therefore, using the Itô formula twice to $\alpha_{-1}x^{-1}$ and $\alpha_1 x + f(x), x \in D$, the Hölder inequality and the fact that the Itô integrals vanish under the conditional expectation, one can show that

$$\begin{aligned}
\|\mathbb{E}(R_i|\mathcal{F}_{i-1})\|_{L^2(\Omega;\mathbb{R})} \le{}& \|X_{t_{i-1}} - \Phi_h(X_{t_{i-1}})\|_{L^2(\Omega;\mathbb{R})} + \int_{t_{i-1}}^{t_i} \Big\|\mathbb{E}\Big(\alpha_{-1}X_s^{-1} - \alpha_{-1}X_{t_i}^{-1}\big|\mathcal{F}_{t_{i-1}}\Big)\Big\|_{L^2(\Omega;\mathbb{R})} ds \\
&+ \int_{t_{i-1}}^{t_i} \Big\|\mathbb{E}\Big(\alpha_1 X_s + f(X_s) - \alpha_1 X_{t_{i-1}} - f(X_{t_{i-1}})\big|\mathcal{F}_{t_{i-1}}\Big)\Big\|_{L^2(\Omega;\mathbb{R})} ds \\
&+ \int_{t_{i-1}}^{t_i} \big\|\alpha_1 X_{t_{i-1}} - \alpha_1\Phi_h(X_{t_{i-1}}) + f(X_{t_{i-1}}) - f(\Phi_h(X_{t_{i-1}}))\big\|_{L^2(\Omega;\mathbb{R})} ds \\
\le{}& (1+\alpha_1 h)\|X_{t_{i-1}} - \Phi_h(X_{t_{i-1}})\|_{L^2(\Omega;\mathbb{R})} \\
&+ \int_{t_{i-1}}^{t_i}\int_s^{t_i} \big\|\alpha_{-1}X_l^{-2}F(X_l) - \alpha_{-1}X_l^{-3}g^2(X_l)\big\|_{L^2(\Omega;\mathbb{R})} dl ds \\
&+ \int_{t_{i-1}}^{t_i}\int_{t_{i-1}}^{s} \big\|(\alpha_1 + f'(X_l))F(X_l) + \tfrac{1}{2}f''(X_l)g^2(X_l)\big\|_{L^2(\Omega;\mathbb{R})} dl ds \\
&+ h\big\|f(X_{t_{i-1}}) - f(\Phi_h(X_{t_{i-1}}))\big\|_{L^2(\Omega;\mathbb{R})} \\
\le{}& Ch^2\Big(1 + \sup_{t\in[0,T]} \big\|X_t^{-3}\big\|_{L^2(\Omega;\mathbb{R})} + \sup_{t\in[0,T]} \big\|X_t\big\|^{4r+1}_{L^{8r+2}(\Omega;\mathbb{R})}\Big),
\end{aligned} \tag{4.17}$$

where we also used (3.3) and Lemma 4.1 in the last inequality. Similar to (4.14), employing Lemma 2.3 for the non-critical case and Lemma 2.4 for the general critical case with $\frac{\alpha_2}{\sigma^2} \geq 4r + \frac{1}{2}$ leads to

$$\|\mathbb{E}[R_{n+1}|\mathcal{F}_n]\|_{L^2(\Omega;\mathbb{R})} \leq Ch^2, \quad \forall\, n \in \{0,1,...,N-1\}, N \in \mathbb{N}. \tag{4.18}$$

The proof is thus completed. $\square$

At the moment, we are well prepared to identify the expected order-one mean-square convergence of the novel schemes.

Theorem 4.3. *Let $\{X_t\}_{t\in[0,T]}$ and $\{Y_n\}_{n=0}^N$ be the solutions of (2.2) and (3.1), respectively. Let Assumption 3.1 hold. If one of the following conditions stands:*

(1) $r+1 > 2\rho$,

(2) $r+1 = 2\rho$, $\frac{\alpha_2}{\sigma^2} \geq 4r + \frac{1}{2}$,

then there exists a uniform constant C independent of h, such that for all $n \in \{1,2,...,N\}$, $N \in \mathbb{N}$,

$$\|X_{t_n} - Y_n\|_{L^2(\Omega;\mathbb{R})} \leq Ch. \tag{4.19}$$

Proof. For brevity, for all $k \in \{0,1,2,...,N\}$ we denote

$$\begin{aligned} e_k := X_{t_k} - Y_k, \quad \Delta\Phi_{h,k}^{X,Y} := \Phi_h(X_{t_k}) - \Phi_h(Y_k), \quad \Delta f_k^{\Phi_h,X,Y} := f(\Phi_h(X_{t_k})) - f(\Phi_h(Y_k)), \\ \Delta g_k^{\Phi_h,X,Y} := g(\Phi_h(X_{t_k})) - g(\Phi_h(Y_k)), \quad \Delta\hat{g}_k^{\Phi_h,X,Y} := \hat{g}(\Phi_h(X_{t_k})) - \hat{g}(\Phi_h(Y_k)). \end{aligned} \tag{4.20}$$

Bearing (4.5) and (4.20) in mind and subtracting (3.1) from (4.5) infer that for all $n \in \{0,1,2,...,N-1\}$,

$$\begin{aligned} e_{n+1} - h\cdot\alpha_{-1}(X_{t_{n+1}}^{-1} - Y_{n+1}^{-1}) = (1+h\alpha_1)\Delta\Phi_{h,n}^{X,Y} + h\Delta f_n^{\Phi_h,X,Y} + \Delta g_n^{\Phi_h,X,Y}\Delta W_n \\ + \tfrac{1}{2}(|\Delta W_n|^2 - h)\Delta\hat{g}_n^{\Phi_h,X,Y} + R_{n+1}. \end{aligned} \tag{4.21}$$

Squaring both sides of (4.21) yields that

$$\begin{aligned} &|e_{n+1} - h\cdot\alpha_{-1}(X_{t_{n+1}}^{-1} - Y_{n+1}^{-1})|^2 \\ &= (1+h\alpha_1)^2|\Delta\Phi_{h,n}^{X,Y}|^2 + h^2|\Delta f_n^{\Phi_h,X,Y}|^2 + |\Delta g_n^{\Phi_h,X,Y}\Delta W_n|^2 + \tfrac{1}{4}|(|\Delta W_n|^2 - h)\Delta\hat{g}_n^{\Phi_h,X,Y}|^2 \\ &\quad + |R_{n+1}|^2 + 2h(1+h\alpha_1)\Delta\Phi_{h,n}^{X,Y}\Delta f_n^{\Phi_h,X,Y} + 2(1+h\alpha_1)\Delta\Phi_{h,n}^{X,Y}\Delta g_n^{\Phi_h,X,Y}\Delta W_n \\ &\quad + (1+h\alpha_1)\Delta\Phi_{h,n}^{X,Y}(|\Delta W_n|^2 - h)\Delta\hat{g}_n^{\Phi_h,X,Y} + 2(1+h\alpha_1)\Delta\Phi_{h,n}^{X,Y}R_{n+1} \\ &\quad + 2h\Delta f_n^{\Phi_h,X,Y}\Delta g_n^{\Phi_h,X,Y}\Delta W_n + h\Delta f_n^{\Phi_h,X,Y}(|\Delta W_n|^2 - h)\Delta\hat{g}_n^{\Phi_h,X,Y} + 2h\Delta f_n^{\Phi_h,X,Y}R_{n+1} \\ &\quad + \Delta g_n^{\Phi_h,X,Y}\Delta W_n(|\Delta W_n|^2 - h)\Delta\hat{g}_n^{\Phi_h,X,Y} + 2\Delta g_n^{\Phi_h,X,Y}\Delta W_n R_{n+1} \\ &\quad + (|\Delta W_n|^2 - h)\Delta\hat{g}_n^{\Phi_h,X,Y}R_{n+1}. \end{aligned} \tag{4.22}$$

Before proceeding further, one first notes that

$$
\begin{aligned}
&|e_{n+1} - h\cdot\alpha_{-1}(X_{t_{n+1}}^{-1} - Y_{n+1}^{-1})|^2 \\
&= |e_{n+1}|^2 - 2h\alpha_{-1}e_{n+1}(X_{t_{n+1}}^{-1} - Y_{n+1}^{-1}) + h^2\alpha_{-1}^2|X_{t_{n+1}}^{-1} - Y_{n+1}^{-1}|^2 \\
&= |e_{n+1}|^2 - 2h\alpha_{-1}|e_{n+1}|^2\int_0^1 (-1)\cdot|Y_{n+1}+\theta(X_{t_{n+1}} - Y_{n+1})|^{-2}d\theta + h^2\alpha_{-1}^2|X_{t_{n+1}}^{-1} - Y_{n+1}^{-1}|^2 \\
&\geq |e_{n+1}|^2.
\end{aligned}
\tag{4.23}
$$

In addition, the facts

$$
\mathbb{E}\big[|\Delta W_n|^2\big] = h, \quad \mathbb{E}\big[(\Delta W_n)^3\big] = 0, \quad \mathbb{E}\big[|\Delta W_n|^4\big] = 3h^2, \tag{4.24}
$$

help us deduce

$$
\begin{aligned}
&\mathbb{E}\Big[\big|\Delta g_n^{\Phi_h,X,Y}\Delta W_n\big|^2\Big] = \mathbb{E}\Big[\big|\Delta g_n^{\Phi_h,X,Y}\big|^2\Big]\cdot\mathbb{E}\Big[\big|\Delta W_n\big|^2\Big] = h\mathbb{E}\Big[\big|\Delta g_n^{\Phi_h,X,Y}\big|^2\Big], \\
&\mathbb{E}\Big[\big|(|\Delta W_n|^2 - h)\Delta\hat{g}_n^{\Phi_h,X,Y}\big|^2\Big] = \mathbb{E}\Big[\big|(|\Delta W_n|^2 - h)\big|^2\Big]\cdot\mathbb{E}\Big[\big|\Delta\hat{g}_n^{\Phi_h,X,Y}\big|^2\Big] = 2h^2\mathbb{E}\Big[\big|\Delta\hat{g}_n^{\Phi_h,X,Y}\big|^2\Big],
\end{aligned}
$$

and

$$
\begin{aligned}
&\mathbb{E}\Big[\Delta\Phi_{h,n}^{X,Y}\Delta g_n^{\Phi_h,X,Y}\Delta W_n\Big] = 0, \quad \mathbb{E}\Big[\Delta\Phi_{h,n}^{X,Y}\big(|\Delta W_n|^2 - h\big)\Delta\hat{g}_n^{\Phi_h,X,Y}\Big] = 0, \\
&\mathbb{E}\Big[\Delta f_n^{\Phi_h,X,Y}\Delta g_n^{\Phi_h,X,Y}\Delta W_n\Big] = 0, \quad \mathbb{E}\Big[\Delta f_n^{\Phi_h,X,Y}\big(|\Delta W_n|^2 - h\big)\Delta\hat{g}_n^{\Phi_h,X,Y}\Big] = 0, \\
&\mathbb{E}\Big[\Delta g_n^{\Phi_h,X,Y}\Delta W_n\big(|\Delta W_n|^2 - h\big)\Delta\hat{g}_n^{\Phi_h,X,Y}\Big] = 0,
\end{aligned}
$$

where we noted the terms $\Delta f_n^{\Phi_h,X,Y}$, $\Delta g_n^{\Phi_h,X,Y}$, $\Delta\hat{g}_n^{\Phi_h,X,Y}$ and $\Delta\Phi_{h,n}^{X,Y}$ are $\mathcal{F}_{t_n}$-measurable by construction. The Young inequality also implies that for all $n \in \{0, 1, 2, ..., N-1\}$,

$$
2h^2\alpha_1\mathbb{E}\Big[\Delta\Phi_{h,n}^{X,Y}\Delta f_n^{\Phi_h,X,Y}\Big] \leq h^2\alpha_1^2\mathbb{E}\Big[|\Delta\Phi_{h,n}^{X,Y}|^2\Big] + h^2\mathbb{E}\Big[|\Delta f_n^{\Phi_h,X,Y}|^2\Big], \tag{4.25}
$$

$$
2h\mathbb{E}\Big[\Delta f_n^{\Phi_h,X,Y}R_{n+1}\Big] \leq h^2\mathbb{E}\Big[|\Delta f_n^{\Phi_h,X,Y}|^2\Big] + \mathbb{E}\Big[|R_{n+1}|^2\Big], \tag{4.26}
$$

$$
\mathbb{E}\Big[\big(|\Delta W_n|^2 - h\big)\Delta\hat{g}_n^{\Phi_h,X,Y}R_{n+1}\Big] \leq \tfrac{1}{2}h^2\mathbb{E}\Big[|\Delta\hat{g}_n^{\Phi_h,X,Y}|^2\Big] + \mathbb{E}\Big[|R_{n+1}|^2\Big], \tag{4.27}
$$

$$
2\mathbb{E}\Big[\Delta g_n^{\Phi_h,X,Y}\Delta W_n R_{n+1}\Big] \leq h(\upsilon-2)\mathbb{E}\Big[|\Delta g_n^{\Phi_h,X,Y}|^2\Big] + \tfrac{1}{\upsilon-2}\mathbb{E}\Big[|R_{n+1}|^2\Big], \tag{4.28}
$$

with $\upsilon > 2$ coming from Lemma 2.1. Accordingly, by taking expectations on both sizes of

(4.22) and utilizing the above estimates, one immediately arrives at

$$
\begin{aligned}
\mathbb{E}\big[|e_{n+1}|^2\big] \leq\ & (1+h\alpha_1)^2\mathbb{E}\Big[\big|\Delta\Phi_{h,n}^{X,Y}\big|^2\Big] + h^2\mathbb{E}\Big[\big|\Delta f_n^{\Phi_h,X,Y}\big|^2\Big] + h\mathbb{E}\Big[\big|\Delta g_n^{\Phi_h,X,Y}\big|^2\Big] \\
& + \tfrac{1}{2}h^2\mathbb{E}\Big[\big|\Delta\hat{g}_n^{\Phi_h,X,Y}\big|^2\Big] + \mathbb{E}\Big[\big|R_{n+1}\big|^2\Big] + 2h(1+h\alpha_1)\mathbb{E}\Big[\Delta\Phi_{h,n}^{X,Y}\Delta f_n^{\Phi_h,X,Y}\Big] \\
& + 2(1+h\alpha_1)\mathbb{E}\Big[\Delta\Phi_{h,n}^{X,Y}R_{n+1}\Big] + 2h\mathbb{E}\Big[\Delta f_n^{\Phi_h,X,Y}R_{n+1}\Big] + 2\mathbb{E}\Big[\Delta g_n^{\Phi_h,X,Y}\Delta W_nR_{n+1}\Big] \\
& + \mathbb{E}\Big[\big(|\Delta W_n|^2-h\big)\Delta\hat{g}_n^{\Phi_h,X,Y}R_{n+1}\Big] \\
\leq\ & \big[(1+h\alpha_1)^2+h^2\alpha_1^2\big]\mathbb{E}\Big[\big|\Delta\Phi_{h,n}^{X,Y}\big|^2\Big] + 3h^2\mathbb{E}\Big[\big|\Delta f_n^{\Phi_h,X,Y}\big|^2\Big] + h(\upsilon-1)\mathbb{E}\Big[\big|\Delta g_n^{\Phi_h,X,Y}\big|^2\Big] \\
& + h^2\mathbb{E}\Big[\big|\Delta\hat{g}_n^{\Phi_h,X,Y}\big|^2\Big] + (3+\tfrac{1}{\upsilon-2})\mathbb{E}\Big[\big|R_{n+1}\big|^2\Big] + 2h\mathbb{E}\Big[\Delta\Phi_{h,n}^{X,Y}\Delta f_n^{\Phi_h,X,Y}\Big] \\
& + 2(1+h\alpha_1)\mathbb{E}\Big[\Delta\Phi_{h,n}^{X,Y}R_{n+1}\Big] \\
=\ & \big[(1+h\alpha_1)^2+h^2\alpha_1^2\big]\mathbb{E}\Big[\big|\Delta\Phi_{h,n}^{X,Y}\big|^2\Big] + (3+\tfrac{1}{\upsilon-2})\mathbb{E}\Big[\big|R_{n+1}\big|^2\Big] \\
& + 2h\bigg(\tfrac{\upsilon-1}{2}\mathbb{E}\Big[\big|\Delta g_n^{\Phi_h,X,Y}\big|^2\Big] + \mathbb{E}\Big[\Delta\Phi_{h,n}^{X,Y}\Delta f_n^{\Phi_h,X,Y}\Big]\bigg) \\
& + h^2\bigg(\mathbb{E}\Big[\big|\Delta\hat{g}_n^{\Phi_h,X,Y}\big|^2\Big] + 3\mathbb{E}\Big[\big|\Delta f_n^{\Phi_h,X,Y}\big|^2\Big]\bigg) \\
& + 2(1+h\alpha_1)\mathbb{E}\Big[\Delta\Phi_{h,n}^{X,Y}R_{n+1}\Big].
\end{aligned}
\tag{4.29}
$$

Furthermore, we note that all conditions required by Lemma 2.1 are fulfilled, as

$$
\tfrac{\alpha_2}{\sigma^2} \geq 4r+\tfrac{1}{2} = \tfrac{r}{8}+\tfrac{31}{8}(r-1)+\tfrac{35}{8} > \tfrac{1}{8}(r+35) > \tfrac{1}{8}\big(r+2+\tfrac{1}{r}\big). \tag{4.30}
$$

Therefore, using Lemma 2.1 and (3.5) shows

$$
\begin{aligned}
\mathbb{E}\big[|e_{n+1}|^2\big] \leq\ & \big[(1+h\alpha_1)^2+h^2\alpha_1^2+2Lh+3L_3h\big]\mathbb{E}\Big[\big|\Delta\Phi_{h,n}^{X,Y}\big|^2\Big] + (3+\tfrac{1}{\upsilon-2})\mathbb{E}\Big[\big|R_{n+1}\big|^2\Big] \\
& + 2(1+h\alpha_1)\mathbb{E}\Big[\Delta\Phi_{h,n}^{X,Y}R_{n+1}\Big] \\
=\ & \big[(1+h\alpha_1)^2+h^2\alpha_1^2+2Lh+3L_3h\big]\mathbb{E}\Big[\big|\Delta\Phi_{h,n}^{X,Y}\big|^2\Big] + (3+\tfrac{1}{\upsilon-2})\mathbb{E}\Big[\big|R_{n+1}\big|^2\Big] \\
& + 2(1+h\alpha_1)\mathbb{E}\Big[\Delta\Phi_{h,n}^{X,Y}\cdot\mathbb{E}\big[R_{n+1}|\mathcal{F}_{t_n}\big]\Big],
\end{aligned}
\tag{4.31}
$$

where we also used the property of the conditional expectation in the last equality. Further, utilizing the Young inequality gives

$$
\begin{aligned}
\mathbb{E}\big[|e_{n+1}|^2\big] \leq\ & \big[(1+h\alpha_1)^2+h^2\alpha_1^2+2Lh+3L_3h\big]\mathbb{E}\Big[\big|\Delta\Phi_{h,n}^{X,Y}\big|^2\Big] + (3+\tfrac{1}{\upsilon-2})\mathbb{E}\Big[\big|R_{n+1}\big|^2\Big] \\
& + h(1+h\alpha_1)\mathbb{E}\Big[\big|\Delta\Phi_{h,n}^{X,Y}\big|^2\Big] + \tfrac{1+h\alpha_1}{h}\mathbb{E}\Big[\big|\mathbb{E}\big[R_{n+1}|\mathcal{F}_{t_n}\big]\big|^2\Big] \\
=\ & \big[(1+h\alpha_1)^2+h^2\alpha_1^2+2Lh+3L_3h+h(1+h\alpha_1)\big]\mathbb{E}\Big[\big|\Delta\Phi_{h,n}^{X,Y}\big|^2\Big] \\
& + (3+\tfrac{1}{\upsilon-2})\mathbb{E}\Big[\big|R_{n+1}\big|^2\Big] + \tfrac{1+h\alpha_1}{h}\mathbb{E}\Big[\big|\mathbb{E}\big[R_{n+1}|\mathcal{F}_{t_n}\big]\big|^2\Big].
\end{aligned}
\tag{4.32}
$$

In view of Lemma 4.2 and (3.4), we obtain

$$\begin{aligned}\mathbb{E}\big[|e_{n+1}|^2\big] &\leq \big[(1+h\alpha_1)^2 + h^2\alpha_1^2 + 2Lh + 3L_3h + h(1+h\alpha_1)\big]\mathbb{E}\big[|e_n|^2\big] \\ &\quad + (3+\tfrac{1}{v-2})Ch^3 + \tfrac{1+h\alpha_1}{h}Ch^4 \\ &\leq (1+Ch)\mathbb{E}\big[|e_n|^2\big] + Ch^3.\end{aligned} \tag{4.33}$$

By iteration and the observation of $e_0 = 0$, we finally arrive at

$$\begin{aligned}\mathbb{E}\big[|e_{n+1}|^2\big] &\leq (1+Ch)^{n+1}\mathbb{E}\big[|e_0|^2\big] + Ch^3\sum_{i=0}^{n}(1+Ch)^i \\ &\leq e^{CT}\mathbb{E}\big[|e_0|^2\big] + Ch^2Te^{CT} \leq Ch^2.\end{aligned} \tag{4.34}$$

The proof is thus completed.

□

5 Numerical experiments

In this section, some numerical experiments are presented to verify the previous theoretical findings, for the approximation of the Aït-Sahalia type model

$$\begin{cases} dX_t = (\alpha_{-1}X_t^{-1} - \alpha_0 + \alpha_1 X_t - \alpha_2 X_t^r)dt + \sigma X_t^\rho dW_t, \quad t \in [0,T], \quad T=1, \\ X_0 = 0.5. \end{cases} \tag{5.1}$$

To be more specific, we conduct numerical experiments by implementing the following semi-implicit projected Milstein method (SIPMM):

$$\begin{aligned} Y_{n+1} = \Phi_h(Y_n) + (\alpha_{-1}Y_{n+1}^{-1} - \alpha_0 + \alpha_1\Phi_h(Y_n) - \alpha_2(\Phi_h(Y_n))^r)h + \sigma(\Phi_h(Y_n))^\rho \Delta W_n \\ + \tfrac{1}{2}\rho\sigma^2(\Phi_h(Y_n))^{2\rho-1}(|\Delta W_n|^2 - h), \quad n \in \{0,1,2,...,N-1\}, \end{aligned} \tag{5.2}$$

where we take $\Phi_h(x) = \min\{1, h^{-\frac{1}{2r-2}}|x|^{-1}\}x$, $x \in D$, so that the mapping coincides with Example 3.3 with $q = \frac{1}{2r-2}$. Three sets of parameters are carefully chosen to meet the required conditions required by Theorem 4.3.

Example 1 (non-critical case $r+1 > 2\rho$): $\alpha_{-1} = \frac{3}{2}, \alpha_0 = 2, \alpha_1 = 1, \alpha_2 = 13, \sigma = 1, r = 4, \rho = 2$;

Example 2 (critical case $r = 3, \rho = 2$): $\alpha_{-1} = \frac{3}{2}, \alpha_0 = 2, \alpha_1 = 1, \alpha_2 = 13, \sigma = 1, r = 3, \rho = 2$;

Example 3 (critical case $r = 2, \rho = 1.5$): $\alpha_{-1} = \frac{3}{2}, \alpha_0 = 2, \alpha_1 = 1, \alpha_2 = 13, \sigma = 1, r = 2, \rho = 1.5$.

The backward Euler method (BEM) investigated in [26] and the semi-implicit tamed Euler method (SITEM) proposed in [17] for the model (1.1), proved to be strongly convergent with order 0.5, are also implemented for comparison. Numerical approximations produced by BEM

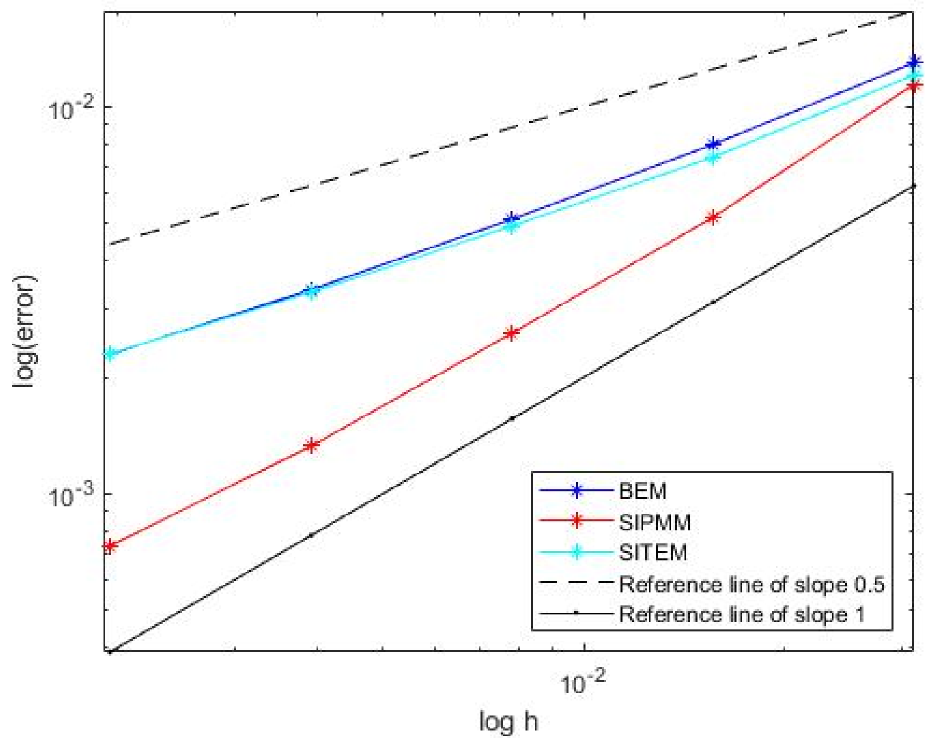

Fig. 1: Mean-square convergence rates for Example 1

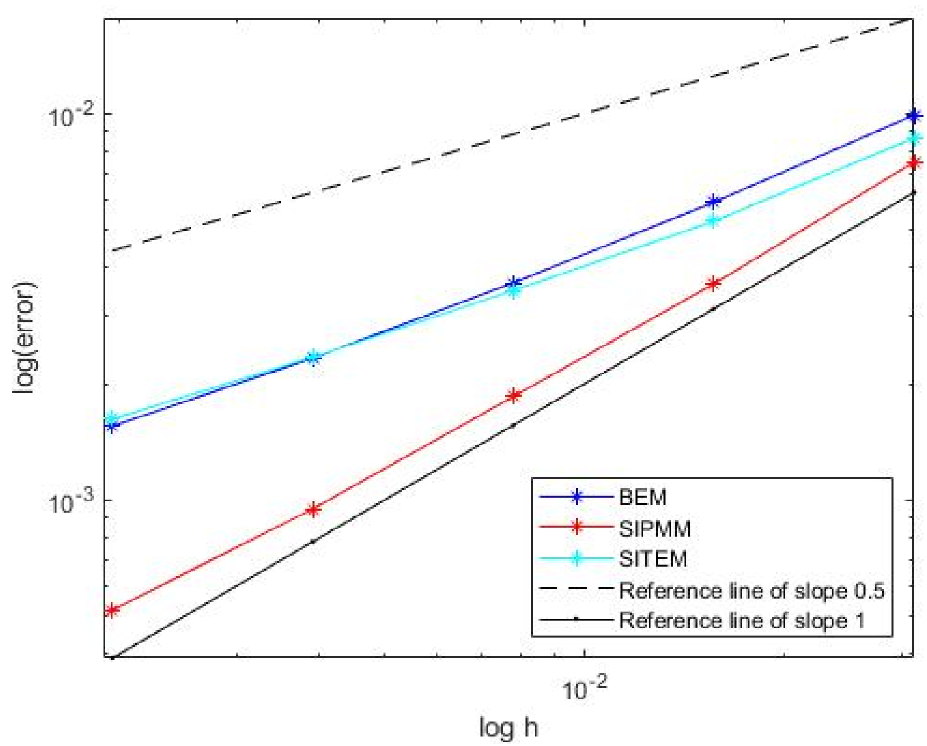

Fig. 2: Mean-square convergence rates for Example 2

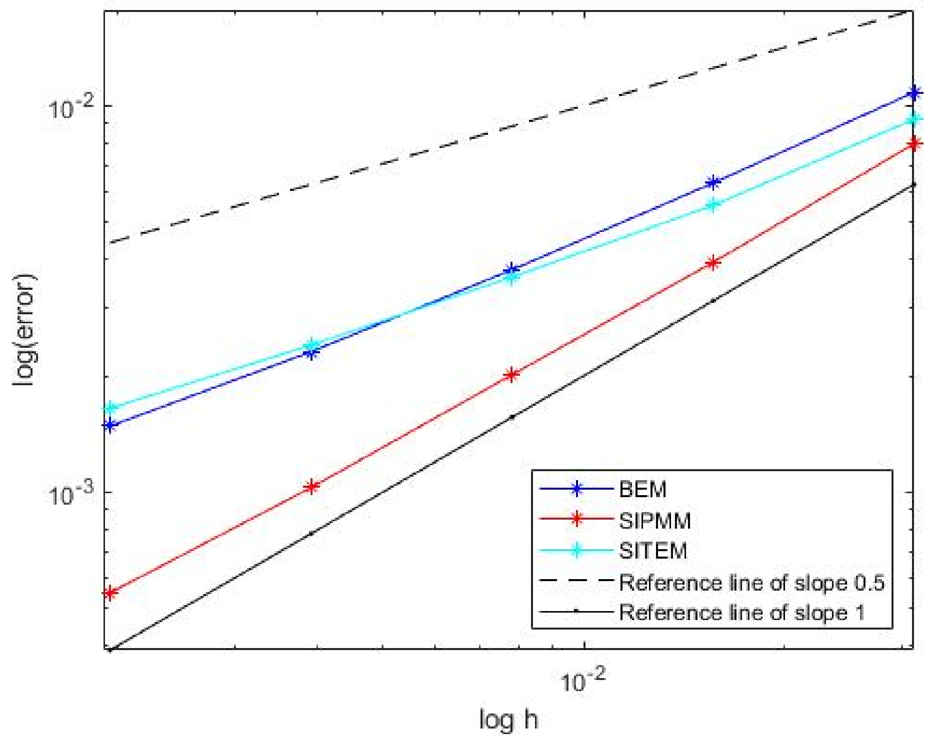

Fig. 3: Mean-square convergence rates for Example 3

with a fine step-size $h_{\text{exact}} = 2^{-14}$ are identified with "exact" solutions, while various step-sizes $h = 2^{-i}, i = 5, 6, 7, 8, 9$ are used for numerical approximations. These two schemes together with our method are tested for the above three examples. The expectation appearing in the mean-square error is approximated by calculating averages over 10000 paths in the following numerical tests.

In Fig.1, Fig.2 and Fig.3, the mean-square convergence rates of three methods (i.e., BEM, SITEM and SIPMM) are depicted on a log-log scale. There one can easily see that the mean-square convergence rates of both BEM and SITEM are close to 0.5, as opposed to a convergence rate close to 1 for SIPMM. This can be detected more transparently from Table 1, which confirms the theoretical results of Theorem 4.3.

Table 1: A least square fit for the convergence rate q

	BEM	SIPMM	SITEM
Example 1	$q = 0.6248$, $resid = 0.0615$	$q = 0.9882$, $resid = 0.1127$	$q = 0.5938$, $resid = 0.0723$
Example 2	$q = 0.6681$, $resid = 0.0696$	$q = 0.9628$, $resid = 0.0685$	$q = 0.5994$, $resid = 0.0749$
Example 3	$q = 0.7184$, $resid = 0.0602$	$q = 0.9651$, $resid = 0.0438$	$q = 0.6163$, $resid = 0.0810$

References

[1] Y. Aït-Sahalia. Testing continuous-time models of the spot interest rate. *The Review of Financial Studies*, 9(2):385–426, 1996.

[2] A. Alfonsi. Strong order one convergence of a drift implicit Euler scheme: Application to the CIR process. *Statistics & Probability Letters*, 83(2):602–607, 2013.

[3] W. Beyn, E. Isaak, and R. Kruse. Stochastic C-stability and B-consistency of explicit and implicit Euler-type schemes. *Journal of Scientific Computing*, 67:955–987, 2016.

[4] W. Beyn, E. Isaak, and R. Kruse. Stochastic C-stability and B-consistency of explicit and implicit Milstein-type schemes. *Journal of Scientific Computing*, 70:1042–1077, 2017.

[5] Y. Cai, Q. Guo, and X. Mao. Positivity preserving truncated scheme for the stochastic Lotka–Volterra model with small moment convergence. *Calcolo*, 60:24, 2023.

[6] J. Chassagneux, A. Jacquier, and I. Mihaylov. An explicit Euler scheme with strong rate of convergence for financial SDEs with non-Lipschitz coefficients. *SIAM Journal on Financial Mathematics*, 7(1):993–1021, 2016.

[7] L. Chen, S. Gan, and X. Wang. First order strong convergence of an explicit scheme for the stochastic SIS epidemic model. *Journal of Computational and Applied Mathematics*, 392:113482, 2021.

[8] S. Deng, C. Fei, W. Fei, and X. Mao. Positivity-preserving truncated Euler–Maruyama method for generalised Ait-Sahalia-type interest model. *BIT Numerical Mathematics*, 63(59), 2023.

[9] S. Dereich, A. Neuenkirch, and L. Szpruch. An Euler-type method for the strong approximation of the Cox–Ingersoll–Ross process. *Proceedings of the royal society A: mathematical, physical and engineering sciences*, 468(2140):1105–1115, 2012.

[10] C. Emmanuel and X. Mao. Truncated EM numerical method for generalised Ait-Sahalia-type interest rate model with delay. *Journal of Computational and Applied Mathematics*, 383:113137, 2021.

[11] N. Halidias and I. S. Stamatiou. Boundary preserving explicit scheme for the Aït-Sahalia model. *Discrete and Continuous Dynamical Systems. Series B*, 28(1):648–664, 2023.

[12] D. J. Higham, X. Mao, and L. Szpruch. Convergence, non-negativity and stability of a new Milstein scheme with applications to finance. *Discrete and Continuous Dynamical Systems. Series B*, 18(8):2083–2100, 2013.

[13] J. Hong, L. Ji, X. Wang, and J. Zhang. Positivity-preserving symplectic methods for the stochastic Lotka–Volterra predator-prey model. *BIT Numerical Mathematics*, 62:493–520, 2022.

[14] M. Hutzenthaler, A. Jentzen, and P. E. Kloeden. Strong and weak divergence in finite time of Euler's method for stochastic differential equations with non-globally Lipschitz continuous coefficients. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 467(2130):1563–1576, 2011.

[15] Z. Lei, S. Gan, and Z. Chen. Strong and weak convergence rates of logarithmic transformed truncated EM methods for SDEs with positive solutions. *Journal of Computational and Applied Mathematics*, 419:114758, 2023.

[16] Y. Li and W. Cao. A positivity preserving Lamperti transformed Euler–Maruyama method for solving the stochastic Lotka–Volterra competition model. *Communications in Nonlinear Science and Numerical Simulation*, 122:107260, 2023.

[17] R. Liu, Y. Cao, and X. Wang. Unconditionally positivity-preserving explicit Euler-type schemes for a generalized Aït-Sahalia model. *Numerical Algorithms*, 2024.

[18] R. Liu and X. Wang. A higher order positivity preserving scheme for the strong approximations of a stochastic epidemic model. *Communications in Nonlinear Science and Numerical Simulation*, 124:107258, 2023.

[19] G. Lord and M. Wang. Convergence of a exponential tamed method for a general interest rate model. *Applied Mathematics and Computation*, 467:128503, 2024.

[20] X. Mao, F. Wei, and T. Wiriyakraikul. Positivity preserving truncated Euler–Maruyama Method for stochastic Lotka–Volterra competition model. *Journal of Computational and Applied Mathematics*, 394:113566, 2021.

[21] G. N. Milstein and M. V. Tretyakov. *Stochastic Numerics for Mathematical Physics*. Springer, 2004.

[22] A. Neuenkirch and L. Szpruch. First order strong approximations of scalar SDEs defined in a domain. *Numerische Mathematik*, 128(1):103–136, 2014.

[23] E. Platen. An introduction to numerical methods for stochastic differential equations. *Acta numerica*, 8:197–246, 1999.

[24] L. Szpruch, X. Mao, D. J. Higham, and J. Pan. Numerical simulation of a strongly nonlinear Ait-Sahalia-type interest rate model. *BIT Numerical Mathematics*, 51:405–425, 2011.

[25] X. Wang. Mean-square convergence rates of implicit Milstein type methods for SDEs with non-Lipschitz coefficients. *Advances in Computational Mathematics*, 49:37, 2023.

[26] X. Wang, J. Wu, and B. Dong. Mean-square convergence rates of stochastic theta methods for SDEs under a coupled monotonicity condition. *BIT Numerical Mathematics*, 60(3):759–790, 2020.

[27] H. Yang and J. Huang. First order strong convergence of positivity preserving logarithmic Euler–Maruyama method for the stochastic SIS epidemic model. *Applied Mathematics Letters*, 121:107451, 2021.

[28] H. Yang and J. Huang. Strong convergence and extinction of positivity preserving explicit scheme for the stochastic SIS epidemic model. *Numerical Algorithms*, 95:1475–1502, 2024.

[29] Y. Yi, Y. Hu, and J. Zhao. Positivity preserving logarithmic Euler–Maruyama type scheme for stochastic differential equations. *Communications in Nonlinear Science and Numerical Simulation*, 101:105895, 2021.