Positivity-preserving truncated Euler-Maruyama method for generalised Ait-Sahalia-type interest model

Shounian Deng · **Chen Fei** · **Weiyin Fei** · **Xuerong Mao**

Received: date / Accepted: date

Abstract The well-known Ait-Sahalia-type interest model, arising in mathematical finance, has some typical features: polynomial drift that blows up at the origin, highly nonlinear diffusion, and positive solution. The known explicit numerical methods including truncated/tamed Euler-Maruyama (EM) applied to it do not preserve its positivity. The main interest of this work is to investigate the numerical conservation of positivity of the solution of generalised Ait-Sahalia-type model. By modifying the truncated EM method to generate positive sequences of numerical approximations, we obtain the rate of convergence of the numerical algorithm not only at time T but also over the time interval $[0,T]$. Numerical experiments confirm the theoretical results.

Keywords Truncated EM method · Strong convergence order · Positivity preservation · Ait-Sahalia model

Mathematics Subject Classification (2000) 60H35 · 60H15 · 65C30

1 Introduction

Modeling the interest rate dynamics and obtaining the term structure of interest rates are important for both practical and theoretical reasons in mathematical finance [4,16]. Many well-known stochastic models have been proposed to explain the behavior of interest rates, for example Merton [21], Vasicek [25], Cox, Ingersoll and Ross [7] and Lewis [18]. Ait-Sahalia [1] investigated several continuous-time interest rate models empirically. Results of the specification test suggested that Ait-Sahalia-type (AIT, for short) model captures well of the dynamics of the spot rate. Now AIT model has been widely used to volatility and other financial quantities besides interest rate [15].

Shounian Deng · Weiyin Fei
Key Laboratory of Advanced Perception and Intelligent Control of High-end Equipment, Ministry of Education, and School of Mathematics-Physics and Finance, Anhui Polytechnic University, Wuhu 241000, China

Chen Fei (✉)
Business School, University of Shanghai for Science and Technology, Shanghai 200093, China
E-mail: chenfei@usst.edu.cn

Xuerong Mao
Department of Mathematics and Statistics, University of Strathclyde, Glasgow G1 1XH, UK

The AIT model and its generalization [1,24], which are stochastic interest rate models, has the form

$$\begin{aligned} dX(t) &= \left[a_{-1}X(t)^{-1} - a_0 + a_1X(t) - a_2X(t)^{\gamma}\right]dt + \lambda X(t)^{\theta}dB(t),\ t > 0 \\ X(0) &= X_0, \end{aligned} \tag{1.1}$$

where X_0, $a_{-1}, a_0, a_1, a_2, \lambda$ are positive constants, $\gamma, \theta > 1$, $B(\cdot)$ is a scalar Brownian motion. Under certain conditions on the parameters, this stochastic differential equation (SDE) has a unique, positive global solution, see [24, Theorem 2.1]. Note that the AIT model of the form (1.1) has a superlinear diffusion and a polynomial drift containing a term $a_{-1}X(t)^{-1}$ which may explode to infinity in finite time at the origin. That is the main source of mathematical and numerical difficulties in the convergence analysis.

In the context of the AIT model, strong L^p-convergence (without rates) of backward Euler-Maruyama (BEM) scheme applied to (1.1) was obtained in [24]. Subsequently, strong convergence rates for the AIT model with Poisson noise were established in [28]. Recently, the following truncated Euler-Maruyama (TEM) scheme was proposed in [12] for this model:

$$Z_{k+1} = Z_k + f(\pi_\Delta(Z_k))\Delta + g(\pi_\Delta(Z_k))\Delta B_k, \qquad Z_0 = X_0, \tag{1.2}$$

where f, g are defined in (2.1) and π_Δ denotes a truncated mapping. The authors proved that the TEM approximation Z_k converges strongly to the true solution $X(t)$ of the AIT model (1.1) without revealing a rate of convergence. Nevertheless, we note that the numerical sequences $\{Z_k\}_{k\geq 0}$ provide negative approximations with positive probability for certain step size Δ.

We then turn our attention to the positivity-preserving schemes for SDEs including the AIT model with nonglobally Lipschitz coefficients. In the classical Euler-Maruyama (EM) scheme [6,8], the approximation can potentially escape the domain of the true solution of the SDE. In contrast, implicit approximations such as BEM [11,22], drift implicit Milstein [14] and stochastic θ [26,23], stay in restricted domains and thus maintain the property of positiveness for approximate solutions. An overview of positive and implicit numerical schemes in mathematical finance can be found in [13]. However, in order to obtain the implicit approximations an additional nonlinear equation needs to be solved at each iteration, which is time consuming to some extent.

In recent years, much effort has focused on deriving the combination of the Lamperti transformation and the Euler schemes for SDEs with positive solutons. Several modifications have been introduced such as Lamperti transformation plus projected EM [4] and logarithmic transformation plus truncated EM [27,17] schemes for financial SDEs including the CIR, the AIT, and the 3/2 models, Lamperti smooth sloping truncation scheme for Wright-Fisher model [5]. However, Lamperti transformation: $Y = X^{1-\theta}$ would translate all the nonsmoothness into the drift. Logarithmic transformation: $Y = \log X$ would cause the coefficients of the transformed SDE growing exponentially. In contrast to the BEM [24, Theorem 6.1], logarithmic truncated EM needs a more restrictive condition: $\gamma > 4\theta - 3$ on the parameters to meet the requirement of Khasminskii-type assumption for the transformed SDE, see [27, Example 4.14].

Based on the above discussions, this paper focuses on the strong convergence rate of a non-transformation and positivity-preserving scheme for the AIT model (1.1). Recently, Mao, Wei and Wiriyakraikul [20] proposed a new positivity-preserving truncated EM scheme for the stochastic Lotka-Volterra model for interacting multi-species in ecology. The authors adopted the truncated method to control the superlinearity of the coefficient of the model and constructed a positive mapping to ensure the numerical sequences to take values in a positive interval. They established strong L^p-convergence (without rates) for their scheme by verifying that the non-truncated-positive sequences and truncated-positive sequences are sufficiently close in the L^p-norm, see [20, Theorem 4.2]. To ensure the property of positiveness for approximate solutions, we follow the similar idea in [20] to construct a variant of truncated EM schemes which is also similar to that in [4] and [12]. Essentially, both the truncated and projected EM methods prevent the numerical approximations from leaving the truncated domain, see e.g., [2,4,19]. We note that the reasons for the failure to obtain the strong convergence rate in [12] lie in two aspects. First, only a Δ-ε type estimate bound for the probability of truncated EM approximation escaping from the truncated interval $(1/R, R)$, i.e., $\mathbb{P}(\rho_{\Delta,R} \leq T) \leq \varepsilon$, was provided, see [12, Lemma 4.4]. Second, the monotonicity condition (3.2) satisfied in the AIT model (1.1) was not fully utilized. To overcome the above two issues essentially caused by the singularity at zero and the superlinearity at the infinity, we construct a truncated projection π_Δ onto a interval $[1/R, R]$ in $\mathbb{R}$, whose size R is expanding with a negative power of the step size, see (2.6) and (3.6). Then we apply the method of Lyapunov function to establish an upper bound for $\mathbb{P}(\rho_{\Delta,R} \leq T)$ which can be expressed as an explicit power function of Δ and R, see Lemma 3.2. Moreover, we apply the monotonicity condition (3.2) to obtain the order of the L^p-error in the truncated domain, see Lemma 3.3. In a similar fashion as in the proofs of [9, Theorem 4.9] or [10, Theorem 4.1], combining Lemma 3.2 with Lemma 3.3 allows us to establish a convergence order of the truncated EM method for the AIT model (1.1), see Theorems 3.1 and 3.2.

The contributions of this paper can be summarised as follows.

- *Positivity preservation and convergence order.* Compared with the convergence results of truncated EM in [12], our scheme maintains the property of positiveness. Furthermore, we successfully obtain the strong convergence order of truncated EM for AIT model.
- *Assumptions*. Compared with the logarithmic transformation EM scheme [27] for the AIT model, our non-transformation EM requires a slightly weaker condition on the parameters γ and θ, namely $\gamma > 2\theta - 1$, while [27, Example 4.14] requires $\gamma > 4\theta - 3$.

The rest of the paper is organised as follows: In Section 2, we introduce a class of truncated EM scheme which can remain positive for the approximation solution and study analytical properties of the truncated mapping. Strong rates of convergence are investigated in Sections 3. In Section 4, we perform some numerical experiments to support the theoretical results. In the final section, we close the paper by our conclusion.

2 Positivity-preserving truncated EM scheme

Throughout this paper, C will be used to denote generic positive constants, which may be different from line to line. Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a complete probability space with a filtration $\{\mathscr{F}_t\}_{t\geq 0}$ satisfying the usual conditions (i.e., it is increasing and right continuous while $\mathscr{F}_0$ contains all $\mathbb{P}$-null sets). If $X \in \mathbb{R}$, then $|X|$ is the Euclidean norm. For two real numbers a and b, $a \vee b := \max(a,b)$ and $a \wedge b := \min(a,b)$. Let $\mathbb{R} = (-\infty, \infty)$ and $\mathbb{R}_+ = (0, \infty)$. For a set G, its indicator function is denoted by $\mathbf{1}_G$. For an empty $\emptyset$, we set $\inf \emptyset = \infty$. Define the coefficients of (1.1):

$$f(x) = a_{-1}x^{-1} - a_0 + a_1 x - a_2 x^\gamma \quad \text{and} \quad g(x) = \lambda x^\theta, \quad \text{for all } x > 0. \tag{2.1}$$

For $H \in C^2(\mathbb{R}_+; \mathbb{R})$, define $\mathbb{L}H : \mathbb{R}_+ \to \mathbb{R}$ by

$$\mathbb{L}H(x) = H'(x)f(x) + \frac{1}{2}H''(x)|g(x)|^2, \tag{2.2}$$

where $H'(x) = \dfrac{dH(x)}{dx}$, $H''(x) = \dfrac{d^2H(x)}{dx^2}$, f and g are from (2.1).

In order to proceed with our analysis, we need an assumption on the parameters. The following assumption allow us to control the superlinear growth from the diffusion term using the dissipative nature of the drift, see [24, p. 408].

Assumption 2.1 *The parameters of SDE (1.1) satisfy*

$$1 + \gamma > 2\theta.$$

Lemmas 2.1 and 2.2 give boundedness properties for the solution to SDE (1.1). While Lemmas 2.3 and 2.4 show that the coefficients of SDE (1.1) satisfy the local Lipschitz and Khasminskii-type conditions on $\mathbb{R}_+$. The proofs of Lemmas 2.1-2.4 can be found in [12, Lemmas 2.4, 2.5, 3.2, 3.3].

Lemma 2.1 *Let Assumption 2.1 hold. Then for any $r \geq 2$, the solution $X(t)$ to SDE (1.1) is upper bounded, i.e.,*

$$\sup_{0\leq t<\infty} \mathbb{E}|X(t)|^r \leq C_1,$$

and

$$\sup_{0\leq t<\infty} \mathbb{E}\left|\frac{1}{X(t)}\right|^r \leq C_2,$$

where C_1 and C_2 are positive constants.

Lemma 2.2 *Let Assumption 2.1 hold. Then for any $r \geq 2$, the solution $X(t)$ to SDE (1.1) satisfies*

$$\mathbb{E}\left[\sup_{0\leq t\leq T} |X(t)|^r\right] \leq C_3,$$

where C_3 is a positive constant.

Lemma 2.3 *For any $R > 0$, there exists a positive constant $L_R > 0$ such that f and g coefficients terms of SDE (1.1) satisfies*

$$|f(x) - f(y)| \vee |g(x) - g(y)| \leq L_R|x-y|, \quad \textit{for all } x, y \in [1/R, R].$$

Lemma 2.4 *Let Assumption 2.1 hold. Then for any $r \geq 2$, there exists K_1 such that the coefficients of SDE (1.1) satisfies*

$$xf(x) + \frac{r-1}{2}|g(x)|^2 \leq K_1(1+|x|^2), \quad \textit{for all } x \in \mathbb{R}_+,$$

where $K_1 = [a_{-1} + K(r)] \vee a_1$ and $K(r) \geq -a_2 x^{\gamma+1} + \dfrac{p-1}{2}\lambda^2 x^{2\theta}$.

To define the truncated EM scheme, we choose a strictly increasing continuous functions $\mu : \mathbb{R}_+ \to \mathbb{R}_+$ such that $\mu(R) \to \infty$ as $R \to \infty$ and

$$\sup_{1/R\leq x\leq R} |f(x)| \vee g(x) \leq \mu(R), \quad \text{for all } R \geq 1. \tag{2.3}$$

Denote by μ^{-1} the inverse function of μ and we see that $\mu^{-1} : [\mu(1), \infty) \to \mathbb{R}_+$ is a strictly increasing continuous function. We then choose a positive constant $\hat{L}$ and a strictly decreasing function $\varphi : (0,1] \to [\mu(1), +\infty)$ such that for all $\Delta \in (0,1]$,

$$\lim_{\Delta\to 0} \varphi(\Delta) = \infty, \quad \varphi(\Delta) \leq \hat{L}\Delta^{-1/4}, \quad \text{and} \quad \mu^{-1}(\varphi(1)) \geq K_0, \tag{2.4}$$

where

$$K_0 := \frac{1}{X_0} \vee X_0 \vee \frac{a_0}{a_{-1}} \vee \left|\frac{a_2}{a_1}\right|^{\frac{1}{\gamma-1}}. \tag{2.5}$$

Let $\Delta \in (0,1]$, define a truncation mapping $\pi_\Delta : \mathbb{R} \to \left[\frac{1}{\mu^{-1}(\varphi(\Delta))}, \mu^{-1}(\varphi(\Delta))\right]$ by

$$\pi_\Delta(x) = \frac{1}{\mu^{-1}(\varphi(\Delta))} \vee (x \wedge \mu^{-1}(\varphi(\Delta))), \quad \text{for all } x \in \mathbb{R}. \tag{2.6}$$

Consequently,

$$\pi_\Delta(x) = \begin{cases} \frac{1}{\mu^{-1}(\varphi(\Delta))} & \text{if } x < \frac{1}{\mu^{-1}(\varphi(\Delta))}, \\ x & \text{if } \frac{1}{\mu^{-1}(\varphi(\Delta))} \le x \le \mu^{-1}(\varphi(\Delta)), \\ \mu^{-1}(\varphi(\Delta)) & \text{if } x > \mu^{-1}(\varphi(\Delta)). \end{cases}$$

It is useful to note that for any $x < \frac{1}{\mu^{-1}(\varphi(\Delta))}$, $|\pi_\Delta(x)| = \frac{1}{\mu^{-1}(\varphi(\Delta))} < 1$ and for any $x \ge \frac{1}{\mu^{-1}(\varphi(\Delta))}$, $|\pi_\Delta(x)| \le |x|$. Thus, we have

$$|\pi_\Delta(x)| = \pi_\Delta(x) \le \max(1, |x|), \quad \text{for all } x \in \mathbb{R}, \tag{2.7}$$

which means that for any $r > 0$,

$$|\pi_\Delta(x)|^r \le 1 + |x|^r, \quad \text{for all } x \in \mathbb{R}. \tag{2.8}$$

Define the truncated functions

$$f_\Delta(x) = f(\pi_\Delta(x)) \quad \text{and} \quad g_\Delta(x) = g(\pi_\Delta(x)), \quad \text{for all } x \in \mathbb{R}. \tag{2.9}$$

We observe from (2.3) that

$$|f_\Delta(x)| \vee g_\Delta(x) \le \varphi(\Delta), \quad \text{for all } x \in \mathbb{R}. \tag{2.10}$$

Let $\Delta = T/N$ for some $N \in \mathbb{N}$, define the following truncated EM scheme by setting $X_\Delta(t_0) = X_0$ and computing

$$\begin{aligned} Y_\Delta(t_k) &= \pi_\Delta(X_\Delta(t_k)), \\ X_\Delta(t_{k+1}) &= X_\Delta(t_k) + f(Y_\Delta(t_k)) + g(Y_\Delta(t_k))\Delta B_k, \end{aligned} \tag{2.11}$$

for $k = 0, 1, 2, \cdots$, where $\Delta B_k = B(t_{k+1}) - B(t_k)$, $t_k = k\Delta$ and π_Δ has been defined in (2.6). By (2.9), (2.11) can be rewritten as the following form:

$$X_\Delta(t_{k+1}) = X_\Delta(t_k) + f_\Delta(X_\Delta(t_k)) + g_\Delta(X_\Delta(t_k))\Delta B_k, \quad k \ge 0. \tag{2.12}$$

We observe that discrete-time auxiliary sequences $\{X_\Delta(t_k)\}_{k=0}^N$ do not preserve positivity, however sequences $\{Y_\Delta(t_k)\}_{k=0}^N$ do. Define the following continuous-time step-processes $\bar{x}_\Delta(t)$ and $\bar{y}_\Delta(t)$ on $t \ge 0$ by

$$\bar{x}_\Delta(t) = \sum_{k=0}^{\infty} X_\Delta(t_k) \mathbf{1}_{[t_k, t_{k+1})}(t) \quad \text{and} \quad \bar{y}_\Delta(t) = \sum_{k=0}^{\infty} Y_\Delta(t_k) \mathbf{1}_{[t_k, t_{k+1})}(t). \tag{2.13}$$

Define the following continuous-time continuous process $x_\Delta(t)$ on $t \ge 0$ by

$$x_\Delta(t) = X_0 + \int_0^t f_\Delta(\bar{x}_\Delta(s))ds + \int_0^t g_\Delta(\bar{x}_\Delta(s))dB(s). \tag{2.14}$$

Clearly, for any $t \in [t_k, t_{k+1})$ with $k \ge 0$, we have $x_\Delta(t_k) = \bar{x}_\Delta(t_k) = X_\Delta(t_k)$, which means that the continuous and discrete truncated EM schemes coincide at the grid points, and

$$x_\Delta(t) - \bar{x}_\Delta(t) = f_\Delta(X_\Delta(t_k))(t - t_k) + g_\Delta(X_\Delta(t_k))(B(t) - B(t_k)). \tag{2.15}$$

The following lemma illustrates that the truncated functions f_Δ and g_Δ preserve the Khasminskii-type condition nicely.

Lemma 2.5 *Let Assumptions 2.1 hold. Then for any $r \geq 2$, the truncated functions satisfy*

$$x f_\Delta(x) + \frac{r-1}{2}|g_\Delta(x)|^2 \leq \overline{K}(1+|x|^2), \quad \textit{for all } x \in \mathbb{R},\ \Delta \in (0,1],$$

where $\overline{K} = K_1 \vee K_2$ is a positive constant independent of Δ.

Proof For $x \in \mathbb{R}$ with $x \in (0, 1/\mu^{-1}(\varphi(\Delta))$, we have $0 < x\mu^{-1}(\varphi(\Delta)) < 1$. By Lemma 2.4, we get that

$$\begin{aligned} x f_\Delta(x) &= x f(1/\mu^{-1}(\varphi(\Delta))) = x\mu^{-1}(\varphi(\Delta))\frac{1}{\mu^{-1}(\varphi(\Delta))} f(1/\mu^{-1}(\varphi(\Delta))) \\ &\leq x\mu^{-1}(\varphi(\Delta))K_1(1+1/|\mu^{-1}(\varphi(\Delta))|^2) \leq 2K_1. \end{aligned}$$

Consequently,

$$x f_\Delta(x) + \frac{r-1}{2}|g_\Delta(x)|^2 \leq 2K_1 + \frac{r-1}{2}|g(1/\mu^{-1}(\varphi(\Delta)))|^2 \leq 2K_1 + \frac{r-1}{2}\lambda^2 \leq K_2(1+|x|^2),$$

where $K_2 = 2K_1 + \frac{r-1}{2}\lambda^2$ and K_1 is from Lemma 2.4. By (2.4), we have

$$f\left(\frac{1}{\mu^{-1}(\varphi(1))}\right) = a_{-1}\mu^{-1}(\varphi(1)) - a_0 + \frac{a_1}{\mu^{-1}(\varphi(1))} - \frac{a_2}{(\mu^{-1}(\varphi(1)))^\gamma} \geq 0. \tag{2.16}$$

For all $x \leq 0$, we conclude from (2.16) that

$$f_\Delta(x) = f(1/\mu^{-1}(\varphi(\Delta))) \geq 0.$$

Moreover,

$$g_\Delta(x) = g\left(\frac{1}{\mu^{-1}(\varphi(\Delta))}\right) = \lambda\left(\frac{1}{\mu^{-1}(\varphi(\Delta))}\right)^\theta < \lambda.$$

Then,

$$x f_\Delta(x) + \frac{r-1}{2}|g_\Delta(x)|^2 \leq \frac{r-1}{2}\lambda^2 \leq K_2 \leq K_2(1+|x|^2).$$

For $x \in [1/\mu^{-1}(\varphi(\Delta)), \mu^{-1}(\varphi(\Delta))]$, we have

$$x f_\Delta(x) + \frac{r-1}{2}|g_\Delta(x)|^2 = x f(x) + \frac{r-1}{2}|g(x)|^2 \leq K_1(1+|x|^2).$$

For $x > \mu^{-1}(\varphi(\Delta))$, we also have $x f_\Delta(x) + \frac{r-1}{2}|g_\Delta(x)|^2 \leq 2K_1(1+|x|^2)$, see [12, Lemma 3.4]. Thus, the proof is finished. □

A central step in establishing convergence of approximation process is to prove uniform bound on the moments of the numerical solution. Similar to the proof of [12, Lemma 4.2], we can show that the following Lemma 2.6 also holds.

Lemma 2.6 *Let Assumption 2.1 hold. Then for any $r \geq 2$,*

$$\sup_{0<\Delta\leq 1}\sup_{0\leq t\leq T} \mathbb{E}|x_\Delta(t)|^r \leq C, \quad \textit{for all } T > 0.$$

where C is a positive constant independent of Δ.

From Lemma 2.6, we have $\sup_{0\le t\le T}\mathbb{E}|\bar{x}_\Delta(t)|^r \le C$. Together this with (2.8) implies that

$$|\bar{y}_\Delta(t)|^r = |\pi_\Delta(\bar{x}_\Delta(t))|^r \le 1 + |\bar{x}_\Delta(t)|^r. \tag{2.17}$$

The next Lemma 2.7 follows directly from Lemma 2.6 and (2.17).

Lemma 2.7 *Let Assumption 2.1 hold. Then for any* $r \ge 2$*,*

$$\sup_{0<\Delta\le 1}\sup_{0\le t\le T}\mathbb{E}|\bar{y}_\Delta(t)|^r \le C, \quad \textit{for all } T>0,$$

where C is a positive constant independent of Δ*.*

The next lemma is a consequence of Lemma 2.7.

Lemma 2.8 *Let Assumption 2.1 hold. Then for any* $r \ge 2$*,*

$$\sup_{0<\Delta\le 1}\mathbb{E}\left[\sup_{0\le t\le T}|x_\Delta(t)|^r\right] \le C, \quad \textit{for all } T>0,$$

where C is a positive constant independent of Δ*.*

Proof By (2.14), we have

$$a_{-1}\int_0^T \frac{1}{\bar{y}_\Delta(s)}ds = x_\Delta(T) - X_0 + a_0 T - a_1\int_0^T \bar{y}_\Delta(s)ds + a_2\int_0^T \bar{y}_\Delta(s)^\gamma ds - \lambda\int_0^T \bar{y}_\Delta(s)^\theta dB(s).$$

Combining this and Lemma 2.7, we have

$$\mathbb{E}\left|\int_0^t \frac{1}{\bar{y}_\Delta(s)}ds\right|^r < \infty. \tag{2.18}$$

It follows from (2.14) and the fact that $f_\Delta(\bar{x}_\Delta(t)) = f(\bar{y}_\Delta(t))$, we have, for all $r\ge 2$,

$$\begin{aligned}\mathbb{E}\left[\sup_{0\le t\le T}|x_\Delta(t)|^r\right] \le& 3^{r-1}X_0^r + 3^{r-1}\mathbb{E}\left|\int_0^T\left(a_1\frac{1}{\bar{y}_\Delta(s)} + a_0 + a_1\bar{y}_\Delta(s) + a_2\bar{y}_\Delta(s)^\gamma\right)ds\right|^r \\ &+ 3^{r-1}\mathbb{E}\left[\sup_{0\le t\le T}\left|\int_0^t \lambda\bar{y}_\Delta(s)^\theta dB(s)\right|^r\right].\end{aligned}$$

This, together with Lemma 2.7 and (2.18) as well as the Burkholder-Davis-Gundy inequality, implies the assertion. □

Lemma 2.9 *Let Assumption 2.1 hold. Then for any* $r \ge 2$*,*

$$\lim_{\Delta\to 0}\mathbb{E}\left[\sup_{0\le t\le T}|x_\Delta(t) - \bar{x}_\Delta(t)|^r\right] = 0, \quad \textit{for all } T>0.$$

Proof By (2.15) and (2.10), we have

$$\begin{aligned}\mathbb{E}\left[\sup_{0\le t\le T}|x_\Delta(t)-\bar{x}_\Delta(t)|^r\right] &\le \mathbb{E}\left[\max_{0\le k\le N}\sup_{t_k\le t<t_{k+1}}|f_\Delta(X_\Delta(t_k))(t-t_k) + g_\Delta(X_\Delta(t_k))(B(t)-B(t_k))|^r\right] \\ &\le C(\varphi(\Delta))^r(\Delta^r + J_1),\end{aligned} \tag{2.19}$$

where

$$J_1 = \mathbb{E}\left[\max_{0\le k\le N}\sup_{t_k\le t<t_{k+1}}|B(t)-B(t_k)|^r\right].$$

According to [19, Lemma 4.5, p. 380], we can show that

$$J_1 \le n^{r/2}\left(\frac{2n}{2n-1}\right)^r (T+1)^{\frac{r}{2n}}\Delta^{\frac{r(n-1)}{2n}}. \tag{2.20}$$

If we take positive integer n to sufficiently large satisfying $\left(\frac{2n}{2n-1}\right)^r (T+1)^{\frac{r}{2n}} \le 2$ and $\frac{n-1}{2n} > \frac{1}{3}$, then $J_1 \le C\Delta^{1/3}$. Inserting this into (2.19), we have

$$\mathbb{E}\left[\sup_{0\le t\le T}|x_\Delta(t)-\bar{x}_\Delta(t)|^r\right] \le C(\varphi(\Delta))^r\Delta^{r/3}. \tag{2.21}$$

Thus, we complete the proof. □

The following lemma shows that the true solution $X(t)$ will remain in a truncated interval $(1/R,R)$ with a large probability.

Lemma 2.10 *Assume that Assumption 2.1 holds. For any real number $R \ge K_0$, define the stopping time*

$$\tau_R = \inf\{t\ge 0 : X(t) \notin (1/R,R)\}, \tag{2.22}$$

Then for any $r \ge 2$,

$$\mathbb{P}(\tau_R \le T) \le \frac{C}{R^r}, \quad \textit{for all } T>0. \tag{2.23}$$

Proof For any $r \ge 2$, we set

$$\hat{H}(x) = x^r + x^{-r}, \quad \text{for all } x>0.$$

Thus,

$$\begin{aligned}\mathbb{L}\hat{H}(x) &= (rx^{r-1} - rx^{-r-1})(a_{-1}x^{-1} - a_0 + a_1 x - a_2 x^\gamma) + 0.5\lambda^2[r(r-1)x^{r-2} + r(r+1)x^{-r-2}]x^{2\theta} \\ &= -a_{-1}rx^{-r-2} + 0.5\lambda^2 r(r+1)x^{-r-2+2\theta} + \cdots - a_2 r x^{r-1+\gamma} + 0.5\lambda^2 r(r-1)x^{r-2+2\theta}.\end{aligned}$$

By Assumption 2.1, there is a positive constant $\hat{K}_3$ such that $\mathbb{L}\hat{H}(x) \le \hat{K}_3$. By the Itô formula, we have that

$$\mathbb{E}[\hat{H}(X(t\wedge\tau_R))] = \hat{H}(X_0) + \mathbb{E}\int_0^{t\wedge\tau_R} L\hat{H}(X(s))ds \le \hat{H}(X_0) + \hat{K}_3 T.$$

Then,

$$R^p\mathbb{P}(\tau_R\le T) \le [\hat{H}(1/R)\wedge\hat{H}(R)]\mathbb{P}(\tau_R\le T) \le \mathbb{E}[\hat{H}(X(\tau_R\wedge T))] \le \hat{H}(X_0) + \hat{K}_3 T,$$

which implies that (2.23) holds. □

Remark 2.1 By the above Lemmas, we can show that for any $r \ge 2$,

$$\lim_{\Delta\to 0}\mathbb{E}\left[\sup_{0\le t\le T}|X(t)-\bar{x}_\Delta(t)|^r\right] = 0, \tag{2.24}$$

see, e.g., [19, Theorem 4.4]. On the other hand, in a similar fashion as [20, Theorem 4.2] was proved, we also can show that

$$\lim_{\Delta\to 0}\mathbb{E}\left[\sup_{0\le t\le T}|\bar{x}_\Delta(t)-\bar{y}_\Delta(t)|^r\right] = 0. \tag{2.25}$$

Thus, by the triangle inequality, we have

$$\lim_{\Delta\to 0}\mathbb{E}\left[\sup_{0\le t\le T}|X(t)-\bar{y}_\Delta(t)|^r\right]=0, \tag{2.26}$$

which establishes the strong convergence of positivity-preserving truncated EM scheme for the AIT model. However, (2.26) does not show a convergence rate. In this paper, we develop new arguments to establish a rate of L^p-convergence of TEM not only at time T but also over the time interval, see Theorems 3.1 and 3.2.

3 Convergence order

In this section, we aim to establish the convergence order of the truncated EM scheme. Let $p\ge 2$ and $1+\gamma>2\theta$. We note that

$$\sup_{x>0}\left(f'(x)+\frac{p-1}{2}|g'(x)|^2\right)=\sup_{x>0}\left(-a_{-1}x^{-2}+a_1-a_2\gamma x^{\gamma-1}+\frac{p-1}{2}\lambda^2\theta^2x^{2\theta-2}\right)<\infty, \tag{3.1}$$

which means that the monotonicity condition is fulfilled in $\mathbb{R}_+$ for all positive parameters, i.e.,

$$\langle x-y,f(x)-f(y)\rangle+\frac{p-1}{2}|g(x)-g(y)|^2\le L^\star|x-y|^2,\quad \text{for all } x,y\in\mathbb{R}_+, \tag{3.2}$$

where $L^\star$ is a positive constant, see [28, Theorem 1] and [26, Theorem 4.5]. Define

$$f_2(x)=-a_0+a_1x-a_2x^\gamma,\quad \text{for all } x\in\mathbb{R}_+.$$

Then

$$f(x)=a_{-1}\frac{1}{x}+f_2(x),\quad \text{for all } x\in\mathbb{R}_+,$$

and

$$|f_2(x)-f_2(y)|\vee|g(x)-g(y)|\le C(1+|x|^{\gamma-1}+|y|^{\gamma-1})|x-y|,\quad \text{for all } x,y\in\mathbb{R}_+. \tag{3.3}$$

In the remaining part of this paper, we provide the reader with a specific procedure to construct a truncated mapping. Now, by Lemma 2.3, (2.3)-(2.5), we set

$$\mu(R)=K_{11}R^\gamma,\quad L_R=K_{12}R^{(\gamma-1)\vee 2},\quad \text{for all } R\in[1,\infty), \tag{3.4}$$

and

$$\varphi(\Delta)=\mu(\hat{K}_0)\Delta^{-\alpha},\quad \text{for all } \Delta\in(0,1], \tag{3.5}$$

where $K_{11}=(a_{-1}+a_0+a_1+a_2)\vee\lambda$, $K_{12}=(a_{-1}+a_1+a_2\gamma)\vee\lambda\theta$, $\hat{K}_0\in\left[\frac{1}{X_0}\vee X_0\vee\frac{a_0}{a_{-1}}\vee\left|\frac{a_2}{a_1}\right|^{\frac{1}{\gamma-1}},\infty\right)$ is a constant, and $\alpha\in(0,1/4]$ is a constant to be determined later on. Thus, we have

$$\mu^{-1}(\varphi(\Delta))=\hat{K}_0\Delta^{-\frac{\alpha}{\gamma}},\quad \text{for all } \Delta\in(0,1]. \tag{3.6}$$

Once μ, L_R and φ have been defined by (3.4) and (3.5) explicitly, then the upper estimate bound for $\mathbb{P}(\rho_{\Delta,R}\le T)$ can be written as a negative power function of Δ, as Lemma 3.2 and (3.34) clarify. Together this with Lemma 3.3, the convergence order can be derived in a similar fashion as in the proofs of [9, Theorem 4.9] or [10, Theorem 4.1].

3.1 Convergence order at time T

The following lemma shows that x_Δ and $\bar{x}_\Delta$ are close to each other in the sense of qth moment.

Lemma 3.1 *Let Assumption 2.1 hold. Then for any $q>0$, we have*

$$\mathbb{E}|x_\Delta(t)-\bar{x}_\Delta(t)|^q \le \hat{c}_p\Big(\Delta^q(\mu^{-1}(\varphi(\Delta)))^q+\Delta^{q/2}\Big), \quad \textit{for all } t\in(0,\infty),\ \Delta\in(0,1], \tag{3.7}$$

where $\hat{c}_q$ is a positive constant dependent of q but independent of Δ.

Proof We conclude from (2.8) that

$$\pi_\Delta(x)\le 1+|x|, \quad \text{for all } x\in\mathbb{R}. \tag{3.8}$$

Consequently,

$$\begin{aligned}|f_\Delta(x)|=|f(\pi_\Delta(x))| &\le a_{-1}\frac{1}{\pi_\Delta(x)}+a_0+a_1\pi_\Delta(x)+a_2|\pi_\Delta(x)|^\gamma\\ &\le a_{-1}\mu^{-1}(\varphi(\Delta))+a_0+a_1(1+|x|)+a_2(1+|x|)^\gamma\\ &\le K_5\Big(\mu^{-1}(\varphi(\Delta))+|x|^\gamma\Big), \quad \text{for all } \in\mathbb{R},\end{aligned} \tag{3.9}$$

where $K_5=2^\gamma(a_{-1}+a_0+a_1+a_2)$ and

$$|g_\Delta(x)|=|g(\pi_\Delta(x))|=\lambda|\pi_\Delta(x)|^\theta\le\lambda(1+|x|)^\theta, \quad \text{for all } x\in\mathbb{R}. \tag{3.10}$$

Let $t\in[t_k,t_{k+1})$. For any $q\ge 2$, by (3.9) and (3.10) as well as Lemma 2.6, we conclude from (2.15) that

$$\begin{aligned}\mathbb{E}|x_\Delta(x)-\bar{x}_\Delta(t)|^q &\le 2^{q-1}\Big(\mathbb{E}|f_\Delta(X_\Delta(t_k)(t-t_k))|^q+\mathbb{E}|g_\Delta(X_\Delta(t_k)(B(t)-B(t_k))|^q\Big)\\ &\le \hat{c}_q\Big(\Delta^q[(\mu^{-1}(\varphi(\Delta)))^q+\mathbb{E}|X_\Delta(t_k)|^{\gamma q}]+\Delta^{q/2}\mathbb{E}[1+|X_\Delta(t_k)|^{\theta q}]\Big)\\ &\le \hat{c}_q\Big(\Delta^q(\mu^{-1}(\varphi(\Delta)))^q+\Delta^{q/2}\Big),\end{aligned} \tag{3.11}$$

which gives (3.7). For $0<q<2$, by the Hölder inequality, we can show that (3.7) also holds. □

Lemma 3.2 *Assume that Assumption 2.1 holds. For any real number $R\ge K_0$, define the stopping time*

$$\rho_{\Delta,R}=\inf\{t\in[0,T]: x_\Delta(t)\notin(1/R,R)\}. \tag{3.12}$$

Then

$$\mathbb{P}(\rho_{\Delta,R}\le T)\le CR^{(2\gamma+1)\vee(\gamma+4)}\Delta^{1/2}, \quad \textit{for all } \Delta\in(0,1], \tag{3.13}$$

where C is a positive constant independent of Δ and R.

Proof Define a C^2-function $H:\mathbb{R}_+\to\mathbb{R}_+$ by

$$H(x)=\frac{1}{x^2}+x^2. \tag{3.14}$$

Clearly, $H(x)\to\infty$ as $x\to\infty$ or $x\to 0$. For any $t\in[0,T]$, we derive from the Itô formula that

$$\begin{aligned}\mathbb{E}[H(x_\Delta(t\wedge\rho_{\Delta,R}))] &= H(X_0)+\mathbb{E}\int_0^{t\wedge\rho_{\Delta,R}}\Big(H'(x_\Delta(s))f_\Delta(\bar{x}_\Delta(s))+\frac{1}{2}H''(x_\Delta(s))|g_\Delta(\bar{x}_\Delta(s))|^2\Big)ds\\ &\le H(X_0)+\mathbb{E}\int_0^{t\wedge\rho_{\Delta,R}}\Big(\mathbb{L}H(x_\Delta(s))+I_2(s)\Big)ds,\end{aligned}$$

where

$$I_2(s) = H'(x_\Delta(s))[f_\Delta(\bar{x}_\Delta(s)) - f_\Delta(x_\Delta(s))] + \frac{1}{2}H''(x_\Delta(s))[|g_\Delta(\bar{x}_\Delta(s))|^2 - |g_\Delta(x_\Delta(s))|^2].$$

By the definition of the truncated functions, we have that for any $s \in [0, t \wedge \rho_{\Delta,R}]$ with $\mu^{-1}(\varphi(\Delta)) \geq R$

$$f_\Delta(\bar{x}_\Delta(s)) = f(\bar{x}_\Delta(s)) \quad \text{and} \quad g_\Delta(\bar{x}_\Delta(s)) = g(\bar{x}_\Delta(s)),$$

as well as

$$|f(\bar{x}_\Delta(s)) - f(x_\Delta(s))| \vee |g(\bar{x}_\Delta(s)) - g(x_\Delta(s))| \leq K_{12}R^{(\gamma-1)\vee 2}|\bar{x}_\Delta(s) - x_\Delta(s)|,$$

moreover

$$\begin{aligned} ||g_\Delta(\bar{x}_\Delta(s))|^2 - |g_\Delta(x_\Delta(s))|^2| &= |g(\bar{x}_\Delta(s)) + g(x_\Delta(s))||g(\bar{x}_\Delta(s)) - g(x_\Delta(s))| \\ &\leq 2\mu(R)K_{12}R^{(\gamma-1)\vee 2}|\bar{x}_\Delta(s) - x_\Delta(s)|. \end{aligned}$$

Thus,

$$I_2(s) \leq A_R|\bar{x}_\Delta(s) - x_\Delta(s)|,$$

where

$$A_R = \max_{1/R \leq x \leq R}[K_{12}R^{(\gamma-1)\vee 2}|H'(x) + H''(x)\mu(R)|],$$

with

$$H'(x) = 2x - 2x^{-3} \quad \text{and} \quad H''(x) = 2 + 6x^{-4}. \tag{3.15}$$

By Assumption 2.1, there is a constant K_3 such that

$$\mathbb{L}H(x_\Delta(s)) = H'(x_\Delta(s))f(x_\Delta(s)) + \frac{1}{2}H''(x_\Delta(s))|g(x_\Delta(s))|^2 \leq K_3.$$

Thus, we get that

$$\begin{aligned} \mathbb{E}[H(x_\Delta(t \wedge \rho_{\Delta,R}))] &\leq H(X_0) + K_3T + A_R\mathbb{E}\int_0^{t \wedge \rho_{\Delta,R}} |\bar{x}_\Delta(s) - x_\Delta(s)|ds \\ &\leq H(X_0) + K_3T + A_R\int_0^T \mathbb{E}|\bar{x}_\Delta(s) - x_\Delta(s)|ds. \end{aligned} \tag{3.16}$$

From Lemma 3.1, we have

$$\mathbb{E}|\bar{x}_\Delta(s) - x_\Delta(s)| \leq C\Big(\Delta^{1/2} + \Delta\mu^{-1}(\varphi(\Delta))\Big) \leq C\Delta^{1/2}. \tag{3.17}$$

Combining this and (3.16), we have

$$\mathbb{P}(\rho_{\Delta,R} \leq T) \leq \frac{H(X_0) + K_3T + C\Delta^{1/2}A_R}{H(1/R) \wedge H(R)}. \tag{3.18}$$

For any $1/R \leq x \leq R$, we conclude from (3.15) that $H'(x) \leq CR^3$ and $H''(x) \leq CR^4$. Recalling (3.4), we have

$$\begin{aligned} A_R &= \max_{1/R \leq x \leq R}[L_RH'(x) + H''(x)\mu(R)L_R] \\ &\leq C\Big(R^{(\gamma-1)\vee 2}R^3 + R^{(\gamma-1)\vee 2}R^\gamma R^4\Big) \leq CR^{(2\gamma+3)\vee(\gamma+6)}. \end{aligned} \tag{3.19}$$

Thus, we conclude from (3.18) that

$$\mathbb{P}(\rho_{\Delta,R} \le T) \le C\frac{R^{(2\gamma+3)\vee(\gamma+6)}}{R^2}\Delta^{1/2} \le CR^{(2\gamma+1)\vee(\gamma+4)}\Delta^{1/2}, \tag{3.20}$$

which completes the proof. □

Let τ_R and $\rho_{\Delta,R}$ be the stopping times as introduced in (2.22) and (3.12). For any $t \in [0,T]$, define

$$\nu_{\Delta,R} = \tau_R \wedge \rho_{\Delta,R} \quad \text{and} \quad e_\Delta(t) = X(t) - x_\Delta(t). \tag{3.21}$$

The following lemma plays a key role in the proof of the convergence rate.

Lemma 3.3 *Let Assumption 2.1 hold and assume that $\mu^{-1}(\varphi(\Delta)) \ge R \ge K_0$ for any $\Delta \in (0,1]$. Then for any $q \ge 1$,*

$$\mathbb{E}|e_\Delta(t\wedge\nu_{\Delta,R})|^q \le C\Delta^{q/2}(\mu^{-1}(\varphi(\Delta)))^{2q}, \quad \text{for all } t \ge 0, \tag{3.22}$$

where C is a positive constant independent of Δ and R.

Proof For any $q \ge 2$, using the Itô formula gives that

$$\mathbb{E}|e_\Delta(t\wedge\nu_{\Delta,R})|^q \le \mathbb{E}\int_0^{t\wedge\nu_{\Delta,R}} q|e_\Delta(s)|^{q-2}\left(e_\Delta(s)[f(X(s)) - f_\Delta(\bar{x}_\Delta(s))] + \frac{q-1}{2}|g(X(s)) - g_\Delta(\bar{x}_\Delta(s))|^2\right)ds.$$

For any $s \in [0, t\wedge\nu_{\Delta,R}]$, we note that $1/R \le \bar{x}_\Delta(s) \le R$. In addition, we have condition $\mu^{-1}(\varphi(\Delta)) \ge R$, thus, $1/\mu^{-1}(\varphi(\Delta)) \le \bar{x}_\Delta(s) \le \mu^{-1}(\varphi(\Delta))$. Thus,

$$f_\Delta(\bar{x}_\Delta(s)) = f(\bar{x}_\Delta(s)) \quad \text{and} \quad g_\Delta(\bar{x}_\Delta(s)) = g(\bar{x}_\Delta(s)), \quad \text{for all } s \in [0, t\wedge\nu_{\Delta,R}].$$

Consequently,

$$\mathbb{E}|e_\Delta(t\wedge\nu_{\Delta,R})|^q \le \mathbb{E}\int_0^{t\wedge\nu_{\Delta,R}} q|e_\Delta(s)|^{q-2}\left(e_\Delta(s)[f(X(s)) - f(\bar{x}_\Delta(s))] + \frac{q-1}{2}|g(X(s)) - g(\bar{x}_\Delta(s))|^2\right)ds.$$

For $p > q$, the elementary inequality

$$(a+b)^2 \le (1+\theta)a^2 + \left(1+\frac{1}{\theta}\right)b^2, \quad \text{for all } a,b,\theta > 0,$$

gives that

$$\frac{q-1}{2}|g(X(s)) - g(\bar{x}_\Delta(s))|^2 \le \frac{p-1}{2}|g(X(s)) - g(x_\Delta(s))|^2 + \frac{(q-1)(p-1)}{2(p-q)}|g(x_\Delta(s)) - g(\bar{x}_\Delta(s))|^2.$$

Thus

$$\mathbb{E}|e_\Delta(t\wedge\nu_{\Delta,R})|^q \le I_{11} + I_{12}, \tag{3.23}$$

where

$$I_{11} = \mathbb{E}\int_0^{t\wedge\nu_{\Delta,R}} q|e_\Delta(s)|^{q-2}\left(e_\Delta(s)[f(X(s)) - f(x_\Delta(s))] + \frac{p-1}{2}|g(X(s)) - g(x_\Delta(s))|^2\right)ds$$

and

$$I_{12} = \mathbb{E}\int_0^{t\wedge\nu_{\Delta,R}} q|e_\Delta(s)|^{q-2}\left(e_\Delta(s)[f(x_\Delta(s)) - f(\bar{x}_\Delta(s))] + \frac{(q-1)(q-1)}{2(p-q)}|g(x_\Delta(s)) - g(\bar{x}_\Delta(s))|^2\right)ds.$$

By the monotonicity condition (3.2), we have

$$I_{11} \le qL^{\star} \int_0^t \mathbb{E}|e_\Delta(s\wedge \nu_{\Delta,R})|^q ds. \tag{3.24}$$

By the elementary inequality, we have

$$\begin{aligned} I_{12} &\le C\int_0^t \mathbb{E}|e_\Delta(s\wedge \nu_{\Delta,R})|^q ds + C\mathbb{E}\int_0^{t\wedge \nu_{\Delta,R}} (|f(x_\Delta(s)) - f(\bar{x}_\Delta(s))|^q + |g(x_\Delta(s)) - g(\bar{x}_\Delta(s))|^q)\, ds \\ &=: C\int_0^t \mathbb{E}|e_\Delta(s\wedge \nu_{\Delta,R})|^q ds + I_{121}. \end{aligned} \tag{3.25}$$

By (3.3), we have

$$\begin{aligned} I_{121} &= C\mathbb{E}\int_0^{t\wedge \nu_{\Delta,R}} (|f(x_\Delta(s)) - f(\bar{x}_\Delta(s))|^q + |g(x_\Delta(s)) - g(\bar{x}_\Delta(s))|^q)\, ds \\ &= C\mathbb{E}\int_0^{t\wedge \nu_{\Delta,R}} \left(\left| a_{-1}\left(\frac{1}{x_\Delta(s)} - \frac{1}{\bar{x}_\Delta(s)}\right) + [f_2(x_\Delta(s)) - f_2(\bar{x}_\Delta(s))] \right|^q + |g(x_\Delta(s)) - g(\bar{x}_\Delta(s))|^q \right) ds \\ &\le C\mathbb{E}\int_0^{t\wedge \nu_{\Delta,R}} \left|\frac{1}{x_\Delta(s)} - \frac{1}{\bar{x}_\Delta(s)}\right|^q ds + C\mathbb{E}\int_0^{t\wedge \nu_{\Delta,R}} (|f_2(x_\Delta(s)) - f_2(\bar{x}_\Delta(s))|^q + |g(x_\Delta(s)) - g(\bar{x}_\Delta(s))|^q)\, ds \\ &=: I_{122} + I_{123}. \end{aligned} \tag{3.26}$$

By (3.3) and the Hölder inequality as well as Lemmas 2.6 and 3.1, we derive that

$$\begin{aligned} I_{123} &= C\mathbb{E}\int_0^{t\wedge \nu_{\Delta,R}} (|f_2(x_\Delta(s)) - f_2(\bar{x}_\Delta(s))|^q + |g(x_\Delta(s)) - g(\bar{x}_\Delta(s))|^q)\, ds \\ &\le C\int_0^T \mathbb{E}\Big[(1 + |x_\Delta(s)|^{(\gamma-1)q} + |\bar{x}_\Delta(s)|^{(\gamma-1)q})|x_\Delta(s) - \bar{x}_\Delta(s)|^q\Big] ds \\ &\le C\int_0^T \Big((1 + \mathbb{E}|x_\Delta(s)|^{2(\gamma-1)q} + \mathbb{E}|\bar{x}_\Delta(s)|^{2(\gamma-1)q})\Big)^{1/2} \Big(\mathbb{E}|x_\Delta(s) - \bar{x}_\Delta(s)|^{2q}\Big)^{1/2} ds \\ &\le C\Big(\Delta^{2q}(\mu^{-1}(\varphi(\Delta)))^{2q} + \Delta^q\Big)^{1/2} \\ &\le C\Big(\Delta^q(\mu^{-1}(\varphi(\Delta)))^q + \Delta^{q/2}\Big). \end{aligned} \tag{3.27}$$

For any $s \in [0, t\wedge \nu_{\Delta,R}]$, we observe that

$$\frac{1}{\mu^{-1}(\varphi(\Delta))} \le \frac{1}{x_\Delta(s)} \le \mu^{-1}(\varphi(\Delta)) \quad \text{and} \quad \frac{1}{\mu^{-1}(\varphi(\Delta))} \le \frac{1}{\bar{x}_\Delta(s)} \le \mu^{-1}(\varphi(\Delta)). \tag{3.28}$$

Thus, by Lemma 3.1, we have

$$\begin{aligned} I_{122} =& C\mathbb{E}\int_0^{t\wedge \nu_{\Delta,R}} \left|\frac{x_\Delta(s) - \bar{x}_\Delta(s)}{x_\Delta(s)\bar{x}_\Delta(s)}\right|^q ds \\ &\le C(\mu^{-1}(\varphi(\Delta)))^{2q}\mathbb{E}\int_0^{t\wedge \nu_{\Delta,R}} |x_\Delta(s) - \bar{x}_\Delta(s)|^q ds \\ &\le C(\mu^{-1}(\varphi(\Delta)))^{2q}\mathbb{E}\int_0^T |x_\Delta(s) - \bar{x}_\Delta(s)|^q ds \\ &\le C(\mu^{-1}(\varphi(\Delta)))^{2q}\int_0^T \mathbb{E}|x_\Delta(s) - \bar{x}_\Delta(s)|^q ds \\ &\le C\Big(\Delta^q(\mu^{-1}(\varphi(\Delta)))^{3q} + \Delta^{q/2}(\mu^{-1}(\varphi(\Delta)))^{2q}\Big) \\ &\le C\Delta^{q/2}(\mu^{-1}(\varphi(\Delta)))^{2q}. \end{aligned} \tag{3.29}$$

By (3.26), (3.27) and (3.29), we conclude from (3.25) that

$$I_{12} \le C\int_0^t \mathbb{E}|e_\Delta(s\wedge \nu_{\Delta,R})|^q ds + C\Delta^{q/2}(\mu^{-1}(\varphi(\Delta)))^{2q}.$$

Thus, we have

$$\mathbb{E}|e_\Delta(t\wedge \nu_{\Delta,R})|^q \le C\int_0^t \mathbb{E}|e_\Delta(s\wedge \nu_{\Delta,R})|^q ds + C\Delta^{q/2}(\mu^{-1}(\varphi(\Delta)))^{2q}.$$

By the Gronwall inequality, we obtain that (3.22) holds for any $q \ge 2$. For any $q \in [1,2)$, (3.22) still holds by the Hölder inequality. Thus, the proof is finished. □

Now we are in a position to establish the strong convergence order of positivity-preserving truncated EM method.

Theorem 3.1 *Let Assumption 2.1 hold and let $q \in [1,\infty)$, $\varepsilon \in \left(0,\frac{1}{2}\right)$, $\hat{K}_0 \in \left[\frac{1}{X_0}\vee X_0 \vee \frac{a_0}{a_{-1}} \vee \left|\frac{a_2}{a_1}\right|^{\frac{1}{\gamma-1}},\infty\right)$. Then the truncated EM scheme (2.11) by setting $\mu^{-1}(\varphi(\Delta)) = \hat{K}_0\Delta^{-\frac{\varepsilon}{(2\gamma+1.5)\vee(\gamma+4.5)}}$ has the property that*

$$\mathbb{E}|X(T)-Y_\Delta(T)|^q \le C\Delta^{\frac{1}{2}-\varepsilon}, \quad \text{for all } T\in(0,\infty),\ \Delta\in(0,1], \tag{3.30}$$

where C is a positive constant independent of Δ.

Proof For any $q \in \left[1, \frac{p}{(4\gamma+4)\vee(2\gamma+10)}\right)$. Using the Young inequality, for any $\kappa > 0$, we have

$$\begin{aligned}\mathbb{E}|e_\Delta(T)|^q &= \mathbb{E}\left[|e_\Delta(T)|^q \mathbf{1}_{\{\tau_R>T \text{ and } \rho_{\Delta,R}>T\}}\right] + \mathbb{E}\left[|e_\Delta(T)|^q \mathbf{1}_{\{\tau_R\le T \text{ or } \rho_{\Delta,R}\le T\}}\right]\\ &\le \mathbb{E}\left[|e_\Delta(T\wedge\nu_{\Delta,R})|^q\right] + \frac{q\Delta^\kappa}{p}\mathbb{E}|e_\Delta(T)|^p + \frac{p-q}{p\Delta^{\kappa q/(p-q)}}\mathbb{P}(\tau_R\le T \text{ or } \rho_{\Delta,R}\le T),\end{aligned} \tag{3.31}$$

where e_Δ, τ_R and $\nu_{\Delta,R}$ have been defined in (3.21), (2.22) and (3.12). Lemmas 2.2 and 2.6 gives that

$$\frac{q\Delta^\kappa}{p}\mathbb{E}|e_\Delta(T)|^p \le C\Delta^\kappa. \tag{3.32}$$

Next, we derives a polynomial upper bound in the time step size on $\mathbb{P}(\tau_R \le T \text{ or } \rho_{\Delta,R}\le T)$. By Lemma 3.2, we get

$$\frac{p-q}{p\Delta^{\kappa q/(p-q)}}\mathbb{P}(\rho_{\Delta,R}\le T) \le C\Delta^{1/2}R^{(2\gamma+1)\vee(\gamma+4)}\Delta^{-\frac{\kappa q}{p-q}}. \tag{3.33}$$

Now, take $\alpha = \dfrac{q}{p-q}\gamma$ in (3.6), and set $\kappa = \dfrac{1}{2}$, $R = \mu^{-1}(\varphi(\Delta))$. Thus, $R = \hat{K}_0\Delta^{-\frac{q}{p-q}}$, which implies

$$\begin{aligned}\frac{p-q}{p\Delta^{\kappa q/(p-q)}}\mathbb{P}(\rho_{\Delta,R}\le T) &\le C\Delta^{\frac{1}{2}-\frac{[(2\gamma+1)\vee(\gamma+4)]q}{p-q}-\frac{\kappa q}{p-q}}\\ &= C\Delta^{\frac{1}{2}-\frac{[(2\gamma+1.5)\vee(\gamma+4.5)]q}{p-q}}.\end{aligned} \tag{3.34}$$

On the other hand, (2.23) gives

$$\frac{p-q}{p\Delta^{\kappa q/(p-q)}}\mathbb{P}(\tau_R\le T) \le \frac{C}{R^p}\Delta^{-\frac{\kappa q}{p-q}} = C\Delta^{\frac{pq}{p-q}-\frac{\kappa q}{p-q}} \le C\Delta^{1-\frac{\kappa q}{p-q}}. \tag{3.35}$$

Consequently,

$$\begin{aligned}\frac{p-q}{p\Delta^{\kappa q/(p-q)}}\mathbb{P}(\tau_R \le T \text{ or } \rho_{\Delta,R} \le T) &\le \frac{p-q}{p\Delta^{\kappa q/(p-q)}}[\mathbb{P}(\rho_{\Delta,R} \le T)+\mathbb{P}(\tau_R \le T)] \\ &\le C\Delta^{\frac{1}{2}-\frac{[(2\gamma+1.5)\vee(\gamma+4.5)]q}{p-q}}. \end{aligned} \tag{3.36}$$

Inserting (3.32) and (3.36) into (3.31) and applying Lemma 3.3, we have

$$\mathbb{E}|e_\Delta(T)|^q \le C\Delta^{q(\frac{1}{2}-2\alpha/\gamma)} + C\Delta^{\frac{1}{2}} + C\Delta^{\frac{1}{2}-\frac{[(2\gamma+1.5)\vee(\gamma+4.5)]q}{p-q}} \le C\Delta^{\frac{1}{2}-\frac{[(2\gamma+1.5)\vee(\gamma+4.5)]q}{p-q}}. \tag{3.37}$$

Combining this and Lemma 3.1 as well as the triangle inequality gives

$$\mathbb{E}|X(T)-\bar{x}_\Delta(T)|^q \le C\Delta^{\frac{1}{2}-\frac{[(2\gamma+1.5)\vee(\gamma+4.5)]q}{p-q}}, \quad \text{for all } T>0,\ \Delta\in(0,1]. \tag{3.38}$$

In the next step we prove that

$$\mathbb{E}|X(T)-\bar{y}_\Delta(T)|^q \le C\Delta^{\frac{1}{2}-\frac{[(2\gamma+1.5)\vee(\gamma+4.5)]q}{p-q}}, \quad \text{for all } T>0,\ \Delta\in(0,1]. \tag{3.39}$$

Let $\rho_{\Delta,R}$ be the stopping time defined in (3.12). Note that for any $\omega \in \{\rho_{\Delta,\mu^{-1}(\varphi(\Delta))} > T\}$, we have

$$\sup_{0\le t\le T}|\bar{x}_\Delta(t)| \le \sup_{0\le t\le T}|x_\Delta(t)| < \mu^{-1}(\varphi(\Delta)),$$

and

$$\inf_{0\le t\le T}|\bar{x}_\Delta(t)| \ge \inf_{0\le t\le T}|x_\Delta(t)| > \frac{1}{\mu^{-1}(\varphi(\Delta))}.$$

In other words, for any $\omega \in \{\rho_{\Delta,\mu^{-1}(\varphi(\Delta))} > T\}$, we have $1/\mu^{-1}(\varphi(\Delta)) < \bar{x}_\Delta(t) < \mu^{-1}(\varphi(\Delta))$, hence $\bar{x}_\Delta(t) = \bar{y}_\Delta(t)$ for any $t\in[0,T]$. In a similar fashion as (3.36) was obtained, we have

$$\begin{aligned}\mathbb{E}|\bar{x}_\Delta(T)-\bar{y}_\Delta(T)|^q &= \mathbb{E}\Big[|\bar{x}_\Delta(T)-\bar{y}_\Delta(T)|^q\mathbf{1}_{\{\rho_{\Delta,\mu^{-1}(\varphi(\Delta))}>T\}}\Big] + \mathbb{E}\Big[|\bar{x}_\Delta(T)-\bar{y}_\Delta(T)|^q\mathbf{1}_{\{\rho_{\Delta,\mu^{-1}(\varphi(\Delta))}\le T\}}\Big] \\ &= \mathbb{E}\Big[|\bar{x}_\Delta(T)-\bar{y}_\Delta(T)|^q\mathbf{1}_{\{\rho_{\Delta,\mu^{-1}(\varphi(\Delta))}\le T\}}\Big] \\ &\le \frac{q\Delta^\kappa}{p}\mathbb{E}|\bar{x}_\Delta(T)-\bar{y}_\Delta(T)|^p + \frac{p-q}{p\Delta^{\kappa q/(p-q)}}\mathbb{P}(\rho_{\Delta,\mu^{-1}(\varphi(\Delta))}\le T) \\ &\le C\Delta^\kappa\mathbb{E}\Big[|\bar{x}_\Delta(T)|^p + |\bar{y}_\Delta(T)|^p\Big] + \frac{p-q}{p\Delta^{\kappa q/(p-q)}}\mathbb{P}(\rho_{\Delta,\mu^{-1}(\varphi(\Delta))}\le T) \\ &\le C\Delta^{\frac{1}{2}-\frac{[(2\gamma+1.5)\vee(\gamma+4.5)]q}{p-q}}, \end{aligned} \tag{3.40}$$

which gives the order of the difference between the non-truncated-positive solution X_Δ and truncated-positive solution Y_Δ in the sense of L^q. Moreover, by the triangle inequality and (3.38) as well as (3.40), we have

$$\mathbb{E}|X(T)-\bar{y}_\Delta(T)|^q \le 2^{q-1}(\mathbb{E}|X(T)-\bar{x}_\Delta(T)|^q + \mathbb{E}|\bar{x}_\Delta(T)-\bar{y}_\Delta(T)|^q) \le C\Delta^{\frac{1}{2}-\frac{[(2\gamma+1.5)\vee(\gamma+4.5)]q}{p-q}}, \tag{3.41}$$

which gives (3.39). Now, taking $\varepsilon = \dfrac{[(2\gamma+1.5)\vee(\gamma+4.5)]q}{p-q}$ in (3.39) completes the proof of Theorem 3.1. □

The next result, Corollary 3.1, follows directly from Theorem 3.1 with $q=1$.

Corollary 3.1 *Consider the notation in Theorem 3.1. Let Assumption 2.1 hold and $\varepsilon\in(0,\frac{1}{2})$. Then the truncated EM scheme (2.11) by setting $\mu^{-1}(\varphi(\Delta)) = \hat{K}_0\Delta^{-\frac{\varepsilon}{(2\gamma+1.5)\vee(\gamma+4.5)}}$ has the property that*

$$\mathbb{E}|X(T)-Y_\Delta(T)| \le C\Delta^{\frac{1}{2}-\varepsilon}, \quad \text{for all } T\in(0,\infty),\ \Delta\in(0,1], \tag{3.42}$$

where C is a positive constant independent of Δ.

3.2 Convergence order over a finite time interval

Sometimes we need to approximate quantities that are path-dependent, for example the European barrier option value. Under these situations, we will need the strong convergence of the numerical approximation over a finite time interval. Let us begin to show the convergence order of the truncated EM over the time interval $[0,T]$.

Lemma 3.4 *Consider the notation in Lemma 3.3. Let Assumption 2.1 hold, $\mu^{-1}(\varphi(\Delta)) \geq R \geq K_0$ for any $\Delta \in (0,1]$, and let $p > 8((2\gamma+1.5)\vee(\gamma+4.5))+4$. Then*

$$\mathbb{E}\Big[\sup_{0\leq u\leq t}|e_\Delta(u\wedge\nu_{\Delta,R})|^2\Big] \leq C\Delta^{\frac{1}{4}-\frac{2[(2\gamma+1.5)\vee(\gamma+4.5)]}{p-4}}, \quad \text{for all } t\geq 0, \tag{3.43}$$

where C is a positive constant independent of Δ.

Proof We use the same notation as in the proof of Lemma 3.3. By the Itô formula, we have that

$$\mathbb{E}\Big[\sup_{0\leq u\leq t\wedge\nu_{\Delta,R}}|e_\Delta(u)|^2\Big] \leq J_6 + J_7, \tag{3.44}$$

where

$$J_6 = \mathbb{E}\Big[\sup_{0\leq u\leq t\wedge\nu_{\Delta,R}}\int_0^u\Big(2[X(s)-x_\Delta(s)][f(X(s))-f(\bar{x}_\Delta(s))]+|g(X(s))-g(\bar{x}_\Delta(s))|^2\Big)ds\Big],$$

$$J_7 = \mathbb{E}\Big[\sup_{0\leq u\leq t\wedge\nu_{\Delta,R}}\int_0^u 2[X(s)-x_\Delta(s)][g(X(s))-g(\bar{x}_\Delta(s))]dB(s)\Big].$$

By (3.2), we have

$$\begin{aligned}
J_6 &\leq \mathbb{E}\Big[\sup_{0\leq u\leq t\wedge\nu_{\Delta,R}}\int_0^u\Big(2[X(s)-x_\Delta(s)][f(X(s))-f(x_\Delta(s))]+2|g(X(s))-g(x_\Delta(s))|^2\Big)ds\Big]\\
&\quad+\mathbb{E}\Big[\sup_{0\leq u\leq t\wedge\nu_{\Delta,R}}\int_0^u\Big(2[X(s)-x_\Delta(s)][f(x_\Delta(s))-f(\bar{x}_\Delta(s))]+2|g(x_\Delta(s))-g(\bar{x}_\Delta(s))|^2\Big)ds\Big]\\
&\leq C\mathbb{E}\Big[\sup_{0\leq u\leq t\wedge\nu_{\Delta,R}}\int_0^u\Big(|X(s)-x_\Delta(s)|^2+|f(x_\Delta(s))-f(\bar{x}_\Delta(s))|^2+|g(x_\Delta(s))-g(\bar{x}_\Delta(s))|^2\Big)ds\Big]\\
&\leq C\int_0^t\mathbb{E}|e_\Delta(s\wedge\nu_{\Delta,R})|^2ds+C\mathbb{E}\Big[\int_0^{t\wedge\nu_{\Delta,R}}\Big(|f(x_\Delta(s))-f(\bar{x}_\Delta(s))|^2+|g(x_\Delta(s))-g(\bar{x}_\Delta(s))|^2\Big)ds\Big].
\end{aligned}\tag{3.45}$$

In the same way as Lemma 3.3 was proved, we can show that

$$J_6 \leq C\int_0^t\mathbb{E}|e_\Delta(s\wedge\nu_{\Delta,R})|^2ds+C\Delta(\mu^{-1}(\varphi(\Delta)))^4. \tag{3.46}$$

Moreover, by the Burkholder-Davis-Gundy inequality and condition (3.3), we derive that

$$\begin{aligned}
J_7 &\leq 8\sqrt{2}\mathbb{E}\Big[\int_0^{t\wedge\nu_{\Delta,R}}|e_\Delta(s)|^2|g(X(s))-g(\bar{x}_\Delta(s))|^2ds\Big]^{1/2}\\
&\leq 8\sqrt{2}\mathbb{E}\Big[\sup_{0\leq u\leq t\wedge\nu_{\Delta,R}}|e_\Delta(u)|^2\int_0^{t\wedge\nu_{\Delta,R}}|g(X(s))-g(\bar{x}_\Delta(s))|^2ds\Big]^{1/2}\\
&\leq \frac{1}{2}\mathbb{E}\left[\sup_{0\leq u\leq t\wedge\nu_{\Delta,R}}|e_\Delta(u)|^2\right]+C\mathbb{E}\int_0^{t\wedge\nu_{\Delta,R}}|g(X(s))-g(\bar{x}_\Delta(s))|^2ds\\
&\leq \frac{1}{2}\mathbb{E}\left[\sup_{0\leq u\leq t\wedge\nu_{\Delta,R}}|e_\Delta(u)|^2\right]+C\int_0^T\mathbb{E}\Big[(1+|X(s)|^{2\gamma-2}+|\bar{x}_\Delta(s)|^{2\gamma-2})|X(s)-\bar{x}_\Delta(s)|^2\Big]ds.
\end{aligned}\tag{3.47}$$

Let $p > 8((2\gamma+1.5)\vee(\gamma+4.5))+4$. By the Hölder inequality and Lemmas 2.1, 2.6 as well as (3.38), we have that for any $0 \le s \le T$,

$$\begin{aligned}&\mathbb{E}\Big[(1+|X(s)|^{2\gamma-2}+|\bar{x}_\Delta(s)|^{2\gamma-2})|X(s)-\bar{x}_\Delta(s)|^2\Big]\\&\le \Big[\mathbb{E}(1+|X(s)|^{4\gamma-4}+|\bar{x}_\Delta(s)|^{4\gamma-4})\Big]^{1/2}\Big[\mathbb{E}|X(s)-\bar{x}_\Delta(s)|^4\Big]^{1/2}\\&\le C\Delta^{\frac{1}{4}-\frac{2[(2\gamma+1.5)\vee(\gamma+4.5)]}{p-4}}.\end{aligned}$$

Inserting (3.46) and (3.47) into (3.44), we have

$$\begin{aligned}\mathbb{E}\Big[\sup_{0\le u\le t\wedge\nu_{\Delta,R}}|e_\Delta(u)|^2\Big] &\le C\int_0^t \mathbb{E}\Big[\sup_{0\le u\le s\wedge\nu_{\Delta,R}}|e_\Delta(u)|^2\Big]ds + C\Delta(\mu^{-1}(\varphi(\Delta)))^4 + C\Delta^{\frac{1}{4}-\frac{2[(2\gamma+1.5)\vee(\gamma+4.5)]}{p-4}}\\&\le C\int_0^t \mathbb{E}\Big[\sup_{0\le u\le s\wedge\nu_{\Delta,R}}|e_\Delta(u)|^2\Big]ds + C\Delta^{\frac{1}{4}-\frac{2[(2\gamma+1.5)\vee(\gamma+4.5)]}{p-4}}.\end{aligned}$$

Finally, the Gronwll inequality yields the required assertion. □

Theorem 3.2 *Let Assumption 2.1 hold and $\varepsilon \in (0,\frac{1}{4})$, $\hat{K}_0 \in \big[\frac{1}{X_0}\vee X_0 \vee \frac{a_0}{a_{-1}} \vee \big|\frac{a_2}{a_1}\big|^{\frac{1}{\gamma-1}},\infty\big)$. Then the truncated EM scheme (2.11) by setting $\mu^{-1}(\varphi(\Delta)) = \hat{K}_0\Delta^{-\frac{\varepsilon}{(2\gamma+1.5)\vee(\gamma+4.5)}}$ has the property that*

$$\mathbb{E}\Big[\sup_{0\le t\le T}|X(t)-\bar{y}_\Delta(t)|^2\Big] \le C\Delta^{\frac{1}{4}-\varepsilon}, \quad \textit{for all } T\in(0,\infty),\ \Delta\in(0,1], \tag{3.48}$$

where and C is a positive constant independent of Δ.

Proof We use the same notation as in the proof of Theorem 3.1. Using the Young inequality, we can show that for any $\delta > 0$ and sufficiently large p,

$$\begin{aligned}\mathbb{E}\left[\sup_{0\le t\le T}|e_\Delta(t)|^2\right] &\le \mathbb{E}\left[\sup_{0\le t\le T}|e_\Delta(t)|^2\mathbf{1}_{\{\tau_R>T \text{ and } \rho_{\Delta,R}>T\}}\right] + \frac{2\delta}{p}\mathbb{E}\left[\sup_{0\le t\le T}|e_\Delta(t)|^2\right]\\&\quad+\frac{p-2}{p\delta^{2/(p-2)}}\mathbb{P}(\tau_R\le T \text{ or } \rho_{\Delta,R}\le T)\\&\le \mathbb{E}\left[\sup_{0\le t\le T}|e_\Delta(t\wedge\nu_{\Delta,R})|^2\right] + \frac{2C}{p}\delta + \frac{p-2}{p\delta^{2/(p-2)}}\mathbb{P}(\tau_R\le T \text{ or } \rho_{\Delta,R}\le T).\end{aligned} \tag{3.49}$$

Set $\delta = \Delta^{\frac{1}{4}}$, $\alpha = \dfrac{2\gamma}{p-4}$ and $\varepsilon = \dfrac{2[(2\gamma+1.5)\vee(\gamma+4.5)]}{p-4}$. Thus, $R = \hat{K}_0\Delta^{-\frac{2}{p-4}} = \hat{K}_0\Delta^{-\frac{\varepsilon}{(2\gamma+1.5)\vee(\gamma+4.5)}}$. By Lemma 3.2, we get that

$$\mathbb{P}(\rho_{\Delta,R}\le T) \le CR^{(2\gamma+1)\vee(\gamma+4)}\Delta^{\frac{1}{2}} \le C\Delta^{\frac{1}{2}-\frac{2[(2\gamma+1)\vee(\gamma+4)]}{p-4}}. \tag{3.50}$$

Thus,

$$\frac{p-2}{p\delta^{2/(p-2)}}\mathbb{P}(\rho_{\Delta,R}\le T) \le C\Delta^{\frac{1}{2}-\frac{2[(2\gamma+1)\vee(\gamma+4)]}{p-4}-\frac{1}{2(p-2)}}. \tag{3.51}$$

Thus, by Lemma 3.4 and (3.51), we have

$$\begin{aligned}\mathbb{E}\left[\sup_{0\le t\le T}|e_\Delta(t)|^2\right] &\le C\Delta^{\frac{1}{4}-\frac{2[(2\gamma+1.5)\vee(\gamma+4.5)]}{p-4}} + C\Delta^{\frac{1}{4}} + C\Delta^{\frac{1}{2}-\frac{2[(2\gamma+1)\vee(\gamma+4)]}{p-4}-\frac{1}{2(p-2)}}\\&\le C\Delta^{\frac{1}{4}-\frac{2[(2\gamma+1.5)\vee(\gamma+4.5)]}{p-4}},\end{aligned}$$

which means that

$$\mathbb{E}\Big[\sup_{0\le t\le T}|X(t)-x_\Delta(t)|^2\Big]\le C\Delta^{\frac{1}{4}-\varepsilon}. \tag{3.52}$$

In the same way as (3.40) was obtained, we can show that

$$\mathbb{E}\Big[\sup_{0\le t\le T}|\bar{x}_\Delta(t)-\bar{y}_\Delta(t)|^2\Big]\le C\Delta^{\frac{1}{4}-\varepsilon}. \tag{3.53}$$

By Lemma 2.9, (3.52) and (3.53) as well as the triangle inequality, we get the assertion (3.48). The proof is finished. □

4 Numerical experiments

In this section, we will test the following schemes: the truncated EM scheme (TEM for short) defined in (2.11) by setting $\pi_\Delta(x)=\hat{K}_0\Delta^{\frac{\varepsilon}{(2\gamma+1.5)\vee(\gamma+4.5)}}\vee x\wedge\hat{K}_0\Delta^{-\frac{\varepsilon}{(2\gamma+1.5)\vee(\gamma+4.5)}}$; the Euler-Maruyama (EM) [6]

$$Y_{k+1}=Y_k+f(Y_k)\Delta+g(|X_k|)\Delta B_k,\quad Y_0=X_0;$$

the backward Euler-Maruyama (BEM) [24]

$$Y_{k+1}=Y_k+f(Y_{k+1})\Delta+g(X_k)\Delta B_k,\quad Y_0=X_0;$$

and logarithmic truncated Euler-Maruyama (log TE) [27,17]

$$Y_{k+1}=Y_k+F(\pi^\star(Y_k))\Delta+G(\pi^\star(Y_k))\Delta B_k,\quad Z_{k+1}=\exp(Y_{k+1}),\quad Y_0=\log(X_0),\quad Z_0=X_0,$$

where

$$F(x)=\exp(-x)f(\exp(x))-0.5\exp(-2x)|g(\exp(x))|^2,\quad G(x)=\exp(-x)g(\exp(x)),$$
$$\pi^\star(x)=(-R)\vee(x\wedge R),\quad R=C-\frac{\log\Delta}{2\gamma}.$$

The aim of the tests is to compare performance of the methods: their convergence orders, positivity preservation and computational costs. The experiments were performed on a Windows desktop computer with an Intel Core CPU i5-9400 (2.90 GHz). To examine the strong convergence in Δ for AIT models below, we run M samples at the final time T to estimate the average absolute error

$$AAE:=\frac{1}{M}\sum_{i=1}^{M}|X^i(T)-Y^i_\Delta(T)|,$$

and the root mean square error

$$RMSE:=\sqrt{\frac{1}{M}\sum_{i=1}^{M}|X^i(T)-Y^i_\Delta(T)|^2},$$

where $X^i(T)$ and $X^i_\Delta(T)$ denote the ith true solution and the ith numerical solution, respectively.

Moreover, we are interested in such scenario that X_Δ does not coincide with Y_Δ defined in TEM scheme (2.11). This event occurs when X_Δ leaves a truncated domain at least once and is then drawn back by the truncation. Thus, we define the following events

$$\mathbf{E}_1(\Delta) := \left\{ \min_{1\le k\le T/\Delta} X_\Delta(t_k) \le 1/R \quad \text{or} \quad \max_{1\le k\le T/\Delta} X_\Delta(t_k) \ge R \right\}, \tag{4.1}$$

$$\mathbf{E}_2(\Delta) := \left\{ \min_{1\le k\le T/\Delta} X_\Delta(t_k) < 0 \right\}. \tag{4.2}$$

In the first part, we explore the strong convergence of TEM in the sense of AAE and RMSE and bounds on $\mathbb{P}(\mathbf{E}_1(\Delta))$. In the second part we compare the positivity preservation and computational costs of the above methods.

4.1 Strong convergence and step size dependence of $\mathbb{P}(\mathbf{E}_1(\Delta))$

Example 4.1 Consider AIT model (1.1) with the following parameters that were estimated from financial data (see [1,3])

$$a_{-1} = 1.041\times 10^{-4}, \quad a_0 = 5.652\times 10^{-3}, \quad a_1 = 9.648\times 10^{-2}, \quad a_2 = 5.349\times 10^{-1},$$
$$\lambda = 1.329\times 10^{-2}, \quad \gamma = 2, \quad \theta = 1.4999, \quad X_0 = 0.06.$$

Clearly, Assumption 2.1 is satisfied in the above setting, while condition $\gamma > 4\theta - 3$ is violated. Thus, log TE is not applicable to this example, see [27, Example 5.5]. Set $q = 2$, $\varepsilon = 13/28$,

$$\pi_\Delta(x) = 40\Delta^{1/14} \vee (x \wedge 40\Delta^{-1/14}), \quad \text{for all } x \in \mathbb{R}. \tag{4.3}$$

According to Theorem 3.1, TEM scheme (2.11) with π_Δ given by (4.3) has the property that

$$\mathbb{E}|X(T) - Y_\Delta(T)|^2 \le C\Delta^{1/2-\varepsilon} = C\Delta^{1/28}. \tag{4.4}$$

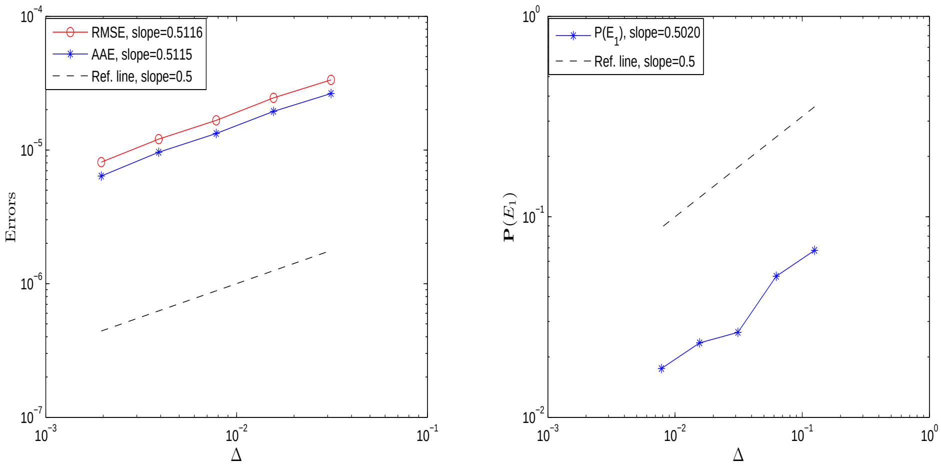
We plot the average absolute and the root mean square errors at $T = 4$ for five different step sizes 2^{-j}, $j = 5, 6, \cdots 9$, in Fig. (4.1a). Here, we set $M = 1000$. Numerical solution with step size 2^{-12} is taken as the reference solution. For the plotted trial, the empirical order of convergence in RMSE is 0.5116 and that is much higher than the theoretical value 0.018. Simulation evidence reveals that TEM scheme performs well in terms of errors in mean absolute and mean square values.

We observe from (2.11) that in some scenarios negative trajectories for X_Δ will appear. For instance, in this experiment for $\Delta = 2^{-3}$, we observed 10 out of 2000 trajectories for which $\mathbf{E}_2$ happened over the time interval $[0,4]$. The percentages of TEM solutions leaving the domain of the true solution to Example 4.1 for different step size is presented in the last row in Table 4.1. The second row in Table 4.1 contains the percentages of TEM solutions escaping from different truncated intervals, that is the probability of the event defined by (4.1) has occurred. We observe from Table 4.1 that the above percentage goes to zero as the step size tends to zero.

Moreover, in order to investigate numerically the dependence of $\mathbb{P}(\mathbf{E}_1(\Delta))$ on the step size, we assume that $\mathbb{P}(\mathbf{E}_1(\Delta))$ obeys a power law relation $\mathbb{P}(\mathbf{E}_1(\Delta)) = C\Delta^\beta$ for $C, \beta > 0$, so that $\log \mathbb{P}(\mathbf{E}_1(\Delta)) = \log C + \beta \log \Delta$. A least square fit for $\log C$ and β produces the value 0.5020 for β, see Fig. 4.1b. On the other hand, by our theoretical estimate for $\mathbb{P}(\mathbf{E}_1(\Delta))$, we conclude from Lemma 3.2 that

$$\mathbb{P}(\mathbf{E}_1(\Delta)) \le CR^6\Delta^{1/2} = C\Delta^{1/2-3/7} \approx \mathbf{C}\Delta^{0.0714}. \tag{4.5}$$

Such a result is seemingly conservative as the treatment in Lemma 3.2 amplifies the upper estimate bound for $\mathbb{P}(\rho_{\Delta,R} \le T)$. We conjecture that adopting an explicit estimate method for $\mathbb{P}(\rho_{\Delta,R} \le T)$ rather than the method of Lyapunov estimate may help to improve this bound and will further achieve a superior convergence order.



(a) AAE and RMSE of TEM

(b) $\mathbb{P}(\mathbf{E}_1(\Delta))$ vs. Δ

Fig. 4.1: Simulations of AIT model (4.1)

Table 4.1: Percentages of TEM solutions escaping from truncated domains for Example 4.1

Δ	2^{-3}	2^{-4}	2^{-5}	2^{-6}	2^{-7}
$(1/R,R)$	$(0.0215,46.4052)$	$(0.0205,48.7605)$	$(0.0195,51.2355)$	$(0.0186,53.8360)$	$(0.0177,56.5685)$
$\mathbb{P}(\mathbf{E}_1(\Delta))$	6.80%	5.05%	2.65%	2.35%	1.75%
$\mathbb{P}(\mathbf{E}_2(\Delta))$	0.05%	0	0	0	0

Table 4.2: CPU times for the selected schemes for Example 4.2

Δ	2^{-4}	2^{-5}	2^{-6}	2^{-7}	2^{-8}	**C**	r
TEM	0.291s	0.569s	1.147s	2.293s	4.526s	0.0188	0.9890
log TE	0.188s	0.355s	0.658s	1.301s	2.590s	0.0115	0.9773
BEM	31.257s	59.932s	115.831s	224.080s	435.493s	2.1796	0.9552

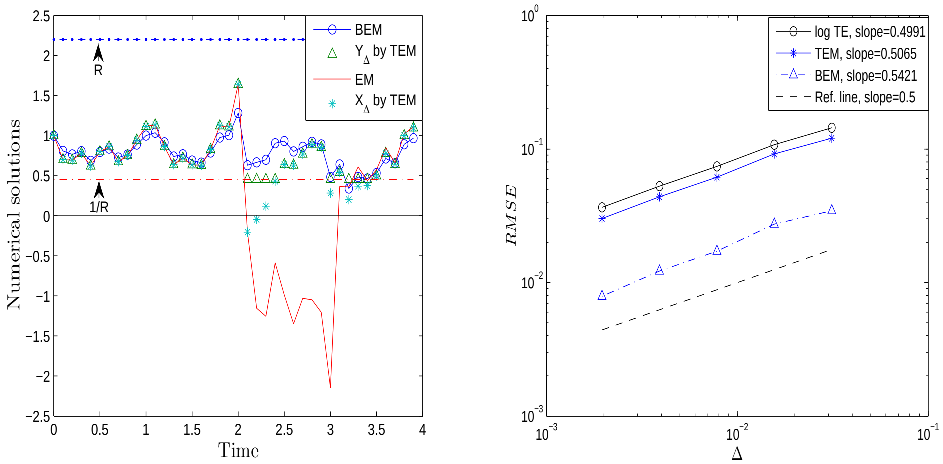
4.2 Positivity preservation and computational costs

Example 4.2 Consider AIT model (1.1) with the following parameters

$$a_{-1} = 1.5, \quad a_0 = 2, \quad a_1 = 1, \quad a_2 = 2, \quad \lambda = 1, \quad \gamma = 4, \quad \theta = 1.5, \quad X_0 = 1,$$

see [17, Example 5.3]. Clearly, the above parameters satisfy the condition $\gamma > 4\theta - 3$, which implies that Assumption 2.1 also holds. Therefore, TEM, EM, BEM, and log TE are applicable in the above setting. For comparison, we plot some trajectories of all methods except log TE in Fig. 4.2a. As expected, that approximations from Y_Δ defined by (2.11) and BEM can preserve positivity, while the EM and X_Δ defined by (2.11) approximations cannot.

For each scheme, we also computed an average of the experimental order of convergence by determining the best fitting line in a least-squares sense for the logarithmically scaled errors. The slopes of these lines are 0.50, 0.51, and 0.54 for log TE, TEM, and BEM, respectively, see Fig 4.2b.



(a) Trajectories of BEM, TEM, and EM

(b) RMSEs of BEM, TEM, and log TE

Fig. 4.2: Simulations of AIT model (4.2)

In Table 4.2, we test the comparable efficiencies as measured by CPU time in seconds. We also assume that the CPU time obeys a power law relation $y = \mathbf{C}\Delta^{-r}$, for all $\Delta \in (0,1]$. Nonlinear fit results for these relations for TEM, BEM, log TE schemes are also illustrated in the last two columns in Table 4.2. We observe that the log TE method is the fastest, because the exponent function performs faster than the negative power function in MATLAB. While the BEM method is the slowest. Under the same computational environment, the explicit TEM scheme (2.11) for Example (4.2) is more than approximately one hundred times faster than the implicit BEM scheme. Numerical experiments demonstrate that explicit schemes outperform the implicit in terms of computational costs.

5 Conclusion

In this work, a modified truncated EM method has been constructed and analyzed for a generalized Ait-Sahalia interest rate model. This kind of positivity-preserving method is proved to possess approximate $1/2$ order strong convergence in L^1-norm under certain reasonable assumptions. The numerical testing carried out shows that the explicit TEM scheme yields superior results over the implicit BEM scheme.

Acknowledgments

The authors would like to thank the National Natural Science Foundation of China (62273003, 12271003, 72301173), the Natural Science Foundation of Universities in Anhui Province (2022AH050993), Natural Science Research Project of Anhui Educational Committee (2023AH010011), the Startup Foundation for Introduction Talent of AHPU (2021YQQ058), and the Enterprise Cooperation Project of Anhui Institute of Future Technology (2023qyhz07) for their financial support.

Declaration

Conflict of interest: The authors declare that there is no conflict of interest regarding the publication of this paper.

References

1. Aït-Sahalia, Y.: Testing continuous-time models of the spot interest rate. Review of Financial Studies **9**(2), 385–426 (1996)
2. Beyn, W.J., Isaak, E., Kruse, R.: Stochastic C-stability and B-consistency of explicit and implicit Euler-type schemes. Journal of Scientific Computing **67**(3), 955–987 (2016). DOI 10.1007/s10915-015-0114-4
3. Chan, K.C., Karolyi, G.A., Longstaff, F.A., Anthony, B.S.: An empirical comparison of alternative models of short-term interest rates. The Journal of Finance **47**(3), 1209–1227 (1992)
4. Chassagneux, J.F., Jacquier, A., Mihaylov, I.: 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). DOI 10.1137/15m1017788
5. Chen, L., Gan, S.: Strong convergence and stationary distribution of an explicit scheme for the Wright-Fisher model. Journal of Computational and Applied Mathematics **424**, 115017 (2023). DOI 10.1016/j.cam.2022.115017
6. Cheng, S.R.: Highly nonlinear model in finance and convergence of Monte Carlo simulations. Journal of Mathematical Analysis and Applications **353**(2), 531–543 (2009). DOI 10.1016/j.jmaa.2008.12.028
7. Cox, J.C., Ingersoll, J.E., Ross, S.A.: A theory of the term structure of interest rates. Econometrica **53**(2), 385–407 (1985)
8. Deng, S., Fei, C., Fei, W., Mao, X.: Generalized Ait-Sahalia-type interest rate model with Poisson jumps and convergence of the numerical approximation. Physica A: Statistical Mechanics and its Applications **533**, 122057 (2019). DOI 10.1016/j.physa.2019.122057
9. Deng, S., Fei, C., Fei, W., Mao, X.: Tamed EM schemes for neutral stochastic differential delay equations with superlinear diffusion coefficients. Journal of Computational and Applied Mathematics **388**, 113269 (2021). DOI 10.1016/j.cam.2020.113269
10. Deng, S., Fei, C., Fei, W., Mao, X.: The truncated EM method for jump-diffusion SDDEs with super-linearly growing diffusion and jump coefficients. Journal of Computational Mathematics (2023). DOI 10.4208/jcm.2204-m2021-0270
11. Dereich, S., Neuenkirch, A., Szpruch, L.: An Euler-type method for the strong approximation of the Cox-Ingersoll-Ross process. Proceedings: Mathematical Physical and Engineering Sciences **468**(2140), 1105–1115 (2012)
12. Emmanuel, C., Mao, X.: Truncated EM numerical method for generalised Ait-Sahalia-type interest rate model with delay. Journal of Computational and Applied Mathematics **383**, 113137 (2021). DOI 10.1016/j.cam.2020.113137
13. Halley, W.G., Malham, S.J.A., Wiese, A.: Positive and implicit stochastic volatility simulation. arXiv **0802.4411** (2008)
14. Higham, D.J., Mao, X., Szpruch, L.: 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)
15. Hong, Y., Lin, H., Wang, S.: Modeling the dynamics of Chinese spot interest rates. Journal of Banking and Finance **34**(5), 1047–1061 (2010). DOI 10.1016/j.jbankfin.2009.11.002
16. Kloeden, P., Neuenkirch, A.: Convergence of Numerical Methods for Stochastic Differential Equations in Mathematical Finance, vol. 14. World Scientific Publishing Co. Pte. Ltd., Singapore (2013)
17. Lei, Z., Gan, S., Chen, Z.: 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). DOI 10.1016/j.cam.2022.114758
18. Lewis, A.L.: Option Valution under Stochastic Volatility II. Finance Press, CA (2009)
19. Mao, X.: The truncated Euler-Maruyama method for stochastic differential equations. Journal of Computational and Applied Mathematics **290**, 370–384 (2015). DOI 10.1016/j.cam.2015.09.035
20. Mao, X., Wei, F., Wiriyakraikul, T.: Positivity preserving truncated Euler-Maruyama method for stochastic Lotka-Volterra competition model. Journal of Computational and Applied Mathematics **394**, 113566 (2021). DOI 10.1016/j.cam.2021.113566
21. Merton, R.C.: Theory of rational option pricing. Bell Journal of Economics **4**(1), 141–183 (1973)
22. Neuenkirch, A., Szpruch, L.: First order strong approximations of scalar SDEs defined in a domain. Numerische Mathematik **128**(1), 103–136 (2014). DOI 10.1007/s00211-014-0606-4
23. Scalone, C.: Positivity preserving stochastic theta-methods for selected SDEs. Applied Numerical Mathematics **172**, 351–358 (2022). DOI 10.1016/j.apnum.2021.10.017
24. Szpruch, L., Mao, X., Higham, D.J., Pan, J.: Numerical simulation of a strongly nonlinear Ait-Sahalia-type interest rate model. BIT Numerical Mathematics **51**(2), 405–425 (2011). DOI 10.1007/s10543-010-0288-y
25. Vasicek, O.: An equilibrium characterization of the term structure. Journal of Financial and Quantitative Analysis **5**(4), 177–188 (1977)
26. Wang, X., Wu, J., Dong, B.: Mean-square convergence rates of stochastic theta methods for SDEs under a coupled monotonicity condition. BIT Numerical Mathematics **60**(3), 759–790 (2020). DOI 10.1007/s10543-019-00793-0
27. Yi, Y., Hu, Y., Zhao, J.: Positivity preserving logarithmic Euler-Maruyama type scheme for stochastic differential equations. Communications in Nonlinear Science and Numerical Simulation **101**, 105895 (2021). DOI 10.1016/j.cnsns.2021.105895
28. Zhao, Y., Wang, X., Wang, M.: On the backward Euler method for a generalized Ait-Sahalia-type rate model with Poisson jumps. Numerical Algorithms **87**(3), 1321–1341 (2020). DOI 10.1007/s11075-020-01009-1