

# An Euler-type method for the strong approximation of the Cox–Ingersoll–Ross process

BY STEFFEN DEREICH<sup>1</sup>, ANDREAS NEUENKIRCH<sup>2,\*</sup> AND LUKASZ SZPRUCH<sup>3</sup>

<sup>1</sup>*Fachbereich Mathematik und Informatik, Philipps-Universität Marburg,  
Hans-Meerwein-Strasse, 35032 Marburg, Germany*

<sup>2</sup>*Fachbereich Mathematik, TU Kaiserslautern, Postfach 3049,  
67663 Kaiserslautern, Germany*

<sup>3</sup>*Mathematical Institute, University of Oxford, 24–29 St Giles’,  
Oxford OX1 3LB, UK*

We analyse the strong approximation of the Cox–Ingersoll–Ross (CIR) process in the regime where the process does not hit zero by a positivity preserving drift-implicit Euler-type method. As an error criterion, we use the  $p$ th mean of the maximum distance between the CIR process and its approximation on a finite time interval. We show that under mild assumptions on the parameters of the CIR process, the proposed method attains, up to a logarithmic term, the convergence of order  $1/2$ . This agrees with the standard rate of the strong convergence for global approximations of stochastic differential equations with Lipschitz coefficients, despite the fact that the CIR process has a non-Lipschitz diffusion coefficient.

**Keywords:** Cox–Ingersoll–Ross process; strong global approximation; square-root coefficient; Lamperti transformation; drift-implicit method; one-sided Lipschitz condition

## 1. Introduction and main result

In computational finance, a lot of effort has been given to the so-called Cox–Ingersoll–Ross (CIR) process recently. The CIR process under consideration has the following form:

$$dX_t = \kappa(\lambda - X_t) dt + \theta\sqrt{X_t} dW_t, \quad X_0 = x_0, \quad t \geq 0. \quad (1.1)$$

Here,  $W = (W_t)_{t \geq 0}$  is a one-dimensional Brownian motion,  $\kappa, \lambda \geq 0$ ,  $\theta > 0$  and  $x_0 > 0$ . It is well known that equation (1.1) admits a unique strong solution that is non-negative (Karatzas & Shreve 1991, ch. 5). The CIR process was originally proposed by Cox *et al.* (1985) as a model for short-term interest rates. Nowadays, this model is widely used in financial modelling, e.g. as the volatility process in the Heston model (Heston 1993).

\*Author for correspondence ([neuenkirch@mathematik.uni-kl.de](mailto:neuenkirch@mathematik.uni-kl.de)).

One of the main objectives in mathematical finance is the pricing of (path-dependent) derivatives. If the asset prices or the interest rates dynamics are modelled by a  $d$ -dimensional stochastic differential equation (SDE) with solution  $(S_t)_{t \in [0, T]}$ , then this corresponds to the quadrature problem

$$p = \mathbf{E}F(S),$$

where  $F: C([0, T], \mathbb{R}^d) \rightarrow \mathbb{R}$  is the discounted payoff of the derivative. Typically, explicit formulae for such quantities are unknown and have to be approximated by Monte Carlo methods that are based on approximate solutions of the SDE on  $[0, T]$ . Here, the knowledge of the global strong convergence rate of the approximation is important, in particular, if the efficient multi-level Monte Carlo method (Giles 2007, 2008) is to be used.

The strong global approximation of equation (1.1) has been studied in several articles. Strong convergence (without a rate or with a logarithmic rate) of several discretization schemes has been shown by Deelstra & Delbaen (1998), Alfonsi (2005), Higham & Mao (2005), Bossy & Diop (2007) and Gyöngy & Rásonyi (2011). In Alfonsi (2005), a general framework for the analysis of strong approximation of the CIR process was presented, along with extensive simulation studies. Moreover, the strong approximation of more general Ait-Sahalia-type interest-rate models was analysed in Higham *et al.* (2011). However, only in Berkaoui *et al.* (2008) were non-logarithmic convergence rates obtained. There, it was shown that a symmetrized Euler method had strong convergence order  $1/2$  under restrictive assumptions on the parameters of the equation (see §2). The difficulties in obtaining strong convergence rates for the approximations of the CIR process are due to its square-root coefficient. Thus, the standard theory that relies on the global Lipschitz assumption does not apply (Milstein 1995; Kloeden & Platen 1999).

In this article, we focus on the regime where the CIR process does not hit zero, i.e. where

$$\mathbf{P}(X_t > 0 \text{ for } t \geq 0) = 1,$$

an assumption that is often fulfilled in interest-rate models. By the Feller test, it is true if and only if  $2\kappa\lambda \geq \theta^2$  (Karatzas & Shreve 1991, ch. 5). We will not directly do a numerical analysis for the CIR process, but for a coordinate transformation thereof. We consider the process  $Y_t = \sqrt{X_t}$ , which, by Itô's formula, satisfies

$$dY_t = \frac{\alpha}{Y_t} dt + \beta Y_t dt + \gamma dW_t, \quad t \geq 0, \quad Y_0 = \sqrt{x_0}, \quad (1.2)$$

with

$$\alpha = \frac{4\kappa\lambda - \theta^2}{8}, \quad \beta = -\frac{\kappa}{2} \quad \text{and} \quad \gamma = \frac{\theta}{2}.$$

This transformation, known as the Lamperti transformation, allows us to shift the nonlinearity from the diffusion coefficient into the drift coefficient. Note that the drift

$$f(x) = \frac{\alpha}{x} + \beta x, \quad x > 0,$$

satisfies for  $\alpha > 0$ ,  $\beta \in \mathbb{R}$  the one-sided Lipschitz condition

$$(x - y)(f(x) - f(y)) \leq \beta(x - y)^2, \quad x, y > 0.$$

This property is crucial to control the error propagation of drift-implicit Euler schemes (Higham *et al.* 2002, 2011).

The drift-implicit Euler method with stepsize  $\Delta > 0$  for equation (1.2) leads to the numerical scheme

$$y_{k+1} = y_k + \left( \frac{\alpha}{y_{k+1}} + \beta y_{k+1} \right) \Delta + \gamma \Delta_k W, \quad k = 0, 1, \dots, \quad (1.3)$$

with  $y_0 = \sqrt{x_0}$  and

$$\Delta_k W = W_{(k+1)\Delta} - W_{k\Delta}, \quad k = 0, 1, \dots$$

Recalling that  $\alpha, \gamma > 0$  and  $\beta < 0$ , equation (1.3) has the unique positive solution

$$y_{k+1} = \frac{y_k + \gamma \Delta_k W}{2(1 - \beta \Delta)} + \sqrt{\frac{(y_k + \gamma \Delta_k W)^2}{4(1 - \beta \Delta)^2} + \frac{\alpha \Delta}{1 - \beta \Delta}},$$

which we call the *drift-implicit square-root Euler method*. Transforming back, i.e.

$$x_k = y_k^2, \quad k = 0, 1, \dots,$$

gives a strictly positive approximation of the original CIR process. This scheme has already been suggested by Alfonsi (2005), but convergence results have not been established. Using piecewise linear interpolation, i.e.

$$\bar{x}_t = \left( k + 1 - \frac{t}{\Delta} \right) x_k + \left( \frac{t}{\Delta} - k \right) x_{k+1}, \quad t \in [k\Delta, (k+1)\Delta],$$

we obtain a global approximation  $(\bar{x}_t)_{t \in [0, T]}$  of the CIR process on  $[0, T]$ . Exploiting the structure of SDE (1.2), we establish the following theorem.

**Theorem 1.1.** *Let  $2\kappa\lambda > \theta^2$ ,  $x_0 > 0$  and  $T > 0$ . Then, for all*

$$1 \leq p < \frac{2\kappa\lambda}{\theta^2},$$

*there exists a constant  $K_p > 0$  such that*

$$\left( \mathbf{E} \max_{t \in [0, T]} |X_t - \bar{x}_t|^p \right)^{1/p} \leq K_p \sqrt{|\log(\Delta)| \sqrt{\Delta}},$$

*for all  $\Delta \in (0, 1/2]$ .*

Hence, we obtain the optimal strong convergence rate for the approximation of SDEs with Lipschitz coefficients (Müller-Gronbach 2002). The only price to be paid is the restriction on  $p$  that arises from the need to control the inverse  $p$ th moments of the CIR process, which are infinite for  $p \geq 2\kappa\lambda/\theta^2$ .

The remainder of this article is structured as follows: in the next section, we give a short overview on discretization schemes for the CIR process, while the proof of theorem 1.1 is given in §3.

## 2. Numerical methods for the Cox–Ingersoll–Ross process

Discretization schemes for the CIR process have been proposed in numerous articles, including Deelstra & Delbaen (1998), Alfonsi (2005, 2010), Higham & Mao (2005), Jäckel & Kahl (2006), Bossy & Diop (2007), Moro & Schurz (2007), Berkaoui *et al.* (2008), Günther *et al.* (2008) and Gyöngy & Rásonyi (2011). In this short summary, we mainly focus on Euler-type methods.

In Berkaoui *et al.* (2008), the authors studied a symmetrized Euler method for the approximation of equation (1.1), i.e.

$$x_{k+1} = |x_k + \kappa(\lambda - x_k)\Delta + \theta\sqrt{x_k}\Delta_k W|, \quad (2.1)$$

for  $k = 0, 1, \dots$ . Under the assumption

$$\frac{2\kappa\lambda}{\theta^2} > 1 + \sqrt{8} \max \left\{ \frac{\sqrt{\kappa}}{\theta} \sqrt{16p-1}, 16p-2 \right\},$$

they found that

$$\mathbf{E} \max_{k=0, \dots, \lceil T/\Delta \rceil} |X_{k\Delta} - x_k|^{2p} \leq C_p \Delta^p,$$

where the constant  $C_p > 0$  depends only  $p, \kappa, \lambda, \theta, x_0$  and  $T$ . As a consequence of lemma 3.5 in §3, the piecewise linear interpolation of this scheme satisfies the same error estimate with respect to the  $p$ th mean maximum distance as our drift-implicit square-root Euler scheme. However, for the symmetrized Euler scheme, the assumptions on the parameters of the CIR process are clearly more restrictive. Moreover, it was shown in Bossy & Diop (2007) that the symmetrized Euler scheme converged weakly with rate 1 if  $2\kappa\lambda \geq \theta^2$ .

The truncated Euler scheme

$$x_{k+1} = x_k + \kappa(\lambda - x_k)\Delta + \theta\sqrt{x_k^+}\Delta_k W, \quad k = 0, 1, \dots, \quad (2.2)$$

was analysed in Deelstra & Delbaen (1998), while the scheme

$$x_{k+1} = x_k + \kappa(\lambda - x_k)\Delta + \theta\sqrt{|x_k|}\Delta_k W, \quad k = 0, 1, \dots, \quad (2.3)$$

was studied in Higham & Mao (2005). Both schemes do not preserve positivity, but satisfy

$$\mathbf{E} \max_{k=0, \dots, \lceil T/\Delta \rceil} |X_{k\Delta} - x_k|^2 \longrightarrow 0,$$

for  $\Delta \rightarrow 0$  without further restrictions on the parameters of the equation. In particular, an application of Gyöngy & Rásonyi (2011, theorem 3.1) yielded a logarithmic convergence rate for the Euler schemes (2.2) and (2.3). Moreover, if  $2\kappa\lambda \geq \theta^2$ , it follows from Gyöngy (1998) that the Euler schemes (2.1)–(2.3) have a pathwise convergence rate of  $1/2 - \varepsilon$  for all  $\varepsilon > 0$ , i.e. we have

$$\frac{1}{\Delta^{1/2-\varepsilon}} \max_{k=0, \dots, \lceil T/\Delta \rceil} |X_{k\Delta} - x_k| \longrightarrow 0 \quad \mathbf{P} - a.s.,$$

for  $\Delta \rightarrow 0$ . The asymptotic error distribution of these schemes can be deduced from Neuenkirch & Zähle (2009): for  $2\kappa\lambda \geq \theta^2$ , it holds that

$$\frac{1}{\sqrt{\Delta}} \max_{k=0, \dots, \lceil T/\Delta \rceil} |X_{k\Delta} - x_k| \xrightarrow{\mathcal{L}} \max_{t \in [0, T]} \left| \frac{\theta^2}{\sqrt{8}} \Phi_t \int_0^t \frac{1}{\Phi_s} dB_s \right|,$$

for  $\Delta \rightarrow 0$ , where  $(B_t)_{t \geq 0}$  is a Brownian motion independent of  $(W_t)_{t \geq 0}$  and  $(\Phi_t)_{t \geq 0}$  is given by

$$\Phi_t = \exp \left( -\kappa t - \frac{\theta^2}{8} \int_0^t \frac{1}{X_s} ds + \frac{\theta}{2} \int_0^t \frac{1}{\sqrt{X_s}} dW_s \right), \quad t \geq 0.$$

While no explicit solution of equation (1.1) is known, the finite-dimensional distributions can be characterized in terms of a non-central  $\chi^2$  distribution (Cox *et al.* 1985). Thus, equation (1.1) can be simulated exactly at a finite number of time points using the Markov property (Glassermann 2004; Broadie & Kaya 2006).

However, the algorithms for the exact simulation of the CIR process are strongly problem dependent: the number of degrees of freedom of the non-central  $\chi^2$  random variable, which has to be simulated in each step, is  $4\kappa\lambda/\theta^2$ . Thus, the computational cost of the algorithms depends strongly on  $\kappa, \lambda$  and  $\theta$ . The same problem, i.e. strong dependence of the computational cost of the algorithm on the parameters of the equation, arises also for the exact sampling algorithm introduced in Beskos & Roberts (2005), which can be also applied to equation (1.1) (Beskos *et al.* 2006).

While for the simulation of the CIR process at a single point, the exact simulation methods are useful, discretization schemes remain superior if a full sample path of the CIR process has to be simulated or if the CIR process is part of a system of SDEs (Higham & Mao 2005). Moreover, results on strong convergence rates for approximations of equation (1.1) have an interest of their own because they are one of the most prominent examples for an SDE, whose coefficients do not satisfy the standard global Lipschitz assumption.

### 3. Proof of theorem 1.1

In the following, we will denote constants by  $c$ , regardless of their value.

#### (a) Preliminaries

For our error analysis, we need to control the inverse moments of the CIR process. As  $X_t$  follows a non-central  $\chi^2$  distribution, we have

$$\begin{aligned} \mathbf{E} X_t^p &= (x_0 \exp(-\kappa t))^p \left( \frac{2\kappa}{\theta^2} \frac{x_0}{\exp(\kappa t) - 1} \right)^{-p} \\ &\quad \times \frac{\Gamma((2\kappa\lambda/\theta^2) + p)}{\Gamma(2\kappa\lambda/\theta^2)} {}_1F_1 \left( -p, \frac{2\kappa\lambda}{\theta^2}, -\frac{2\kappa}{\theta^2} \frac{x_0}{\exp(\kappa t) - 1} \right), \end{aligned}$$

for  $p > -(2\kappa\lambda/\theta^2)$ , where  ${}_1F_1$  denotes the confluent hypergeometric function, and

$$\mathbf{E} X_t^p = \infty$$

else (Hurd & Kuznetsov 2008, theorem 3.1). As

$${}_1F_1(a, b, z) = \frac{\Gamma(b)}{\Gamma(b-a)} |z|^{-a} (1 + O(|z|^{-1})), \quad z \rightarrow -\infty,$$

see formula 13.1.5 in Abramowitz & Stegun (1964, p. 504), it follows for  $p > -(2\kappa\lambda/\theta^2)$  that  $t \mapsto \mathbf{E}X_t^p$  is bounded on  $[0, T]$ , i.e.

$$\sup_{t \in [0, T]} \mathbf{E}X_t^p < \infty, \quad \text{for } p > -\frac{2\kappa\lambda}{\theta^2}. \quad (3.1)$$

Moreover, from Hurd & Kuznetsov (2008, theorem 3.1) or Bossy & Diop (2007, lemma A.2), we also have

$$\mathbf{E} \exp \left( \varepsilon \int_0^T X_s^{-1} \mathrm{d}s \right) < \infty,$$

if and only if  $0 \leq \varepsilon \leq (\theta^2/8) ((2\kappa\lambda/\theta^2) - 1)^2$ . In general, one can estimate polynomial moments against exponential moments because for  $q, \varepsilon > 0$ , there exists  $c > 0$  such that  $x^q \leq ce^{\varepsilon x}$  for  $x \geq 0$ . Hence, we arrive at the following.

**Lemma 3.1.** *Let  $2\kappa\lambda > \theta^2$ ,  $T > 0$  and  $q \geq 0$ . Therefore,*

$$\mathbf{E} \left( \int_0^T X_s^{-1} \mathrm{d}s \right)^q < \infty.$$

Furthermore, we need a smoothness result for equation (1.2).

**Lemma 3.2.** *Let  $2\kappa\lambda > \theta^2$  and  $T > 0$ . Then, for all  $q \geq 1$ , we have*

$$\begin{aligned} \mathbf{E} |Y_t - Y_s|^q &\leq c |t - s|^{q/2}, \quad \text{for } s, t \in [0, T], \\ \mathbf{E} \sup_{0 \leq s < t \leq T, |t-s| \leq \Delta} |Y_t - Y_s|^q &\leq c (|\log(\Delta)| \Delta)^{q/2}, \quad \text{for } \Delta \in (0, 1/2], \end{aligned}$$

and

$$\mathbf{E} \sup_{t \in [0, T]} |Y_t|^q < \infty.$$

*Proof.* For  $0 \leq s \leq t$ , we have

$$Y_t - Y_s = \int_s^t \frac{\alpha}{Y_u} \mathrm{d}u + \int_s^t \beta Y_u \mathrm{d}u + \gamma (W_t - W_s).$$

By the Cauchy–Schwarz inequality, one has

$$|Y_t - Y_s| \leq \gamma |W_t - W_s| + |t - s|^{1/2} \left( \int_0^T \frac{\alpha^2}{Y_u^2} \mathrm{d}u \right)^{1/2} + |t - s|^{1/2} \left( \int_0^T \beta^2 Y_u^2 \mathrm{d}u \right)^{1/2}.$$

As  $Y_u = \sqrt{X_u}$ , the first assertion follows now from equation (3.1), lemma 3.1, Minkowski's inequality and the smoothness of Brownian motion in the  $q$ th mean. For the second assertion, in addition, we use the well-known fact that the modulus

of continuity of Brownian motion satisfies

$$\mathbf{E} \sup_{0 \leq s < t \leq T, |t-s| \leq \Delta} |W_t - W_s|^q \leq c(|\log(\Delta)|\Delta)^{q/2},$$

for  $\Delta \in (0, 1/2]$  (e.g. Fischer & Nappo 2010). The third assertion follows from

$$|Y_t| \leq y_0 + \gamma |W_t| + t^{1/2} \left( \int_0^T \frac{\alpha^2}{Y_u^2} du \right)^{1/2} + t^{1/2} \left( \int_0^T \beta^2 Y_u^2 du \right)^{1/2}$$

and

$$\mathbf{E} \sup_{t \in [0, T]} |W_t|^q < \infty. \quad \blacksquare$$

(b) Error bound for the implicit Euler scheme for  $Y$

**Proposition 3.3.** Let  $2\kappa\lambda > \theta^2$ . For  $T > 0$  and  $1 \leq p < 2\kappa\lambda/\theta^2$ , there exists  $c > 0$  such that

$$\left( \mathbf{E} \sup_{k=0, \dots, \lceil T/\Delta \rceil} |Y_{k\Delta} - y_k|^p \right)^{1/p} \leq c\sqrt{\Delta},$$

for  $\Delta \in (0, 1/2]$ .

*Proof.* Without loss of generality, we assume that  $\Delta < T$ . For the error

$$e_k = Y_{k\Delta} - y_k$$

at time point  $k\Delta$ , we have the recursion

$$e_0 = 0,$$

$$e_{k+1} = e_k + (f(Y_{(k+1)\Delta}) - f(y_{k+1}))\Delta + r_k,$$

with

$$r_k = - \int_{k\Delta}^{(k+1)\Delta} (f(Y_{(k+1)\Delta}) - f(Y_t)) dt.$$

Multiplying both sides by  $e_{k+1}$ , we obtain

$$e_{k+1}^2 \leq \frac{1}{2} e_k^2 + \frac{1}{2} e_{k+1}^2 + e_{k+1} (f(Y_{(k+1)\Delta}) - f(y_{k+1}))\Delta + e_{k+1} r_k.$$

As

$$e_{k+1} (f(Y_{(k+1)\Delta}) - f(y_{k+1}))\Delta \leq \beta e_{k+1}^2 \Delta \leq 0,$$

we have

$$0 \leq e_n^2 \leq 2 \sum_{k=0}^{n-1} e_{k+1} r_k, \quad n = 1, 2, \dots,$$

and it follows that

$$\sup_{k=0, \dots, \lceil T/\Delta \rceil} |e_k| \leq 2 \sum_{k=0}^{\lceil T/\Delta \rceil - 1} |r_k|. \quad (3.2)$$

Now, it remains to analyse the local error

$$|r_k| = \left| \int_{k\Delta}^{(k+1)\Delta} (f(Y_{(k+1)\Delta}) - f(Y_t)) dt \right|.$$

Note that

$$f(a) - f(b) = \beta(a - b) + \frac{\alpha}{ab}(b - a),$$

for  $a, b > 0$ . Thus,

$$|f(a) - f(b)| \leq c \left(1 + \frac{1}{ab}\right) |a - b|,$$

and we obtain

$$|r_k| \leq c \int_{k\Delta}^{(k+1)\Delta} \left(1 + \frac{1}{Y_t Y_{(k+1)\Delta}}\right) |Y_{(k+1)\Delta} - Y_t| dt.$$

An application of Hölder's and Minkowski's inequality yields

$$(\mathbf{E}|r_k|^p)^{1/p} \leq c \int_{k\Delta}^{(k+1)\Delta} \left(1 + \left(\mathbf{E} \frac{1}{|Y_t Y_{(k+1)\Delta}|^{pq}}\right)^{1/pq}\right) (\mathbf{E}|Y_{(k+1)\Delta} - Y_t|^{pq'})^{1/pq'} dt,$$

for  $q, q' > 1$  with  $1/q + 1/q' = 1$ . Now, lemma 3.2 gives

$$(\mathbf{E}|r_k|^p)^{1/p} \leq c\sqrt{\Delta} \int_{k\Delta}^{(k+1)\Delta} \left(1 + \left(\mathbf{E} \frac{1}{|Y_t Y_{(k+1)\Delta}|^{pq}}\right)^{1/pq}\right) dt, \quad (3.3)$$

for all  $q > 1$ . As

$$\mathbf{E} \frac{1}{|Y_t Y_{(k+1)\Delta}|^{pq}} \leq \mathbf{E} \frac{1}{|Y_t|^{2pq}} + \mathbf{E} \frac{1}{|Y_{(k+1)\Delta}|^{2pq}},$$

applying (3.1) with  $q > 1$  such that  $pq < 2\kappa\lambda/\theta^2$  yields

$$(\mathbf{E}|r_k|^p)^{1/p} \leq c\Delta^{3/2}, \quad \text{for } p < \frac{2\kappa\lambda}{\theta^2},$$

and using (3.2) completes the proof of the proposition. ■

### (c) Moment bounds for the implicit Euler scheme for $Y$

Now we show that all moments of the approximation scheme are uniformly bounded. Multiplying (1.3) by  $y_{k+1}$  yields

$$y_{k+1}^2 = (\alpha + \beta y_{k+1}^2)\Delta + y_{k+1}(y_k + \gamma\Delta_k W).$$

It follows that

$$y_{k+1}^2 \leq (2\alpha + \gamma^2)\Delta + y_k^2 + M_k, \quad (3.4)$$

with

$$M_k = 2\gamma y_k \Delta_k W + \gamma^2(\Delta_k W^2 - \Delta), \quad (3.5)$$

and we obtain by induction that

$$\mathbf{E} y_{k+1}^2 \leq y_0^2 + (2\alpha + \gamma^2)(k+1)\Delta, \quad (3.6)$$

for all  $k = 0, 1, \dots, \lceil T/\Delta \rceil$ . This allows us to show the following.

**Lemma 3.4.** *Let  $2\kappa\lambda \geq \theta^2$  and  $\Delta > 0$ ,  $T > 0$ . Then, for all  $p \geq 1$ , we have*

$$\mathbf{E} \sup_{k=0, \dots, \lceil T/\Delta \rceil} |y_k|^p < \infty.$$



*Proof.* From (3.4) and (3.5), we obtain

$$\sup_{k=0,\dots,\lceil T/\Delta \rceil} |y_k|^2 \leq c + c \sup_{k=0,\dots,\lceil T/\Delta \rceil} \left| \int_0^{k\Delta} m_t \, dW_t \right|, \quad (3.7)$$

with

$$m_t = 2\gamma y_\ell + 2\gamma^2(W_t - W_{\ell\Delta}), \quad t \in [\ell\Delta, (\ell+1)\Delta).$$

As

$$\sup_{t \in [0, \lceil T/\Delta \rceil \Delta]} \mathbf{E} |m_t|^p \leq c + c \sup_{\ell=0,\dots,\lceil T/\Delta \rceil} \mathbf{E} |y_\ell|^p,$$

the Burkholder–Davis–Gundy inequality, i.e.

$$\mathbf{E} \sup_{s \in [0, t]} \left| \int_0^s m_\tau \, dW_\tau \right|^p \leq c \mathbf{E} \left| \int_0^t m_\tau^2 \, d\tau \right|^{p/2},$$

Jensen's inequality and (3.7) give

$$\mathbf{E} \sup_{k=0,\dots,\lceil T/\Delta \rceil} |y_k|^{2p} \leq c + c \sup_{\ell=0,\dots,\lceil T/\Delta \rceil} \mathbf{E} |y_\ell|^p.$$

Therefore, (3.6) now yields

$$\mathbf{E} \sup_{k=0,\dots,\lceil T/\Delta \rceil} |y_k|^4 < \infty,$$

and the assertion follows from an induction procedure in  $p$ . ■

(d) *Error bound for the drift-implicit square-root Euler scheme*

Now denote by  $\bar{X}$  the piecewise linear interpolation of the CIR process with stepsize  $\Delta > 0$ , i.e.

$$\bar{X}_t = \left(k + 1 - \frac{t}{\Delta}\right) X_{k\Delta} + \left(\frac{t}{\Delta} - k\right) X_{(k+1)\Delta}, \quad t \in [k\Delta, (k+1)\Delta].$$

**Lemma 3.5.** *Let  $2\kappa\lambda > \theta^2$ ,  $T > 0$  and  $\Delta \in (0, 1/2]$ . Then, for all  $q \geq 1$ , we have*

$$\mathbf{E} \max_{t \in [0, T]} |X_t - \bar{X}_t|^q \leq c(|\log(\Delta)|\Delta)^{q/2}.$$

*Proof.* Combining the equality

$$X_t - X_s = (Y_t + Y_s)(Y_t - Y_s)$$

with the Cauchy–Schwarz inequality and lemma 3.2, we get

$$\mathbf{E} \sup_{0 \leq s < t \leq T, |t-s| \leq \Delta} |X_t - X_s|^q \leq c(|\log(\Delta)|\Delta)^{q/2}.$$

Now, the assertion follows from

$$\sup_{t \in [0, T]} |X_t - \bar{X}_t| \leq \sup_{0 \leq s < t \leq T, |t-s| \leq \Delta} |X_t - X_s|. \quad \blacksquare$$

Owing to

$$\sup_{t \in [0, T]} |X_t - \bar{x}_t| \leq \sup_{t \in [0, T]} |X_t - \bar{X}_t| + \sup_{k=0, \dots, \lceil T/\Delta \rceil} |X_{k\Delta} - y_k^2|$$

and lemma 3.5, it only remains to control

$$\left( \mathbf{E} \sup_{k=0, \dots, \lceil T/\Delta \rceil} |X_{k\Delta} - y_k^2|^p \right)^{1/p}.$$

As

$$\mathbf{E} \sup_{k=0, \dots, \lceil T/\Delta \rceil} |X_{k\Delta} - y_k^2|^p \leq \mathbf{E} \left( \sup_{k=0, \dots, \lceil T/\Delta \rceil} |Y_{k\Delta} + y_k|^p \sup_{k=0, \dots, \lceil T/\Delta \rceil} |Y_{k\Delta} - y_k|^p \right),$$

an application of Hölder's inequality with  $\varepsilon > 0$  such that  $(1 + \varepsilon)p < 2\kappa\lambda/\theta^2$  and proposition 3.3 give

$$\mathbf{E} \sup_{k=0, \dots, \lceil T/\Delta \rceil} |X_{k\Delta} - y_k^2|^p \leq \left( \mathbf{E} \sup_{k=0, \dots, \lceil T/\Delta \rceil} |Y_{k\Delta} + y_k|^{p(1+\varepsilon)/\varepsilon} \right)^{\varepsilon/(1+\varepsilon)} \Delta^{p/2}.$$

It remains to apply lemmas 3.2 and 3.4 to finish the proof of theorem 1.1.

The authors thank Martin Altmayer for valuable comments on an earlier version of the manuscript.

## References

- Abramowitz, M. & Stegun, I. A. 1964 *Handbook of mathematical functions with formulas, graphs and mathematical tables*. Washington, DC: US Department of Commerce.
- Alfonsi, A. 2005 On the discretization schemes for the CIR (and Bessel squared) processes. *Monte Carlo Methods Appl.* **11**, 355–384. (doi:10.1515/156939605777438569)
- Alfonsi, A. 2010 High order discretization schemes for the CIR process: application to affine term structure and Heston models. *Math. Comput.* **79**, 209–237. (doi:10.1090/S0025-5718-09-02252-2)
- Berkaoui, A., Bossy, M. & Diop, A. 2008 Euler scheme for SDEs with non-Lipschitz diffusion coefficient: strong convergence. *ESAIM: Probab. Stat.* **12**, 1–11. (doi:10.1051/ps:2007030)
- Beskos, A. & Roberts, G. O. 2005 Exact simulation of diffusions. *Ann. Appl. Probab.* **15**, 2422–2444. (doi:10.1214/105051605000000485)
- Beskos, A., Papaspiliopoulos, O., Roberts, G. O. & Fearnhead, P. 2006 Exact and computationally efficient likelihood-based estimation for discretely observed diffusion processes (with discussion). *J. R. Stat. Soc., Ser. B. Stat. Methodol.* **68**, 333–382. (doi:10.1111/j.1467-9868.2006.00552.x)
- Bossy, M. & Diop, A. 2007 An efficient discretization scheme for one dimensional SDEs with a diffusion coefficient function of the form  $|x|^a$ ,  $a \in [1/2, 1)$ . Research report RR-5396, INRIA, Sophia Antipolis.
- Broadie, M. & Kaya, Ö. 2006 Exact simulation of stochastic volatility and other affine jump diffusion processes. *Oper. Res.* **54**, 217–231. (doi:10.1287/opre.1050.0247)
- Cox, J., Ingersoll, J. & Ross, S. 1985 A theory of the term structure of interest rates. *Econometrica* **53**, 385–408. (doi:10.2307/1911242)
- Deelstra, G. & Delbaen, F. 1998 Convergence of discretized stochastic (interest rate) processes with stochastic drift term. *Appl. Stoch. Models Data Anal.* **14**, 77–84. (doi:10.1002/(SICI)1099-0747(199803)14:1<77::AID-ASM338>3.0.CO;2-2)
- Fischer, M. & Nappo, G. 2010 On the moments of the modulus of continuity of Itô processes. *Stoch. Anal. Appl.* **28**, 103–122. (doi:10.1080/07362990903415825)

- Giles, M. 2007 Improved multilevel Monte Carlo convergence using the Milstein scheme. In *Proc. Monte Carlo and Quasi-Monte Carlo Methods, Ulm, Germany, 14–18 August 2006*. (eds A. Keller, S. Heinrich & H. Niederreiter), pp. 343–354. Berlin, Germany: Springer.
- Giles, M. 2008 Multi-level Monte Carlo path simulation. *Oper. Res.* **56**, 607–617. (doi:10.1287/opre.1070.0496)
- Glasserman, P. 2004 *Monte Carlo methods in financial engineering*. New York, NY: Springer.
- Günther, M., Kahl, C. & Rossberg, T. 2008 Structure preserving stochastic integration schemes in interest rate derivative modeling. *Appl. Numer. Math.* **58**, 284–295. (doi:10.1016/j.apnum.2006.11.013)
- Gyöngy, I. 1998 A note on Euler’s approximations. *Potential Anal.* **8**, 205–216. (doi:10.1023/A:1016557804966)
- Gyöngy, I. & Rásonyi, M. 2011 A note on Euler approximations for SDEs with Hölder continuous diffusion coefficients. *Stoch. Proc. Appl.* **121**, 2189–2200. (doi:10.1016/j.spa.2011.06.008)
- Heston, S. L. 1993 A closed form solution for options with stochastic volatility, with applications to bonds and currency options. *Rev. Financ. Stud.* **6**, 327–343. (doi:10.1093/rfs/6.2.327)
- Higham, D. J. & Mao, X. 2005 Convergence of Monte Carlo simulations involving the mean-reverting square root process. *J. Comp. Fin.* **8**, 35–61.
- 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. (doi:10.1137/S0036142901389530)
- Higham, D. J., Mao, X., Szpruch, L. & Pan, J. 2011 Numerical simulation of a strongly nonlinear Ait-Sahalia-type interest rate model. *BIT* **51**, 405–425. (doi:10.1007/s10543-010-0288-y)
- Hurd, T. R. & Kuznetsov, A. 2008 Explicit formulas for Laplace transforms of stochastic integrals. *Markov Process. Relat. Fields* **14**, 277–290.
- Jäckel, P. & Kahl, C. 2006 Fast strong approximation Monte-Carlo schemes for stochastic volatility models. *Quant. Finance* **6**, 513–536.
- Karatzas, I. & Shreve, S. E. 1991 *Brownian motion and stochastic calculus*, 2nd edn. New York, NY: Springer.
- Kloeden, P. E. & Platen, E. 1999 *Numerical solution of stochastic differential equations*, 3rd edn. Berlin, Germany: Springer.
- Milstein, G. N. 1995 *Numerical integration of stochastic differential equations*. Dordrecht, The Netherlands: Kluwer.
- Müller-Gronbach, T. 2002 The optimal uniform approximation of systems of stochastic differential equations. *Ann. Appl. Probab.* **12**, 664–690. (doi:10.1214/aoap/1026915620)
- Moro, E. & Schurz, H. 2007 Boundary preserving semianalytic numerical algorithms for stochastic differential equations. *SIAM J. Sci. Comput.* **29**, 1525–1549. (doi:10.1137/05063725X)
- Neuenkirch, A. & Zähle, H. 2009 Asymptotic error distribution of the Euler method for SDEs with non-Lipschitz coefficients. *Monte Carlo Methods Appl.* **15**, 333–351. (doi:10.1515/MCMA.2009.018)