IMA Journal of Numerical Analysis (2018) **00**, 1–19
doi:10.1093/imanum/dry067

Strong order 1/2 convergence of full truncation Euler approximations to the Cox–Ingersoll–Ross process

ANDREI COZMA* AND CHRISTOPH REISINGER
Mathematical Institute, University of Oxford, Oxford OX2 6GG, UK
*Corresponding author: andrei.s.cozma@gmail.com

[Received on 4 May 2017; revised on 8 September 2018]

We study convergence properties of the full truncation Euler scheme for the Cox–Ingersoll–Ross (CIR) process in the regime where the boundary point zero is inaccessible. Under some conditions on the model parameters (precisely, when the so-called Feller ratio is greater than three) we establish the strong order 1/2 convergence in L^p of the scheme to the exact solution. For the global error criterion studied in this paper this is consistent with the optimal rate of strong convergence for approximations to the CIR process based on sequential evaluations of the driving Brownian motion.

Keywords: Cox–Ingersoll–Ross process; explicit full truncation Euler scheme; strong convergence order.

1. Introduction

Let $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geqslant 0}, \mathbb{P}\right)$ be a filtered probability space and $W = \left(W_t\right)_{t\geqslant 0}$ be a one-dimensional $\mathcal{F}_t$-adapted Brownian motion. A Cox–Ingersoll–Ross (CIR) process is defined by the stochastic differential equation (SDE):

$$dv_t = k(\theta - v_t)\mathrm{d}t + \xi\sqrt{v_t}\,\mathrm{d}W_t, \tag{1.1}$$

where v_0, k, θ and ξ are strictly positive real numbers. This SDE admits a unique strong solution, which is strictly positive when $2k\theta \geqslant \xi^2$ by the Feller test (see Karatzas & Shreve, 1991). In other words, the boundary point zero is inaccessible, i.e., $\mathbb{P}(\forall\, t > 0 : v_t > 0) = 1$, if and only if $2k\theta \geqslant \xi^2$. The CIR process was originally introduced in finance to model short-term interest rates (Cox *et al.*, 1985). Nowadays, due to its desirable properties like non-negativity, mean-reversion and analytical tractability, it plays a key role in the field of option pricing, for instance when modeling squared volatilities in the Heston model (see Heston, 1993). For a given $t \geqslant 0$ the CIR process v_t has a noncentral chi-squared conditional distribution and its increments can be simulated exactly (see Broadie & Kaya, 2006). However, when pricing financial derivatives written on an underlying process $S = (S_t)_{t\in[0,T]}$ modeled by a d-dimensional SDE, with CIR dynamics in one or more components, we need to evaluate

$$\mathbb{E}\left[f(S)\right], \tag{1.2}$$

where $f : \mathcal{C}([0,T], \mathbb{R}^d) \to \mathbb{R}$ is the discounted payoff. This expectation is rarely available in closed form, nor can S be sampled exactly. In this case we use Monte Carlo simulation methods (see Glasserman, 2004) and approximate the solution to the SDE using a suitable discretization scheme. Due to the non-zero probability of the approximation process becoming negative the standard Euler–Maruyama scheme applied to (1.1) is not well defined. Possible remedies are to set the process equal to zero when it turns negative (absorption fix), e.g., the full truncation Euler (FTE) scheme studied in this paper or reflect it in the origin (reflection fix). An overview of the explicit Euler schemes with different fixes at the boundary can be found in Lord *et al.* (2010). Alternatively, we can employ an implicit Euler or a Milstein scheme to discretize the CIR process.

Weak convergence is important when estimating expectations of payoffs like the one in (1.2). However, strong convergence plays a crucial role in multilevel Monte Carlo methods (see Heinrich, 1998; Kebaier, 2005; Giles, 2008; Giles *et al.*, 2009) and may be required for some complex path-dependent derivatives. Moreover, pathwise convergence follows automatically (see Kloeden & Neuenkirch, 2007).

The classical convergence theory (see Kloeden & Platen, 1995; Higham *et al.*, 2002) does not apply to the CIR process because the square-root diffusion coefficient is not Lipschitz. Consequently, a considerable amount of research has been devoted to the numerical approximation of (1.1) and alternative approaches have been employed to prove the strong convergence of various discretizations for the CIR process (for different error criteria).

Results in the literature concerned with positive strong convergence rates for numerical approximations to the CIR process were restricted to the regime where the boundary point is inaccessible until just recently. Strong convergence, at best with a logarithmic rate, of different Euler schemes including the partial truncation, the full truncation, the reflection and the symmetrized Euler schemes was established in Deelstra & Delbaen (1998), Alfonsi (2005), Higham & Mao (2005), Lord *et al.* (2010) and Gyöngy & Rásonyi (2011). The first non logarithmic rate was obtained in Berkaoui *et al.* (2008), where it was shown that the symmetrized Euler scheme converges strongly with the standard order 1/2 to the exact solution, although in a very restricted parameter regime. Strong convergence with order 1/2 of the backward (drift-implicit) Euler–Maruyama (BEM) scheme was later established in Dereich *et al.* (2012), and this rate was improved to 1 in Alfonsi (2013) and Neuenkirch & Szpruch (2014). Recently, Bossy & Olivero Quinteros (2017) proved the strong convergence with order 1 of the symmetrized Milstein scheme under some restrictive conditions on the parameters, whereas Chassagneux *et al.* (2016) proved the strong convergence of a modified Euler–Maruyama scheme with an order between 1/6 and 1 that depends on the parameters.

In the past few years there has been significant development in the accessible boundary case. Polynomial rates of strong convergence for an order of up to 1/2 were established in Hutzenthaler *et al.* (2014) and Hefter & Herzwurm (2018) for the BEM scheme and the truncated Milstein scheme, respectively. For a particular instance of a CIR process, namely the one-dimensional squared Bessel process, polynomial convergence rates equal to infinity and 1/2 were established in Hefter & Herzwurm (2017) for the class of adaptive algorithms and for the class of algorithms using equidistant grids, respectively.

In this paper we study the FTE scheme proposed in Lord *et al.* (2010). This scheme preserves the positivity of the original process, is easy to implement and, hence, a widely used scheme in practice. Perhaps most importantly it is found empirically to produce the smallest bias of various explicit Euler schemes with different fixes at the boundary (see Lord *et al.*, 2010). For a fixed time horizon $T > 0$, let $N \in \mathbb{N}$ and consider a uniform grid

$$T = N\delta t, t_n = n\delta t, \forall\, n \in \{0, 1, \dots, N\}. \tag{1.3}$$

We introduce the discrete-time auxiliary process

$$\tilde{v}^N_{t_{n+1}} = \tilde{v}^N_{t_n} + k\Big(\theta - \big(\tilde{v}^N_{t_n}\big)^+\Big)\delta t + \xi\sqrt{\big(\tilde{v}^N_{t_n}\big)^+}\,\delta W_{t_n}, \tilde{v}^N_0 = v_0, \forall\, n \in \{0, 1, \dots, N-1\}, \tag{1.4}$$

where $v^+ = \max(0, v)$ and $\delta W_{t_n} = W_{t_{n+1}} - W_{t_n}$, its continuous-time interpolation

$$\tilde{v}_t^N = \tilde{v}_{t_n}^N + k\left(\theta - \left(\tilde{v}_{t_n}^N\right)^+\right)(t - t_n) + \xi\sqrt{\left(\tilde{v}_{t_n}^N\right)^+}(W_t - W_{t_n}), \tag{1.5}$$

and, finally, the FTE scheme

$$\bar{v}_t^N = \left(\tilde{v}_{t_n}^N\right)^+, \tag{1.6}$$

whenever $t \in [t_n, t_{n+1})$. The convergence in L^1 of this scheme was proved in Lord *et al.* (2010). The convergence rate, however, remained an open question and our Theorem 1.1 is the first result to address it, to the best of our knowledge. For convenience we define in this paper the Feller ratio

$$\lambda = \frac{2k\theta}{\xi^2}. \tag{1.7}$$

We establish the strong convergence in L^p with order 1/2 of the scheme in the inaccessible boundary case, specifically, for a Feller ratio above three. Hence, for the global error criterion studied in this paper we obtain the optimal strong convergence rate for numerical approximations to the CIR process based on N sequential evaluations of the Brownian driver (Hefter *et al.*, 2017). The main and novel idea of the proof is to weigh the difference between the process v and its approximation $\tilde{v}^N$ by the former raised to a suitably chosen negative power, and prove the strong convergence with a rate of the weighted error. This, in turn, allows us to derive an upper bound for the actual error.

THEOREM 1.1 Suppose that $\lambda > 3$ and let $2 \leqslant p < \lambda - 1$. Then the FTE scheme converges strongly in L^p with order 1/2, i.e., there exists a constant $C > 0$ such that, for all $N \geqslant 1$,

$$\sup_{t\in[0,T]} \left(\mathbb{E}\left[|v_t - \bar{v}_t^N|^p\right]\right)^{\frac{1}{p}} \leqslant CN^{-\frac{1}{2}}. \tag{1.8}$$

To the best of our knowledge there is no simple extension of Theorem 1.1 to more general drift coefficients in SDE (1.1), like in Bossy & Olivero Quinteros (2017). We mention that the assumption on the Feller ratio from Theorem 1.1, i.e., $\lambda > 3$, appears in the literature as a sufficient condition for the strong convergence with a rate of several discretization schemes for the CIR process (for different error criteria). For example, this condition ensures in Corollary 4.1 in Chassagneux *et al.* (2016) the strong convergence with order 1/2 (and order 1 if $\lambda > 5$) of the modified Euler–Maruyama scheme, and in Proposition 3.1 in Neuenkirch & Szpruch (2014) the strong convergence with order 1 of the BEM scheme.

In Hefter & Jentzen (2017) a lower error bound was recently derived for all discretization schemes for the CIR process based on equidistant evaluations of the Brownian driver in the regime where the boundary point is accessible. In light of this result we demonstrate numerically that the FTE scheme achieves an optimal performance—in the L^1 sense—in half of the regime where the boundary point is accessible, where by optimal we mean that the empirical L^1 convergence rate is the best possible for equidistant discretization schemes for the CIR process.

The remainder of the paper is structured as follows. In Section 2 we prove the convergence with a rate of the scheme. In Section 3 we conduct numerical tests for the rate of convergence that validate and complement our theoretical findings. Finally, Section 4 contains a short discussion.

2. Convergence analysis

We need to control the polynomial moments of the CIR process and its FTE discretization.

LEMMA 2.1 The CIR process has bounded moments, i.e.,

$$\sup_{t\in[0,T]} \mathbb{E}\left[v_t^p\right] < \infty, \forall p > -\lambda. \tag{2.1}$$

Proof. Follows from Dereich *et al.* (2012) or Theorem 3.1 in Hurd & Kuznetsov (2008). □

LEMMA 2.2 The auxiliary process $\tilde{v}^N$ has uniformly bounded moments, i.e.,

$$\sup_{N\geqslant 1} \mathbb{E}\left[\sup_{t\in[0,T]} |\tilde{v}_t^N|^p\right] < \infty, \forall p \geqslant 1. \tag{2.2}$$

Proof. Integrating the auxiliary process $\tilde{v}^N$ defined in (1.5) we deduce that

$$\tilde{v}_t^N = v_0 + k\theta t - k\int_0^t \bar{v}_u^N \,\mathrm{d}u + \xi\int_0^t \sqrt{\bar{v}_u^N}\,\mathrm{d}W_u. \tag{2.3}$$

Using Hölder's inequality we get

$$\sup_{t\in[0,T]} |\tilde{v}_t^N|^p \leqslant 3^{p-1}\left(v_0 + k\theta T\right)^p + 3^{p-1}k^pT^{p-1}\int_0^T \left(\bar{v}_u^N\right)^p \mathrm{d}u + 3^{p-1}\xi^p \sup_{t\in[0,T]}\left|\int_0^t \sqrt{\bar{v}_u^N}\,\mathrm{d}W_u\right|^p. \tag{2.4}$$

Taking expectations on both sides and using the Burkholder–Davis–Gundy inequality yields

$$\mathbb{E}\left[\sup_{t\in[0,T]} |\tilde{v}_t^N|^p\right] \leqslant 3^{p-1}\left(v_0 + k\theta T\right)^p + \frac{1}{2}3^{p-1}\xi^pC_p + 3^{p-1}\left(k^p + \frac{1}{2}\xi^pC_p\right)T^p \sup_{t\in[0,T]} \mathbb{E}\left[\left(\bar{v}_t^N\right)^p\right], \tag{2.5}$$

for some constant $C_p > 0$. The conclusion follows from Proposition 3.8 in Cozma *et al.* (2018). □

By construction the FTE approximation $\bar{v}^N$ is non-negative. However, an important step in the convergence analysis lies in analyzing the behavior of the auxiliary process $\tilde{v}^N$ at the boundary. The next result derives a polynomial upper bound in the time step size on the probability of $\tilde{v}^N$ becoming negative. Similar results were established for the symmetrized Euler scheme, in Lemma 3.7 in Bossy & Diop (2007), and for the symmetrized Milstein scheme, in Lemma 2.2 in Bossy & Olivero Quinteros (2017). However, the FTE scheme has led to different technical challenges, and the arguments employed in the proofs of the aforementioned results do not apply here.

Suppose that $\lambda > 2$ and define

$$\bar{\lambda} = \inf\left\{x > 0 : \frac{(4x \vee \lambda)(\lambda - x)}{\lambda x(\lambda - x - 1)} < 1.99\sqrt{\lambda\pi}e^{\frac{\lambda}{2}} - 1\right\}. \tag{2.6}$$

We first show that $\bar{\lambda}$ exists and derive some bounds. Note that as the value of λ increases the left-hand side term of the inequality in (2.6) decreases and the right-hand side term increases. Hence, $\bar{\lambda}$ decreases as λ increases. In particular $0 < \bar{\lambda} < \bar{\lambda}|_{\lambda=2} \approx 0.176$.

Proposition 2.3 Suppose that $\lambda > 2$. Then there exists a constant $C > 0$ such that for all $N \geqslant 1$,

$$\sup_{0\leqslant n\leqslant N} \mathbb{P}\left(\tilde{v}^N_{t_n} \leqslant 0\right) \leqslant CN^{-\lambda+\bar{\lambda}+1}, \tag{2.7}$$

with $\bar{\lambda}$ as defined in (2.6).

Proof. Note that the exponent in (2.7) is negative. Fix $N > \lfloor kT \rfloor$ and define, for brevity,

$$\alpha_N = \frac{1}{2}\left(1 - \frac{kT}{N}\right) = \frac{1 - k\delta t}{2}. \tag{2.8}$$

First, consider the sequence $(c_j)_{0\leqslant j\leqslant N}$ given by

$$c_0 = \alpha_N, c_1 = \alpha_N - \alpha_N^2 \text{ and } c_{j+1} = c_j^2 + \alpha_N - \alpha_N^2, \forall\, 1 \leqslant j \leqslant N-1. \tag{2.9}$$

As $\alpha_N \in (0, 0.5)$ one can clearly see that $c_j \in (0, \alpha_N)$ for all $1 \leqslant j \leqslant N$. We will show by induction that

$$c_j \leqslant 1 - \alpha_N - \frac{\varphi_\lambda}{j - 1 + \eta_\lambda}, \forall\, 1 \leqslant j \leqslant N, \tag{2.10}$$

where

$$\varphi_\lambda = 1 - \frac{\bar{\lambda}}{\lambda} \text{ and } \eta_\lambda = \frac{(\lambda - \bar{\lambda})(4\bar{\lambda} \vee \lambda)}{\lambda\bar{\lambda}}. \tag{2.11}$$

Since

$$\frac{\varphi_\lambda}{\eta_\lambda} = \frac{\bar{\lambda}}{4\bar{\lambda} \vee \lambda} \leqslant \frac{1}{4} \leqslant (1 - \alpha_N)^2, \tag{2.12}$$

(2.10) holds when $j = 1$. Suppose that (2.10) holds for some $1 \leqslant j \leqslant N - 1$, then

$$c_{j+1} \leqslant \alpha_N - \alpha_N^2 + \left(1 - \alpha_N - \frac{\varphi_\lambda}{j - 1 + \eta_\lambda}\right)^2 \tag{2.13}$$

and some simple computations lead to the following sufficient condition for the induction step,

$$(j - 1 + \eta_\lambda)^2(1 - 2\alpha_N) + (j - 1 + \eta_\lambda)(1 - 2\alpha_N) + (j - 1)(1 - \varphi_\lambda) + \eta_\lambda(1 - \varphi_\lambda) - \varphi_\lambda \geqslant 0, \tag{2.14}$$

which clearly holds. For convenience define another sequence $(a_j)_{0\leqslant j\leqslant N}$ given by

$$a_j = \frac{2(\alpha_N - c_j)}{\xi^2\delta t}, \forall\, 0 \leqslant j \leqslant N, \tag{2.15}$$

such that

$$a_0 = 0, \quad a_1 = \frac{2\alpha_N^2}{\xi^2 \delta t} \quad \text{and} \quad a_{j+1} = 2\alpha_N a_j - \frac{1}{2} a_j^2 \xi^2 \delta t, \quad \forall\, 1 \leqslant j \leqslant N-1. \tag{2.16}$$

A sequence similar to $(a_j)_{0\leqslant j\leqslant N}$ was analyzed in Bossy & Diop (2007). However, a sharper lower bound than the one obtained in Lemma 3.6 in Bossy & Diop (2007) is needed for our purposes. We now show that, for N large enough,

$$\mathscr{S}_1 = \sum_{n=0}^{N-2} \prod_{j=0}^{n} \exp\left\{-a_j k\theta \delta t\right\} \leqslant 2\sqrt{\lambda(1-k\delta t)\pi}\, e^{\frac{\lambda}{2}(1-k\delta t)}, \tag{2.17}$$

a bound that will be of use later in the proof. Using (2.10) we get

$$\begin{aligned}
\mathscr{S}_1 &= \sum_{n=0}^{N-2} \exp\left\{\lambda \sum_{j=1}^{n} (c_j - \alpha_N)\right\} \\
&\leqslant \sum_{n=0}^{N-2} \exp\left\{\lambda \sum_{j=1}^{n} \left(1 - 2\alpha_N - \frac{\varphi_\lambda}{j-1+\eta_\lambda}\right)\right\} \\
&\leqslant \sum_{n=0}^{N-2} \exp\left\{\lambda k T \frac{n}{N} - \lambda\varphi_\lambda \int_0^n (x+\eta_\lambda)^{-1} \mathrm{d}x\right\} \\
&\leqslant \eta_\lambda^{\lambda\varphi_\lambda} \sum_{n=0}^{N-1} e^{\lambda k T \frac{n}{N}} (\eta_\lambda + n)^{-\lambda\varphi_\lambda}.
\end{aligned} \tag{2.18}$$

To improve readability and avoid further notations let $\epsilon = 0.002$ for the remainder of the proof. Since $\lambda\varphi_\lambda = \lambda - \bar{\lambda} > 1$ the Hurwitz (generalized Riemann) zeta function

$$\zeta(\lambda\varphi_\lambda, \eta_\lambda) = \sum_{n=0}^{\infty} (\eta_\lambda + n)^{-\lambda\varphi_\lambda} \tag{2.19}$$

converges and hence, for N large enough, we have

$$\sum_{n=1}^{N-1} \left(e^{\lambda k T \frac{n}{N}} - 1 - \epsilon\right)(\eta_\lambda + n)^{-\lambda\varphi_\lambda} \leqslant \sum_{n > \frac{\log(1+\epsilon)}{\lambda k T} N} \left(e^{\lambda k T} - 1\right)(\eta_\lambda + n)^{-\lambda\varphi_\lambda} \leqslant \epsilon\eta_\lambda^{-\lambda\varphi_\lambda}, \tag{2.20}$$

which implies that

$$\sum_{n=0}^{N-1} e^{\lambda k T \frac{n}{N}} (\eta_\lambda + n)^{-\lambda\varphi_\lambda} \leqslant (1+\epsilon) \sum_{n=0}^{N-1} (\eta_\lambda + n)^{-\lambda\varphi_\lambda} \leqslant (1+\epsilon)\zeta(\lambda\varphi_\lambda, \eta_\lambda). \tag{2.21}$$

However,

$$\zeta(\lambda\varphi_\lambda,\eta_\lambda) = \eta_\lambda^{-\lambda\varphi_\lambda} + \sum_{n=1}^{\infty}(\eta_\lambda+n)^{-\lambda\varphi_\lambda} \leqslant \eta_\lambda^{-\lambda\varphi_\lambda} + \int_0^\infty (\eta_\lambda+x)^{-\lambda\varphi_\lambda}\mathrm{d}x, \tag{2.22}$$

and hence,

$$\zeta(\lambda\varphi_\lambda,\eta_\lambda) \leqslant \eta_\lambda^{-\lambda\varphi_\lambda} + \frac{\eta_\lambda^{-\lambda\varphi_\lambda+1}}{\lambda\varphi_\lambda - 1}. \tag{2.23}$$

Combining (2.18), (2.21) and (2.23) and using (2.6) and (2.11) we deduce that

$$\mathscr{S}_1 \leqslant 1.99(1+\epsilon)\sqrt{\lambda\pi}e^{\frac{\lambda}{2}}. \tag{2.24}$$

However, for N large enough we have

$$\sqrt{\lambda\pi}e^{\frac{\lambda}{2}} \leqslant (1+\epsilon)\sqrt{\lambda(1-k\delta t)\pi}e^{\frac{\lambda}{2}(1-k\delta t)} \tag{2.25}$$

and the upper bound on $\mathscr{S}_1$ in (2.17) follows.

Secondly, recall from (1.4) that for all $0 \leqslant j \leqslant N-1$,

$$\tilde{v}^N_{t_{N-j}} = \tilde{v}^N_{t_{N-j-1}} + k\theta\delta t - k\left(\tilde{v}^N_{t_{N-j-1}}\right)^+\delta t + \xi\sqrt{\left(\tilde{v}^N_{t_{N-j-1}}\right)^+}(W_{t_{N-j}} - W_{t_{N-j-1}}), \tag{2.26}$$

and note that

$$\mathbb{P}\left(\tilde{v}^N_{t_{N-j}} \leqslant 0\right) = \mathbb{P}\left(\tilde{v}^N_{t_{N-j-1}} \leqslant -k\theta\delta t\right) + \mathbb{P}\left(\tilde{v}^N_{t_{N-j}} \leqslant 0, \tilde{v}^N_{t_{N-j-1}} > 0\right). \tag{2.27}$$

Let $(\mathcal{F}^v_t)_{t\geqslant 0}$ be the natural filtration generated by W and consider the shorthand notations $\mathbb{E}_t[\,\cdot\,] = \mathbb{E}\left[\,\cdot\,|\mathcal{F}^v_t\right]$ and $\mathbb{P}_t(\cdot) = \mathbb{P}\left(\cdot\,|\mathcal{F}^v_t\right)$ for the conditional expectation and probability. Conditioning on $\mathcal{F}^v_{t_{N-j-1}}$ we get

$$\begin{aligned}\mathbb{P}\left(\tilde{v}^N_{t_{N-j}} \leqslant 0, \tilde{v}^N_{t_{N-j-1}} > 0\right) &= \mathbb{E}\left[\mathbb{1}_{\tilde{v}^N_{t_{N-j-1}}>0}\,\mathbb{E}_{t_{N-j-1}}\left[\mathbb{1}_{\tilde{v}^N_{t_{N-j}}\leqslant 0}\right]\right]\\ &= \mathbb{E}\left[\mathbb{1}_{w>0}\,\mathbb{P}_{t_{N-j-1}}\left(Z \leqslant -\frac{k\theta\delta t + w^+(1-k\delta t)}{\xi\sqrt{w^+\delta t}}\right)\right],\end{aligned} \tag{2.28}$$

where $w = \tilde{v}^N_{t_{N-j-1}}$ and $Z \sim \mathcal{N}(0,1)$ independent of $\mathcal{F}^v_{t_{N-j-1}}$. Using a standard inequality for the lower tail of the normal distribution, namely

$$\mathbb{P}\left(Z \leqslant -x\right) \leqslant \frac{1}{\sqrt{2\pi}x}e^{-\frac{1}{2}x^2}, \forall x > 0, \tag{2.29}$$

and the arithmetic mean–geometric mean (AM-GM) inequality we deduce that

$$\mathbb{P}\left(\tilde{v}^N_{t_{N-j}} \leqslant 0, \tilde{v}^N_{t_{N-j-1}} > 0\right) \leqslant \frac{1}{2\sqrt{\lambda(1-k\delta t)\pi}}\,\mathbb{E}\left[\mathbb{1}_{w>0}\exp\left\{-\frac{(k\theta\delta t + w^+(1-k\delta t))^2}{2\xi^2 w^+\delta t}\right\}\right]$$
$$\leqslant \frac{e^{-\frac{\lambda}{2}(1-k\delta t)}}{2\sqrt{\lambda(1-k\delta t)\pi}}\,\mathbb{E}\left[\exp\left\{-a_1 \operatorname{abs}_\infty(\tilde{v}^N_{t_{N-j-1}})\right\}\right], \tag{2.30}$$

where $\operatorname{abs}_\infty(w) = w$ if $w > 0$ and $\operatorname{abs}_\infty(w) = \infty$ otherwise. Let $1 \leqslant i,j \leqslant N-1$ then

$$\mathbb{E}\left[\exp\left\{-a_i \operatorname{abs}_\infty\left(\tilde{v}^N_{t_{N-j}}\right)\right\}\right] = \mathbb{E}\left[\exp\left\{-a_i \operatorname{abs}_\infty\left(\tilde{v}^N_{t_{N-j}}\right)\right\}\left(\mathbb{1}_{\tilde{v}^N_{t_{N-j-1}}>0} + \mathbb{1}_{\tilde{v}^N_{t_{N-j-1}}\leqslant 0}\right)\right]$$
$$\leqslant \mathbb{E}\left[\mathbb{1}_{\tilde{v}^N_{t_{N-j-1}}>0}\,\mathbb{E}_{t_{N-j-1}}\left[\exp\left\{-a_i \operatorname{abs}_\infty\left(\tilde{v}^N_{t_{N-j}}\right)\right\}\right]\right]$$
$$+\,\mathbb{P}\left(-k\theta\delta t < \tilde{v}^N_{t_{N-j-1}} \leqslant 0\right). \tag{2.31}$$

Denote $w = \tilde{v}^N_{t_{N-j-1}}$ as before and let $\mathscr{I}$ be the inner expectation on the right-hand side of (2.31), i.e.,

$$\mathscr{I} = \mathbb{E}_{t_{N-j-1}}\left[\exp\left\{-a_i \operatorname{abs}_\infty\left(w + k\theta\delta t - kw^+\delta t + \xi\sqrt{w^+\delta t}Z\right)\right\}\right], \tag{2.32}$$

where $Z \sim \mathcal{N}(0,1)$ independent of $\mathscr{F}^v_{t_{N-j-1}}$. There are two possible outcomes, namely $w \leqslant 0$, in which case $\mathscr{I} \leqslant 1$, and $w > 0$, which is treated now:

$$\mathscr{I} = \int_{z_0}^{\infty}\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{1}{2}z^2 - a_i\left(k\theta\delta t + w(1-k\delta t) + \xi\sqrt{w\delta t}z\right)\right\}\mathrm{d}z$$
$$= \exp\left\{-a_i k\theta\delta t - wa_i(1-k\delta t) + \frac{1}{2}wa_i^2\xi^2\delta t\right\}\Phi\left(-z_0 - a_i\xi\sqrt{w\delta t}\right), \tag{2.33}$$

where

$$z_0 = -\frac{k\theta\delta t + w(1-k\delta t)}{\xi\sqrt{w\delta t}} \tag{2.34}$$

and Φ is the standard normal cumulative distribution function (CDF). We deduce from (2.16) that

$$\mathscr{I} \leqslant \exp\left\{-a_i k\theta\delta t - wa_{i+1}\right\}, \tag{2.35}$$

and hence, that

$$\mathbb{E}\left[\exp\left\{-a_i \operatorname{abs}_\infty\left(\tilde{v}^N_{t_{N-j}}\right)\right\}\right] \leqslant \exp\left\{-a_i k\theta\delta t\right\}\mathbb{E}\left[\exp\left\{-a_{i+1} \operatorname{abs}_\infty\left(\tilde{v}^N_{t_{N-j-1}}\right)\right\}\right]$$
$$+\,\mathbb{P}\left(-k\theta\delta t < \tilde{v}^N_{t_{N-j-1}} \leqslant 0\right). \tag{2.36}$$

For N large enough, putting together (2.17), (2.27), (2.30) and (2.36), we can prove by induction that, for all $0 \leqslant l \leqslant N-1$,

$$\sup_{0\leqslant n\leqslant l} \mathbb{P}\left(\tilde{v}^N_{t_{N-n}} \leqslant 0\right) \leqslant \frac{e^{-\frac{\lambda}{2}(1-k\delta t)}}{2\sqrt{\lambda(1-k\delta t)\pi}} \sum_{n=0}^{l} \mathbb{E}\left[\exp\left\{-a_{n+1}\,\mathrm{abs}_\infty\left(\tilde{v}^N_{t_{N-l-1}}\right)\right\}\right] \prod_{j=0}^{n} \exp\left\{-a_j k\theta\delta t\right\} + \mathbb{P}\left(\tilde{v}^N_{t_{N-l-1}} \leqslant 0\right). \tag{2.37}$$

Taking $l = N-1$ in (2.37) and since $\tilde{v}^N_{t_0} = v_0 > 0$ we obtain

$$\sup_{0\leqslant n\leqslant N} \mathbb{P}\left(\tilde{v}^N_{t_n} \leqslant 0\right) \leqslant \frac{e^{-\frac{\lambda}{2}(1-k\delta t)}}{2\sqrt{\lambda(1-k\delta t)\pi}} \sum_{n=0}^{N-1} \exp\left\{-a_{n+1}v_0\right\} \prod_{j=0}^{n} \exp\left\{-a_j k\theta\delta t\right\}. \tag{2.38}$$

Using (2.10) yields

$$\begin{aligned}
\mathscr{S}_2 &= \sum_{n=0}^{N-1} \exp\left\{-a_{n+1}v_0\right\} \prod_{j=0}^{n} \exp\left\{-a_j k\theta\delta t\right\} \\
&\leqslant \sum_{n=0}^{N-1} \exp\left\{\lambda \sum_{j=1}^{n}\left(1 - 2\alpha_N - \frac{\varphi_\lambda}{j-1+\eta_\lambda}\right) + \frac{2v_0}{\xi^2\delta t}\left(1 - 2\alpha_N - \frac{\varphi_\lambda}{n+\eta_\lambda}\right)\right\} \\
&\leqslant \eta_\lambda^{\lambda\varphi_\lambda} \exp\left\{\lambda\left(\frac{v_0}{\theta} + kT\right)\right\} \sum_{n=0}^{N-1} (\eta_\lambda + n)^{-\lambda\varphi_\lambda} \exp\left\{-\frac{2v_0\varphi_\lambda N}{\xi^2 T(n+\eta_\lambda)}\right\}.
\end{aligned} \tag{2.39}$$

However, for all $x, y > 0$ we have that $\log(y/x) \geqslant 1 - x/y$ such that $e^{-x} \leqslant e^{-y}(y/x)^y$ and hence,

$$\begin{aligned}
\mathscr{S}_2 &\leqslant \eta_\lambda^{\lambda\varphi_\lambda} \exp\left\{\lambda\left(\frac{v_0}{\theta} + kT\right)\right\} \sum_{n=0}^{N-1} e^{-\lambda\varphi_\lambda} \left(\frac{k\theta T}{v_0}\right)^{\lambda\varphi_\lambda} N^{-\lambda\varphi_\lambda} \\
&= \left(\frac{k\theta T\eta_\lambda}{v_0}\right)^{\lambda\varphi_\lambda} \exp\left\{\lambda\left(\frac{v_0}{\theta} + kT - \varphi_\lambda\right)\right\} N^{-\lambda+\bar{\lambda}+1}.
\end{aligned} \tag{2.40}$$

Combining (2.25), (2.38) and (2.40) we deduce that

$$\sup_{0\leqslant n\leqslant N} \mathbb{P}\left(\tilde{v}^N_{t_n} \leqslant 0\right) \leqslant \frac{(1+\epsilon)e^{-\frac{\lambda}{2}}}{2\sqrt{\lambda\pi}} \left(\frac{k\theta T\eta_\lambda}{v_0}\right)^{\lambda\varphi_\lambda} \exp\left\{\lambda\left(\frac{v_0}{\theta} + kT - \varphi_\lambda\right)\right\} N^{-\lambda+\bar{\lambda}+1}, \tag{2.41}$$

whence the conclusion. □

Next we bound the L^p difference between the two continuous-time approximations $\tilde{v}^N$ and $\bar{v}^N$.

Proposition 2.4 Suppose that $\lambda > 2$ and let $1 \leqslant p < 2(\lambda - \bar{\lambda} - 1)$. Then there exists a constant $C > 0$ such that, for all $N \geqslant 1$,

$$\sup_{t\in[0,T]} \left(\mathbb{E}\left[|\tilde{v}_t^N - \bar{v}_t^N|^p\right]\right)^{\frac{1}{p}} \leqslant CN^{-\frac{1}{2}}. \tag{2.42}$$

Proof. Let $\epsilon = 2(\lambda - \bar{\lambda} - 1) - p > 0$. For convenience of notation define

$$\Delta\tilde{v}_t^N = \tilde{v}_t^N - \tilde{v}_{\bar{t}}^N, \tag{2.43}$$

where $\bar{t} = \delta t \lfloor \frac{t}{\delta t} \rfloor$ for all $t \in [0, T]$. From the triangle inequality we have

$$|\tilde{v}_t^N - \bar{v}_t^N| \leqslant \left|\tilde{v}_{\bar{t}}^N - \left(\tilde{v}_{\bar{t}}^N\right)^+\right| + |\Delta\tilde{v}_t^N| = |\tilde{v}_{\bar{t}}^N|\,\mathbb{1}_{\tilde{v}_{\bar{t}}^N \leqslant 0} + |\Delta\tilde{v}_t^N|, \tag{2.44}$$

and hence, using Hölder's inequality, we get

$$\begin{aligned}\sup_{t\in[0,T]} \mathbb{E}\left[|\tilde{v}_t^N - \bar{v}_t^N|^p\right] \leqslant & 2^{p-1} \sup_{N\geqslant 1}\sup_{0\leqslant n\leqslant N} \mathbb{E}\left[|\tilde{v}_{t_n}^N|^{\frac{p(p+\epsilon)}{\epsilon}}\right]^{\frac{\epsilon}{p+\epsilon}} \sup_{0\leqslant n\leqslant N} \mathbb{P}\left(\tilde{v}_{t_n}^N \leqslant 0\right)^{\frac{p}{p+\epsilon}} \\ & + 2^{p-1}\sup_{t\in[0,T]} \mathbb{E}\left[|\Delta\tilde{v}_t^N|^p\right].\end{aligned} \tag{2.45}$$

We can bound the last term on the right-hand side from above as follows:

$$\begin{aligned}|\Delta\tilde{v}_t^N|^p & \leqslant \left(k\theta\delta t + k\tilde{v}_t^N\delta t + \xi\sqrt{\bar{v}_t^N}\left|W_t - W_{\bar{t}}\right|\right)^p \\ & \leqslant 3^{p-1}k^p\theta^p(\delta t)^p + 3^{p-1}k^p|\tilde{v}_{\bar{t}}^N|^p(\delta t)^p + 3^{p-1}\xi^p|\tilde{v}_{\bar{t}}^N|^{\frac{p}{2}}\left|W_t - W_{\bar{t}}\right|^p,\end{aligned} \tag{2.46}$$

and hence,

$$\begin{aligned}\sup_{t\in[0,T]} \mathbb{E}\left[|\Delta\tilde{v}_t^N|^p\right] \leqslant & 3^{p-1}(k\theta T)^p N^{-p} + 3^{p-1}(kT)^p \sup_{N\geqslant 1}\sup_{0\leqslant n\leqslant N} \mathbb{E}\left[|\tilde{v}_{t_n}^N|^p\right]N^{-p} \\ & + 3^{p-1}\xi^p T^{\frac{p}{2}}\,\mathbb{E}\left[|Z|^p\right] \sup_{N\geqslant 1}\sup_{0\leqslant n\leqslant N} \mathbb{E}\left[|\tilde{v}_{t_n}^N|^{\frac{p}{2}}\right]N^{-\frac{p}{2}},\end{aligned} \tag{2.47}$$

where $Z \sim \mathcal{N}(0, 1)$. Substituting back into (2.45) with (2.47) and using Lemma 2.2 and Proposition 2.3 concludes the proof. □

Before we prove the main result of this paper we need the following technical auxiliary result.

Lemma 2.5 Suppose that $2 \leqslant q < \lambda - 1$. Then we can find $\beta > 0$ such that

$$\lambda > 2\beta + 1 > \lambda + 2q - \sqrt{(\lambda + 2q - 1)^2 - 4q(q-1)} \tag{2.48}$$

and

$$2(\lambda - \bar{\lambda} - 1)(\lambda - \beta - 1) > \lambda q. \tag{2.49}$$

Proof. Note that (2.48) and (2.49) are equivalent to

$$(\lambda q) \vee (\lambda - 1)(\lambda - \bar{\lambda} - 1) < (\lambda - \bar{\lambda} - 1)\left(\lambda - 2q - 1 + \sqrt{(\lambda + 2q - 1)^2 - 4q(q-1)}\right) \tag{2.50}$$

since $\beta \in (0, \infty)$ spans the interval of values associated with (2.50) for

$$2(\lambda - \bar{\lambda} - 1)(\lambda - \beta - 1). \tag{2.51}$$

Furthermore, one can easily check that

$$2q < \sqrt{(\lambda + 2q - 1)^2 - 4q(q-1)}, \tag{2.52}$$

which implies that

$$1 + \bar{\lambda} < \lambda - \frac{\lambda q}{\lambda - 2q - 1 + \sqrt{(\lambda + 2q - 1)^2 - 4q(q-1)}} \tag{2.53}$$

is equivalent to (2.50). However, we can rewrite (2.53) as

$$1 + \bar{\lambda} < \lambda\left(1 - \frac{2q + 1 - \lambda + \sqrt{(\lambda + 2q - 1)^2 - 4q(q-1)}}{4(2\lambda - q - 1)}\right). \tag{2.54}$$

Let $x = \lambda - q - 1 > 0$ and define the right-hand side function

$$\begin{aligned} f_q(x) &= (q + 1 + x)\left(1 - \frac{4q - (3q + x) + \sqrt{(3q + x)^2 - 4q(q-1)}}{4(q + 1 + 2x)}\right) \\ &= 1 + \frac{(2q + 1 + 2x)x}{q + 1 + 2x} + \frac{(2q + 2 + 2x)q(q-1)}{2(q + 1 + 2x)(3q + x + \sqrt{(3q + x)^2 - 4q(q-1)})}. \end{aligned} \tag{2.55}$$

Hence, we deduce that

$$f_q(x) \geqslant 1 + x + \frac{q(q-1)}{4(3q + x)} \geqslant 1 + \frac{q-1}{12} \geqslant \frac{13}{12} > 1 + \bar{\lambda}|_{\lambda=3} \geqslant 1 + \bar{\lambda}, \tag{2.56}$$

which concludes the proof. □

With these results at our disposal we are now ready to prove the main theorem.

Proof. (of Theorem 1.1) Let $p < q < \lambda - 1$ and fix any $\beta > 0$ which satisfies (2.48) and (2.49). For convenience of notation, define

$$e_t^v = v_t - \tilde{v}_t^N, e_0^v = 0, \tag{2.57}$$

such that

$$de_t^v = -k\left(v_t - \tilde{v}_t^N\right) \mathrm{d}t + \xi\left(\sqrt{v_t} - \sqrt{\tilde{v}_t^N}\right) \mathrm{d}W_t. \tag{2.58}$$

Since $\lambda \geqslant 1$ the CIR process v has almost surely strictly positive paths, applying Itô's formula to the $\mathcal{C}^{2,2}$ function $f(v_t, e_t^v) = v_t^{-\beta}|e_t^v|^q$ yields

$$v_t^{-\beta}|e_t^v|^q = -\beta \int_0^t v_u^{-(\beta+1)}|e_u^v|^q \,\mathrm{d}v_u + q\int_0^t v_u^{-\beta}|e_u^v|^{q-1}\operatorname{sgn}(e_u^v)\,\mathrm{d}e_u^v + \frac{1}{2}\beta(\beta+1)\int_0^t v_u^{-(\beta+2)}|e_u^v|^q \,\mathrm{d}\langle v\rangle_u$$
$$+ \frac{1}{2}q(q-1)\int_0^t v_u^{-\beta}|e_u^v|^{q-2}\,\mathrm{d}\langle e^v\rangle_u - \beta q\int_0^t v_u^{-(\beta+1)}|e_u^v|^{q-1}\operatorname{sgn}(e_u^v)\,\mathrm{d}\langle v, e^v\rangle_u, \tag{2.59}$$

where $\operatorname{sgn}(e^v) = 1$ if $e^v > 0$ and $\operatorname{sgn}(e^v) = -1$ otherwise and hence,

$$v_t^{-\beta}|e_t^v|^q = -\beta k\theta \int_0^t v_u^{-(\beta+1)}|e_u^v|^q \,\mathrm{d}u + \beta k\int_0^t v_u^{-\beta}|e_u^v|^q \,\mathrm{d}u - \beta\xi\int_0^t v_u^{-\left(\beta+\frac{1}{2}\right)}|e_u^v|^q \,\mathrm{d}W_u$$
$$- qk\int_0^t v_u^{-\beta}|e_u^v|^{q-1}\operatorname{sgn}(e_u^v)(v_u - \bar{v}_u^N)\,\mathrm{d}u + q\xi\int_0^t v_u^{-\beta}|e_u^v|^{q-1}\operatorname{sgn}(e_u^v)\big(\sqrt{v_u} - \sqrt{\bar{v}_u^N}\big)\,\mathrm{d}W_u$$
$$+ \frac{1}{2}\beta(\beta+1)\xi^2\int_0^t v_u^{-(\beta+1)}|e_u^v|^q \,\mathrm{d}u + \frac{1}{2}q(q-1)\xi^2\int_0^t v_u^{-\beta}|e_u^v|^{q-2}\big|\sqrt{v_u} - \sqrt{\bar{v}_u^N}\big|^2 \,\mathrm{d}u$$
$$- \beta q\xi^2\int_0^t v_u^{-\left(\beta+\frac{1}{2}\right)}|e_u^v|^{q-1}\operatorname{sgn}(e_u^v)\big(\sqrt{v_u} - \sqrt{\bar{v}_u^N}\big)\,\mathrm{d}u. \tag{2.60}$$

We can show that the two stochastic integrals in (2.60) are true martingales by a simple application of Hölder's inequality and Lemmas 2.1, 2.2 and 2.5. Taking expectations on both sides, since

$$v_u - \bar{v}_u^N = e_u^v + \tilde{v}_u^N - \bar{v}_u^N \text{ and } \operatorname{sgn}(e_u^v)e_u^v = |e_u^v|, \tag{2.61}$$

we deduce that

$$\mathbb{E}\left[v_t^{-\beta}|e_t^v|^q\right] = (\beta - q)k\,\mathbb{E}\left[\int_0^t v_u^{-\beta}|e_u^v|^q \,\mathrm{d}u\right] - qk\,\mathbb{E}\left[\int_0^t v_u^{-\beta}|e_u^v|^{q-1}\operatorname{sgn}(e_u^v)\left(\tilde{v}_u^N - \bar{v}_u^N\right)\mathrm{d}u\right]$$
$$- \beta q\xi^2\,\mathbb{E}\left[\int_0^t v_u^{-\left(\beta+\frac{1}{2}\right)}|e_u^v|^{q-2}e_u^v\left(\sqrt{v_u} - \sqrt{\bar{v}_u^N}\right)\mathrm{d}u\right]$$
$$+ \frac{1}{2}q(q-1)\xi^2\,\mathbb{E}\left[\int_0^t v_u^{-\beta}|e_u^v|^{q-2}\big|\sqrt{v_u} - \sqrt{\bar{v}_u^N}\big|^2 \,\mathrm{d}u\right]$$
$$- \frac{1}{2}\beta\xi^2(\lambda - \beta - 1)\,\mathbb{E}\left[\int_0^t v_u^{-(\beta+1)}|e_u^v|^q \,\mathrm{d}u\right]. \tag{2.62}$$

However, note that

$$\begin{aligned}\sqrt{v_u}e_u^v\left(\sqrt{v_u}-\sqrt{\bar{v}_u^N}\right) &= \sqrt{v_u}\left(v_u-\bar{v}_u^N+\bar{v}_u^N-\tilde{v}_u^N\right)\left(\sqrt{v_u}-\sqrt{\bar{v}_u^N}\right)\\ &\geqslant v_u\left|\sqrt{v_u}-\sqrt{\bar{v}_u^N}\right|^2-\sqrt{v_u}\left|\sqrt{v_u}-\sqrt{\bar{v}_u^N}\right||\tilde{v}_u^N-\bar{v}_u^N|\\ &\geqslant v_u\left|\sqrt{v_u}-\sqrt{\bar{v}_u^N}\right|^2-|e_u^v||\tilde{v}_u^N-\bar{v}_u^N|-|\tilde{v}_u^N-\bar{v}_u^N|^2. \end{aligned} \tag{2.63}$$

Substituting back into (2.62) with (2.63) leads to

$$\begin{aligned}\mathbb{E}\left[v_t^{-\beta}|e_t^v|^q\right] \leqslant{}& (\beta-q)^+k\,\mathbb{E}\left[\int_0^t v_u^{-\beta}|e_u^v|^q\,\mathrm{d}u\right]+qk\,\mathbb{E}\left[\int_0^t v_u^{-\beta}|e_u^v|^{q-1}|\tilde{v}_u^N-\bar{v}_u^N|\,\mathrm{d}u\right]\\ &+\beta q\xi^2\,\mathbb{E}\left[\int_0^t v_u^{-(\beta+1)}|e_u^v|^{q-1}|\tilde{v}_u^N-\bar{v}_u^N|\,\mathrm{d}u\right]+\beta q\xi^2\,\mathbb{E}\left[\int_0^t v_u^{-(\beta+1)}|e_u^v|^{q-2}|\tilde{v}_u^N-\bar{v}_u^N|^2\,\mathrm{d}u\right]\\ &-\frac{1}{2}q\xi^2(2\beta+1-q)\,\mathbb{E}\left[\int_0^t v_u^{-\beta}|e_u^v|^{q-2}\left|\sqrt{v_u}-\sqrt{\bar{v}_u^N}\right|^2\mathrm{d}u\right]\\ &-\frac{1}{2}\beta\xi^2(\lambda-\beta-1)\,\mathbb{E}\left[\int_0^t v_u^{-(\beta+1)}|e_u^v|^q\,\mathrm{d}u\right]. \end{aligned} \tag{2.64}$$

Let $\eta>0$. For any $a,b\geqslant 0$ and $j\in\{1,2\}$ Young's inequality yields

$$a^{q-j}b^j=\left(\eta^{\frac{j(q-j)}{q}}a^{q-j}\right)\left(\eta^{-\frac{j(q-j)}{q}}b^j\right)\leqslant\frac{q-j}{q}\eta^j a^q+\frac{j}{q}\eta^{j-q}b^q. \tag{2.65}$$

Using (2.65) and after some rearrangements we deduce that

$$\begin{aligned}\mathbb{E}\left[v_t^{-\beta}|e_t^v|^q\right] \leqslant{}& \big((\beta-q)^++\eta(q-1)\big)k\,\mathbb{E}\left[\int_0^t v_u^{-\beta}|e_u^v|^q\,\mathrm{d}u\right]+\eta^{1-q}k\,\mathbb{E}\left[\int_0^t v_u^{-\beta}|\tilde{v}_u^N-\bar{v}_u^N|^q\,\mathrm{d}u\right]\\ &+\eta^{1-q}(1+2\eta)\beta\xi^2\,\mathbb{E}\left[\int_0^t v_u^{-(\beta+1)}|\tilde{v}_u^N-\bar{v}_u^N|^q\,\mathrm{d}u\right]\\ &-\frac{1}{2}q\xi^2(2\beta+1-q)\,\mathbb{E}\left[\int_0^t v_u^{-\beta}|e_u^v|^{q-2}\left|\sqrt{v_u}-\sqrt{\bar{v}_u^N}\right|^2\mathrm{d}u\right]\\ &-\frac{1}{2}\beta\xi^2\big(\lambda-\beta-1-2\eta(q-1)-2\eta^2(q-2)\big)\,\mathbb{E}\left[\int_0^t v_u^{-(\beta+1)}|e_u^v|^q\,\mathrm{d}u\right]. \end{aligned} \tag{2.66}$$

There are two cases to consider. First, if $q\leqslant 2\beta+1$, as we can find $\eta>0$ small enough such that

$$2\eta(q-1)+2\eta^2(q-2)\leqslant\beta \tag{2.67}$$

and since $2\beta + 1 < \lambda$ from Lemma 2.5 we get

$$\mathbb{E}\left[v_t^{-\beta}|e_t^v|^q\right] \leqslant \big((\beta-q)^+ + \eta(q-1)\big)k\,\mathbb{E}\left[\int_0^t v_u^{-\beta}|e_u^v|^q\,\mathrm{d}u\right] + \eta^{1-q}k\,\mathbb{E}\left[\int_0^t v_u^{-\beta}|\tilde{v}_u^N - \bar{v}_u^N|^q\,\mathrm{d}u\right] + \eta^{1-q}(1+2\eta)\beta\xi^2\,\mathbb{E}\left[\int_0^t v_u^{-(\beta+1)}|\tilde{v}_u^N - \bar{v}_u^N|^q\,\mathrm{d}u\right]. \tag{2.68}$$

Secondly, if $q > 2\beta + 1$, using the triangle and AM–GM inequalities yields

$$v_u\big|\sqrt{v_u} - \sqrt{\bar{v}_u^N}\big|^2 \leqslant |v_u - \bar{v}_u^N|^2 \leqslant (1+\eta)|e_u^v|^2 + \eta^{-1}(1+\eta)|\tilde{v}_u^N - \bar{v}_u^N|^2, \tag{2.69}$$

and hence, using Young's inequality, we get

$$qv_u|e_u^v|^{q-2}\big|\sqrt{v_u} - \sqrt{\bar{v}_u^N}\big|^2 \leqslant \big(q + \eta(q-2)\big)(1+\eta)|e_u^v|^q + 2\eta^{1-q}(1+\eta)|\tilde{v}_u^N - \bar{v}_u^N|^q. \tag{2.70}$$

Substituting back into (2.66) with (2.70) leads to

$$\begin{aligned}\mathbb{E}\left[v_t^{-\beta}|e_t^v|^q\right] \leqslant{}& \big((\beta-q)^+ + \eta(q-1)\big)k\,\mathbb{E}\left[\int_0^t v_u^{-\beta}|e_u^v|^q\,\mathrm{d}u\right] + \eta^{1-q}k\,\mathbb{E}\left[\int_0^t v_u^{-\beta}|\tilde{v}_u^N - \bar{v}_u^N|^q\,\mathrm{d}u\right]\\ &+ \eta^{1-q}\xi^2\big((1+2\eta)\beta + (1+\eta)(q-2\beta-1)\big)\,\mathbb{E}\left[\int_0^t v_u^{-(\beta+1)}|\tilde{v}_u^N - \bar{v}_u^N|^q\,\mathrm{d}u\right]\\ &- \frac{1}{2}\xi^2\big(\beta(\lambda-\beta-1) - q(q-2\beta-1) - 2\eta(q-1)(q-\beta-1) - \eta^2(q-1)(q-2)\big)\\ &\times \mathbb{E}\left[\int_0^t v_u^{-(\beta+1)}|e_u^v|^q\,\mathrm{d}u\right].\end{aligned} \tag{2.71}$$

However, note that

$$\beta(\lambda-\beta-1) - q(q-2\beta-1) > 0 \tag{2.72}$$

from Lemma 2.5 and hence, we can find $\eta > 0$ small enough such that

$$2\eta(q-1)(q-\beta-1) + \eta^2(q-1)(q-2) \leqslant \beta(\lambda-\beta-1) - q(q-2\beta-1). \tag{2.73}$$

Going back to (2.71) we deduce that

$$\mathbb{E}\left[v_t^{-\beta}|e_t^v|^q\right] \leqslant \big((\beta-q)^+ + \eta(q-1)\big)k\,\mathbb{E}\left[\int_0^t v_u^{-\beta}|e_u^v|^q\,\mathrm{d}u\right] + \eta^{1-q}k\,\mathbb{E}\left[\int_0^t v_u^{-\beta}|\tilde{v}_u^N - \bar{v}_u^N|^q\,\mathrm{d}u\right] + \eta^{1-q}\xi^2\big((1+2\eta)\beta + (1+\eta)(q-2\beta-1)\big)\,\mathbb{E}\left[\int_0^t v_u^{-(\beta+1)}|\tilde{v}_u^N - \bar{v}_u^N|^q\,\mathrm{d}u\right]. \tag{2.74}$$

Let $0 \leqslant s \leqslant T$. Combining (2.68) and (2.74) taking the supremum over $[0, s]$ and using Fubini's theorem leads to

$$\begin{aligned}\sup_{t\in[0,s]} \mathbb{E}\left[v_t^{-\beta}|e_t^v|^q\right] \leqslant &\big((\beta-q)^+ + \eta(q-1)\big)k\int_0^s \sup_{t\in[0,u]} \mathbb{E}\left[v_t^{-\beta}|e_t^v|^q\right]\mathrm{d}u\\ &+ \eta^{1-q}\xi^2\big((1+2\eta)\beta + (1+\eta)(q-2\beta-1)^+\big)T \sup_{t\in[0,T]} \mathbb{E}\left[v_t^{-(\beta+1)}|\tilde{v}_t^N - \bar{v}_t^N|^q\right]\\ &+ \eta^{1-q}kT \sup_{t\in[0,T]} \mathbb{E}\left[v_t^{-\beta}|\tilde{v}_t^N - \bar{v}_t^N|^q\right]. \end{aligned} \tag{2.75}$$

Note from Lemma 2.5 that we can find $r > 1$ such that

$$\frac{\lambda}{\lambda-\beta-1} < \frac{r}{r-1} < \frac{2(\lambda-\bar{\lambda}-1)}{q}. \tag{2.76}$$

Applying Hölder's inequality yields for $i \in \{0, 1\}$,

$$\sup_{t\in[0,T]} \mathbb{E}\left[v_t^{-(\beta+i)}|\tilde{v}_t^N - \bar{v}_t^N|^q\right] \leqslant \sup_{t\in[0,T]} \mathbb{E}\left[v_t^{-r(\beta+i)}\right]^{\frac{1}{r}} \sup_{t\in[0,T]} \mathbb{E}\left[|\tilde{v}_t^N - \bar{v}_t^N|^{\frac{rq}{r-1}}\right]^{\frac{r-1}{r}}. \tag{2.77}$$

Substituting back into (2.75) with (2.77) noticing that the expectation on the left-hand side of (2.75) is finite and using Gronwall's inequality leads to

$$\begin{aligned}\sup_{t\in[0,T]} \mathbb{E}\left[v_t^{-\beta}|e_t^v|^q\right] \leqslant &\sup_{t\in[0,T]} \mathbb{E}\left[|\tilde{v}_t^N - \bar{v}_t^N|^{\frac{rq}{r-1}}\right]^{\frac{r-1}{r}} e^{((\beta-q)^+ + \eta(q-1))kT}\Bigg\{\eta^{1-q}kT \sup_{t\in[0,T]} \mathbb{E}\left[v_t^{-r\beta}\right]^{\frac{1}{r}}\\ &+ \eta^{1-q}\xi^2\big((1+2\eta)\beta + (1+\eta)(q-2\beta-1)^+\big)T \sup_{t\in[0,T]} \mathbb{E}\left[v_t^{-r(\beta+1)}\right]^{\frac{1}{r}}\Bigg\}. \end{aligned} \tag{2.78}$$

Using Lemma 2.1, Proposition 2.4 and (2.76) we conclude that there exists a constant $C > 0$ such that, for all $N \geqslant 1$,

$$\sup_{t\in[0,T]} \mathbb{E}\left[v_t^{-\beta}|e_t^v|^q\right] \leqslant CN^{-\frac{q}{2}}. \tag{2.79}$$

From Hölder's inequality, Lemma 2.1 and (2.79), we have

$$\sup_{t\in[0,T]} \mathbb{E}\left[|e_t^v|^p\right] \leqslant \sup_{t\in[0,T]} \mathbb{E}\left[v_t^{-\beta}|e_t^v|^q\right]^{\frac{p}{q}} \sup_{t\in[0,T]} \mathbb{E}\left[v_t^{\frac{\beta p}{q-p}}\right]^{1-\frac{p}{q}} \leqslant CN^{-\frac{p}{2}}. \tag{2.80}$$

Finally, upon noticing that

$$|v_t - \bar{v}_t^N| \leqslant |v_t - \tilde{v}_t^N| \leqslant |e_t^v| + |\Delta\tilde{v}_t^N|, \tag{2.81}$$

the conclusion follows from (2.47) and (2.80). □

3. Numerical results

In this section we perform a numerical analysis of the strong convergence of the FTE scheme. Recall that $\bar{v}_T^N$ is the value at time T of the approximation process corresponding to an equidistant discretization with N time steps. We study the L^1 error

$$\varepsilon(N) = \mathbb{E}\left[|v_T - \bar{v}_T^N|\right]. \tag{3.1}$$

This error criterion is weaker than the one in Theorem 1.1 since, for all $p \geqslant 1$,

$$\mathbb{E}\left[|v_T - \bar{v}_T^N|\right] \leqslant \sup_{t\in[0,T]} \left(\mathbb{E}\left[|v_t - \bar{v}_t^N|^p\right]\right)^{\frac{1}{p}}. \tag{3.2}$$

Due to the difficulty in computing this quantity we estimate as proxy the difference between the values of the approximation process corresponding to N time steps ($\bar{v}_T^N$) and $2N$ time steps ($\bar{v}_T^{2N}$) for the same Brownian path and use the fact that for any $\alpha > 0$, assuming that $\bar{v}_T^N$ converges to v_T in L^1 (without a rate),

$$\varepsilon(N) = \mathcal{O}\big(N^{-\alpha}\big) \Leftrightarrow \mathbb{E}\left[|\bar{v}_T^N - \bar{v}_T^{2N}|\right] = \mathcal{O}\big(N^{-\alpha}\big). \tag{3.3}$$

A proof of (3.3) can be found, for instance, in Alfonsi (2005). The data in Fig. 1 suggest an empirical L^1 order of $\lambda \wedge \frac{1}{2}$, which is in line with our theoretical results when $\lambda > 3$. We mention that this L^1 convergence order was demonstrated theoretically (for all $\lambda > 0$) and numerically (for $\lambda = 0.25$) for the truncated Milstein scheme in Hefter & Herzwurm (2018). We now recall the main result in Hefter & Jentzen (2017), which established a lower error bound for all discretization schemes for the CIR process based on equidistant evaluations of the Brownian driver in the case of an accessible boundary point.

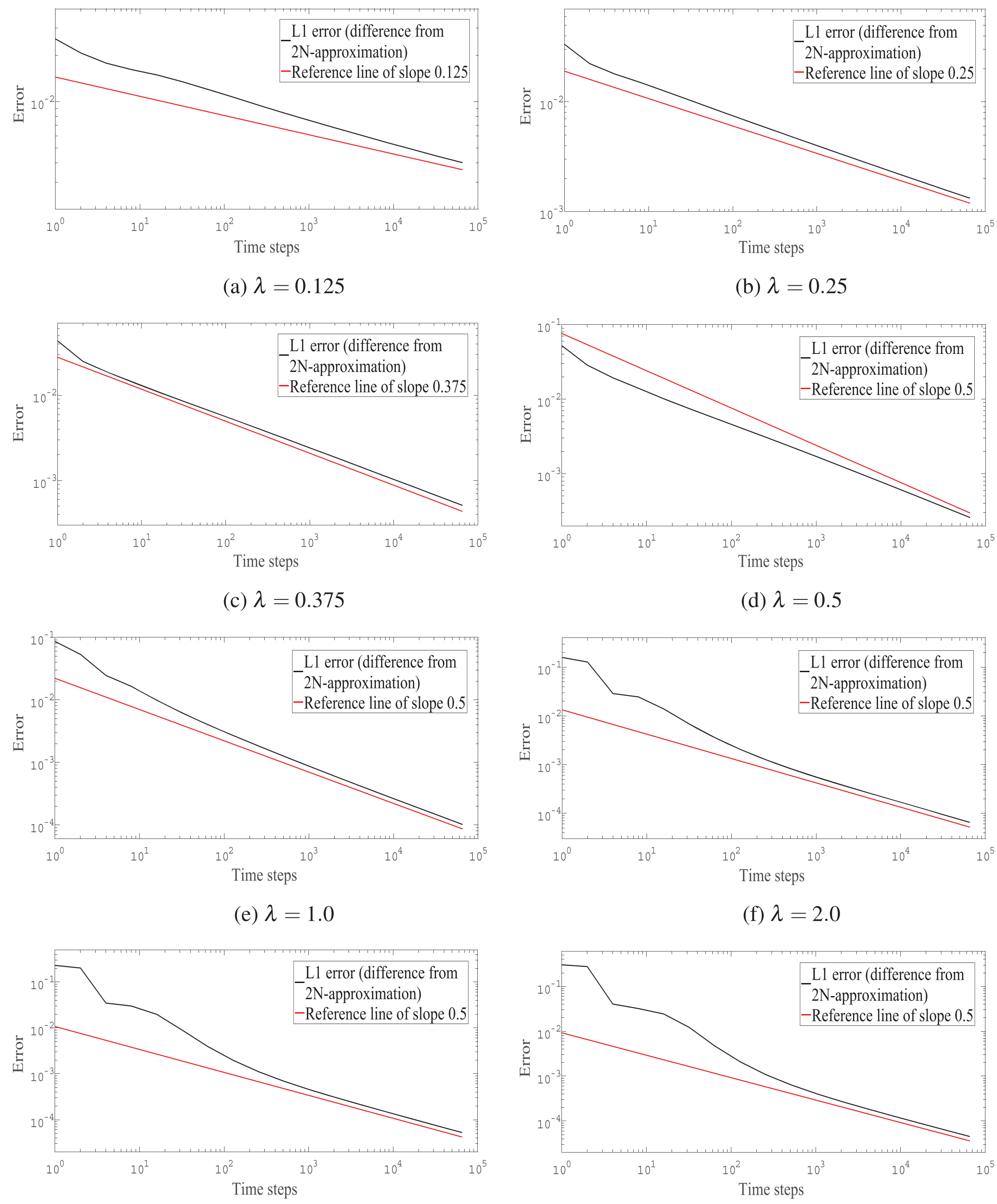
Proposition 3.1 (Theorem 1 in Hefter & Jentzen 2017). If $\lambda < 1$, then discretization methods for the CIR process v based on equidistant evaluations of the driving noise process achieve at most a strong convergence order of λ, i.e., there exists a positive constant C such that, for all $N \geqslant 1$,

$$\inf_{\varphi:\mathbb{R}^N\to\mathbb{R}} \mathbb{E}\left[\left|v_T - \varphi\big(W_{t_1}, W_{t_2}, \ldots, W_{t_N}\big)\right|\right] \geqslant CN^{-\lambda}. \tag{3.4}$$

In particular, the bound in (3.4) suggests an optimal performance—in the L^1 sense—of the FTE scheme in half of the regime where the boundary point is accessible, specifically, when $\lambda \leqslant \frac{1}{2}$.

4. Conclusions

This work has answered questions concerning the convergence order of the FTE scheme for the CIR process. This scheme is often encountered in the mathematical finance literature in the context of Monte Carlo simulations for multi-dimensional models with CIR dynamics in one or more components, like the Heston model. One consequence of this work is that we can establish positive strong convergence rates for approximations of these models (see, e.g., Cozma & Reisinger, 2017). Inevitably, this work also raises some questions, like whether we can relax the condition on the parameters in the inaccessible boundary case without losing the convergence, or whether we can deduce similar properties of the scheme in the accessible boundary case.



(a) $\lambda = 0.125$ (b) $\lambda = 0.25$

(c) $\lambda = 0.375$ (d) $\lambda = 0.5$

(e) $\lambda = 1.0$ (f) $\lambda = 2.0$

(g) $\lambda = 3.0$ (h) $\lambda = 4.0$

FIG. 1. The loglog-plots of the L^1 errors against the number of time steps when $v_0 = 0.02$, $\theta = 0.02$, $k \in \{2.0, 4.0, 6.0, 8.0, 16.0, 32.0, 48.0, 64.0\}$, $\xi = 0.8$ and $T = 1.0$, computed using 2×10^6 Monte Carlo paths (for a relative error less than 10 basis points).

Funding

Engineering and Physical Sciences Research Council (research grant EPSRC EP/N509711/1 to A.C.).

References

Alfonsi, A. (2005) On the discretization schemes for the CIR (and Bessel squared) processes. *Monte Carlo Methods Appl.*, **11**, 355–384.

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

Berkaoui, A., Bossy, M. & Diop, A. (2008) Euler scheme for SDEs with non-Lipschitz diffusion coefficient: strong convergence. *ESAIM: P&S*, **12**, 1–11.

Bossy, M. & Diop, A. (2007) *An efficient discretization scheme for one dimensional SDEs with a diffusion coefficient function of the form* $|x|^\alpha$, $\alpha \epsilon [1/2, 1)$. INRIA Research Report, 5396.

Bossy, M. & Olivero Quinteros, H. (2017) Strong convergence of the symmetrized Milstein scheme for some CEV-like SDEs. *Bernoulli J.*, **24**, 1995–2042.

Broadie, M. & Kaya, O. (2006) Exact simulation of stochastic volatility and other affine jump diffusion processes. *Oper. Res.*, **54**, 217–231.

Chassagneux, J.-F., Jacquier, A. & Mihaylov, I. (2016) An explicit Euler scheme with strong rate of convergence for financial SDEs with non-Lipschitz coefficients. *SIAM J. Financial Math.*, **7**, 993–1021.

Cox, J., Ingersoll, J. & Ross, S. (1985) A theory of the term structure of interest rates. *Econometrica*, **53**, 385–407.

Cozma, A., Mariapragassam, M. & Reisinger, C. (2018) Convergence of an Euler scheme for a hybrid stochastic-local volatility model with stochastic rates in foreign exchange markets. *SIAM J. Financial Math.*, **9**, 127–170.

Cozma, A. & Reisinger, C. (2017) Strong convergence rates for Euler approximations to a class of stochastic path-dependent volatility models. *SIAM J. Numer. Anal., forthcoming*.

Deelstra, G. & Delbaen, F. (1998) Convergence of discretized stochastic (interest rate) processes with stochastic drift term. *Appl. Stochastic Models and Data Anal.*, **14**, 77–84.

Dereich, S., Neuenkirch, A. & Szpruch, L. (2012) An Euler-type method for the strong approximation of the Cox–Ingersoll–Ross process. *Proceedings of the Royal Society of London A,* **468**, 1105–1115.

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

Giles, M. B., Higham, D. J. & Mao, X. (2009) Analysing multi-level Monte Carlo for options with non-globally Lipschitz payoff. *Finance and Stochastics*, **13**, 403–413.

Glasserman, P. (2004) *Monte Carlo Methods in Financial Engineering. Stochastic Modelling and Applied Probability, vol. 53*. New York: Springer.

Gyöngy, I. & Rásonyi, M. (2011) A note on Euler approximations for SDEs with Hölder continuous diffusion coefficients. *Stochastic Process. Appl.*, **121**, 2189–2200.

Hefter, M., Herzwurm, A. & Müller-Gronbach, T. (2017) Lower error bounds for strong approximation of scalar SDEs with non-Lipschitzian coefficients. *Working paper, arXiv*: 1710.08707.

Hefter, M. & Herzwurm, A. (2017) Optimal strong approximation of the one-dimensional squared Bessel process. *Comm. Math. Sci.*, **15**, 2121–2141.

Hefter, M. & Herzwurm, A. (2018) Strong convergence rates for Cox–Ingersoll–Ross processes – Full parameter range. *J. Math. Anal. Appl.*, **459**, 1079–1101.

Hefter, M. & Jentzen, A. (2017) On arbitrarily slow convergence rates for strong numerical approximations of Cox–Ingersoll–Ross processes and squared Bessel processes. *Financ. Stoch., forthcoming*.

Heinrich, S. (1998) Monte Carlo complexity of global solution of integral equations. *J. Complexity*, **14**, 151–175.

Heston, S. (1993) A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Rev. Financial Stud.*, **6**, 327–343.

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

HIGHAM, D. J. & MAO, X. (2005) Convergence of Monte Carlo simulations involving the mean-reverting square root process. *J. Comput. Finance*, **8**, 35–62.

HURD, T. R. & KUZNETSOV, A. (2008) Explicit formulas for Laplace transforms of stochastic integrals. *Markov Process Related Fields*, **14**, 277–290.

HUTZENTHALER, M., JENTZEN, A. & NOLL, M. (2014) Strong convergence rates and temporal regularity for Cox–Ingersoll–Ross processes and Bessel processes with accessible boundaries. *Working paper, arXiv*: 1403.6385.

KARATZAS, I. & SHREVE, S. E. (1991) *Brownian Motion and Stochastic Calculus, 2nd edn*. New York: Springer.

KEBAIER, A. (2005) Statistical Romberg extrapolation: a new variance reduction method and applications to option pricing. *Ann. Appl. Probab.*, **15**, 2681–2705.

KLOEDEN, P. E. & PLATEN, E. (1995) *Numerical Solution of Stochastic Differential Equations, 2nd edn*. Berlin: Springer.

KLOEDEN, P. & NEUENKIRCH, A. (2007) The pathwise convergence of approximation schemes for stochastic differential equations. *LMSJ. Comput. Math.*, **10**, 235–253.

LORD, R., KOEKKOEK, R. & VAN DIJK, D. (2010) A comparison of biased simulation schemes for stochastic volatility models. *Quant. Financ.*, **10**, 177–194.

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