AN EULER-TYPE METHOD FOR THE STRONG APPROXIMATION OF THE COX-INGERSOLL-ROSS PROCESS

STEFFEN DEREICH, ANDREAS NEUENKIRCH AND LUKASZ SZPRUCH

ABSTRACT. We analyze 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 (SDEs) 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
MATHEMATICS SUBJECT CLASSIFICATION (2000). 65C30; 60H35

1. INTRODUCTION AND MAIN RESULT

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

$$dX_t = \kappa(\lambda - X_t)\,dt + \theta\sqrt{X_t}\,dW_t, \quad X_0 = x_0, \qquad t \geq 0. \qquad (1)$$

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) admits a unique strong solution which is non-negative, see e.g. Chapter 5 in [24]. The CIR process was originally proposed by Cox et al. [9] in 1985 as a model for short-term interest rates. Nowadays, this model is widely used in financial modeling, e.g. as volatility process in the Heston model [18].

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 modeled by a d-dimensional SDE with solution $(S_t)_{t\in[0,T]}$, then this corresponds to the quadrature problem

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

where $F : C([0,T], \mathbb{R}^d) \to \mathbb{R}$ 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 stochastic differential equation 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 [12, 13] is to be used.

The strong global approximation of equation (1) has been studied in several articles. Strong convergence (without a rate or with a logarithmic rate) of several discretization schemes has been shown in [2, 7, 10, 17, 19]. In [2], a general framework for the analysis of strong approximation of the CIR process is presented along with extensive simulation studies. Moreover, the strong approximation of more general Ait-Sahalia-type interest rate models is analyzed in [21]. However, only in [4] non-logarithmic convergence rates are obtained. In [4] it is shown that a symmetrized Euler method has strong convergence order 1/2 under restrictive assumptions on the parameters of the equation, see Section 2. The difficulties to obtain strong convergence rates for the approximations of the CIR process are due to its square-root coefficient. Thus the standard theory which relies on the global Lipschitz assumption does not apply [25, 26].

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 which is often fulfilled in interest rate models. By the Feller test, it is true if and only if $2\kappa\lambda \geq \theta^2$, see e.g. Chapter 5 in [24]. 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

$$dY_t = \frac{\alpha}{Y_t}\,dt + \beta Y_t\,dt + \gamma\,dW_t, \quad t \geq 0, \qquad Y_0 = \sqrt{x_0}, \tag{2}$$

with

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

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

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

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

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

This property is crucial to control the error propagation of drift-implicit Euler schemes, see [20, 21].

The drift-implicit Euler method with stepsize $\Delta > 0$ for equation (2) leads to the numerical scheme

$$y_{k+1} = y_k + \left(\frac{\alpha}{y_{k+1}} + \beta y_{k+1}\right)\Delta + \gamma\Delta_k W, \qquad k = 0, 1, \dots \tag{3}$$

with $y_0 = \sqrt{x_0}$ and

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

Recalling that $\alpha, \gamma > 0$ and $\beta < 0$, equation (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 *drift-implicit square-root Euler method*. Transforming back, i.e.

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

gives a strictly positive approximation of the original CIR process. This scheme has already been suggested in [2], but convergence results have not been established. Using piecewise linear interpolation, i.e.

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

we obtain a global approximation $(\overline{x}_t)_{t\in[0,T]}$ of the CIR process on $[0,T]$. Exploiting the structure of SDE (2) we establish the following theorem.

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

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

there exists a constant $K_p > 0$ such that

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

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

Hence we obtain the optimal strong convergence rate for the approximation of SDEs with Lipschitz coefficients, see [27]. 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 $p \ge \frac{2\kappa\lambda}{\theta^2}$.

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 Section 3.

2. Numerical Methods for the CIR Process

Discretization schemes for the CIR process were proposed in numerous articles, among these are [2, 3, 4, 7, 10, 15, 17, 19, 23, 28]. In this short summary we mainly focus on Euler-type methods.

In [4] the authors study a symmetrized Euler method for the approximation of equation (1), i.e.

$$x_{k+1} = |x_k + \kappa(\lambda - x_k)\Delta + \theta\sqrt{x_k}\,\Delta_k W| \tag{4}$$

for $k = 0, 1, \ldots$. 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

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

where the constant $C_p > 0$ depends only $p, \kappa, \lambda, \theta, x_0$ and T. As a consequence of Lemma 3.5 in Section 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 is shown in [7] that the symmetrized Euler scheme converges weakly with rate 1 if $2\kappa\lambda \ge 2\theta^2$.

The truncated Euler scheme

$$x_{k+1} = x_k + \kappa(\lambda - x_k)\Delta + \theta\sqrt{x_k^+}\,\Delta_k W, \qquad k = 0, 1, \dots \tag{5}$$

was analyzed in [10], while the scheme

$$x_{k+1} = x_k + \kappa(\lambda - x_k)\Delta + \theta\sqrt{|x_k|}\,\Delta_k W, \qquad k = 0, 1, \dots \tag{6}$$

was studied in [19]. 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 $\Delta \to 0$ without further restrictions on the parameters of the equation. In particular an application of Theorem 3.1 in [17] yields a logarithmic convergence rate for the Euler schemes (5) and (6). Moreover, if $2\kappa\lambda \geq \theta^2$ it follows from [16] that the Euler schemes (4) – (6) have a pathwise convergence rate of $1/2 - \varepsilon$ for all $\varepsilon > 0$, i.e. we have

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

for $\Delta \to 0$. The asymptotic error distribution of these schemes can be deduced from [29]: For $2\kappa\lambda \geq \theta^2$ it holds that

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

for $\Delta \to 0$, 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

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

While no explicit solution of equation (1) is known, the finite dimensional distributions can be characterized in terms of a non-central chi-square distribution, see e.g. [9]. Thus, equation (1) can be simulated exactly at a finite number of time points using the Markov property, see e.g. [8, 14].

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-square 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 κ, λ and θ. 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 [5], which can be also applied to equation (1), see [6].

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 stochastic differential equations, see e.g. [19]. Moreover, results on strong convergence rates for approximations of equation (1) have an interest of its own, since it is one of the most prominent examples for a stochastic differential equation, whose coefficients do not satisfy the standard global Lipschitz assumption.

3. Proof of Theorem 1.1

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

3.1. Preliminaries.

For our error analysis we need to control the inverse moments of the CIR process. Since X_t follows a non-central chi-square distribution, we have

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

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

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

else, see e.g. Theorem 3.1 in [22]. Since

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

see formula 13.1.5 on page 504 in [1], it follows for $p > -\frac{2\kappa\lambda}{\theta^2}$ that $t \mapsto \mathbf{E}X_t^p$ is bounded on $[0, T]$, i.e.

$$\sup_{t \in [0,T]} \mathbf{E}X_t^p < \infty \qquad \text{for} \quad p > -\frac{2\kappa\lambda}{\theta^2}. \tag{7}$$

Moreover, from Theorem 3.1 [22] or Lemma A.2 in [7] we also have

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

if and only if $0 \le \varepsilon \le \frac{\theta^2}{8} \left(\frac{2\kappa\lambda}{\theta^2} - 1 \right)^2$. In general, one can estimate polynomial moments against exponential moments, since for $q, \varepsilon > 0$ there exists $c > 0$ such that $x^q \le ce^{\varepsilon x}$ for $x \ge 0$. Hence we arrive at

Lemma 3.1. *Let $2\kappa\lambda > \theta^2$, $T > 0$ and $q \ge 0$. It holds*

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

Further we need a smoothness result for equation (2):

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

$$\mathbf{E}|Y_t - Y_s|^q \le c \cdot |t - s|^{q/2}, \qquad\qquad for \quad s, t \in [0, T],$$
$$\mathbf{E} \sup_{0 \le s < t \le T, |t - s| \le \Delta} |Y_t - Y_s|^q \le c \cdot (|\log(\Delta)|\Delta)^{q/2}, \qquad for \quad \Delta \in (0, 1/2],$$

and

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

Proof. For $0 \le s \le t$, we have

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

By the Cauchy-Schwarz inequality one has

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

Since $Y_u = \sqrt{X_u}$ the first assertion follows now from (7), 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

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

for $\Delta \in (0, 1/2]$, see e.g. [11]. The third assertion follows from

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

and

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

$\square$

3.2. **Error Bound for the implicit Euler Scheme for Y.**

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

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

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

$$e_0 = 0$$

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

with

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

Multiplying both sides with e_{k+1} we obtain

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

Since

$$e_{k+1} \left(f(Y_{(k+1)\Delta}) - f(y_{k+1}) \right)\Delta \le \beta e_{k+1}^2 \Delta \le 0$$

we have

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

and it follows that

$$\sup_{k=0,\dots,\lceil T/\Delta \rceil} |e_k| \le 2 \sum_{k=0}^{\lceil T/\Delta \rceil - 1} |r_k|. \tag{8}$$

Now, it remains to analyze 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

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

and we obtain

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

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

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

for $q, q' > 1$ with $\frac{1}{q} + \frac{1}{q'} = 1$. Now, Lemma 3.2 gives

$$\left(\mathbf{E}|r_k|^p\right)^{1/p} \le c \cdot \sqrt{\Delta} \cdot \int_{k\Delta}^{(k+1)\Delta} \left(1 + \left(\mathbf{E} \frac{1}{|Y_t Y_{(k+1)\Delta}|^{pq}} \right)^{\frac{1}{pq}} \right) dt, \tag{9}$$

for all $q > 1$. Since

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

applying (7) with $q > 1$ such that $pq < \frac{2\kappa\lambda}{\theta^2}$ yields

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

and using (8) completes the proof of the Proposition. $\qquad \square$

3.3. **Moment Bounds for the implicit Euler Scheme for Y.** Now we show that all moments of the approximation scheme are uniformly bounded. Multiplying (3) with 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

$$y_{k+1}^2 \le (2\alpha + \gamma^2)\Delta + y_k^2 + M_k \qquad (10)$$

with

$$M_k = 2\gamma y_k \Delta_k W + \gamma^2(\Delta_k W^2 - \Delta), \qquad (11)$$

and we obtain by induction that

$$\mathbf{E}y_{k+1}^2 \le y_0^2 + (2\alpha + \gamma^2)(k+1)\Delta \qquad (12)$$

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

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

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

Proof. From (10) and (11) we obtain that

$$\sup_{k=0,\ldots,\lceil T/\Delta \rceil} |y_k|^2 \le c + c \sup_{k=0,\ldots,\lceil T/\Delta \rceil} \left| \int_0^{k\Delta} m_t \, dW_t \right| \qquad (13)$$

with

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

Since

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

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

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

Jensen's inequality and (13) give that

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

So (12) now yields

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

and the assertion follows from an induction procedure in p.   □

3.4. **Error Bound for the drift-implicit square-root Euler Scheme.** Now denote by $\overline{X}$ the piecewise linear interpolation of the CIR process with stepsize $\Delta > 0$, i.e.

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

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*

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

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

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

$\square$

Due to

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

and the above Lemma it only remains to control

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

Since

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

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

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

It remains to apply Lemma 3.2 and Lemma 3.4 to finish the proof of Theorem 1.1.

References

[1] Abramowitz, M., Stegun, I.A.: Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables. U.S. Department of Commerce, Washington (1964)

[2] Alfonsi, A.: On the discretization schemes for the CIR (and Bessel squared) processes. Monte Carlo Methods Appl. **11**(4), 355-384 (2005)

[3] Alfonsi, A.: High order discretization schemes for the CIR process: Application to affine term structure and Heston models. Math. Comput. **79**(269), 209-237 (2010)

[4] Berkaoui, A., Bossy, M., Diop, A.: Euler scheme for SDEs with non-Lipschitz diffusion coefficient: strong convergence. ESAIM, Probab. Stat. **12**, 1-11 (2008)

[5] Beskos, A., Roberts, G.O.: Exact simulation of diffusions. Ann. Appl. Probab. **15**(4), 2422-2444 (2005)

[6] Beskos, A., Papaspiliopoulos, O., Roberts, G.O., Fearnhead, P.: Exact and computationally efficient likelihood-based estimation for discretely observed diffusion processes (With discussion). J. R. Stat. Soc., Ser. B, Stat. Methodol. **68**(3), 333-382 (2006)

[7] Bossy, M., Diop, A.: An efficient discretization scheme for one dimensional SDEs with a diffusion coefficient function of the form $|x|^a, a \in [1/2, 1)$. Working paper, INRIA (2007)

[8] Broadie, M., Kaya, Ö.: Exact simulation of stochastic volatility and other affine jump diffusion processes. Oper. Res. **54**, 217-231 (2006)

[9] Cox, J., Ingersoll, J., Ross, S.: A theory of the term structure of interest rates. Econometrica **53**, 385-408 (1985)

[10] Deelstra, G., Delbaen, F.: Convergence of discretized stochastic (interest rate) processes with stochastic drift term. Appl. Stochastic Models Data Anal. **14**, 77-84 (1998)

[11] Fischer, M., Nappo, G.: On the moments of the modulus of continuity of Itô processes. Stochastic Anal. Appl. **28**(1), 103-122 (2010)

[12] Giles, M.: Multi-level Monte Carlo path simulation, Oper. Res. **56**(3), 607-617 (2008)

[13] Giles, M.: Improved multilevel Monte Carlo convergence using the Milstein scheme. In: Keller, A. (ed.) et al., Monte Carlo and Quasi-Monte Carlo Methods 2006, Proceedings. Berlin, Springer, 343-354 (2007)

[14] Glassermann, P.: Monte Carlo Methods in Financial Engineering. Springer, New York (2004)

[15] Günther, M., Kahl, C., Roßberg, T.: Structure preserving stochastic integration schemes in interest rate derivative modeling. Appl. Numer. Math. **58**(3), 284-295 (2008)

[16] Gyöngy, I.: A note on Euler's approximations. Potential Anal. **8**, 205-216 (1998)

[17] Gyöngy, I., Rásonyi, M.: A note on Euler approximations for SDEs with Hölder continuous diffusion coefficients. Stoch. Proc. Appl. **121**(10), 2189-2200 (2011)

[18] Heston, S.L.: A closed form solution for options with stochastic volatility, with applications to bonds and currency options. Rev. Financial Stud. **6**, 327-343 (1993)

[19] Higham, D.J., Mao, X.: Convergence of Monte Carlo simulations involving the mean-reverting square root process. J. Comp. Fin. **8**, 35-61 (2005)

[20] Higham, D.J., Mao, X., Stuart, A.M.: Strong convergence of Euler-type methods for nonlinear stochastic differential equations. SIAM J. Numer. Anal. **40**(3), 1041-1063 (2002)

[21] Higham, D.J., Mao, X., Szpruch, L., Pan. J.: Numerical simulation of a strongly nonlinear Ait-Sahalia-type interest rate model. BIT **51**(2), 405-425 (2011)

[22] Hurd, T.R., Kuznetsov, A.: Explicit formulas for Laplace transforms of stochastic integrals. Markov Process. Relat. Fields **14**(2), 277-290 (2008)

[23] Jäckel, P., Kahl, C.: Fast strong approximation Monte-Carlo schemes for stochastic volatility models. Quant. Finance **6**, 513-536 (2006)

[24] Karatzas, I., Shreve, S.E.: Brownian Motion and Stochastic Calculus. Springer, New York, 2nd edition (1991)

[25] Kloeden, P.E., Platen, E.: Numerical Solution of Stochastic Differential Equations. Springer, Berlin, 3rd edition (1999)

[26] Milstein, G.N.: Numerical Integration of Stochastic Differential Equations. Kluwer, Doordrecht (1995)

[27] Müller-Gronbach, T.: The optimal uniform approximation of systems of stochastic differential equations. Ann. Appl. Probab. **12**(2), 664-690 (2002)

[28] Moro, E., Schurz, H.: Boundary preserving semianalytic numerical algorithms for stochastic differential equations. SIAM J. Sci. Comput. **29**, 1525-1549 (2007)

[29] Neuenkirch, A., Zähle, H.: Asymptotic error distribution of the Euler method for SDEs with non-Lipschitz coefficients. Monte Carlo Methods Appl. **15**(4), 333-351 (2009)

STEFFEN DEREICH, FACHBEREICH MATHEMATIK UND INFORMATIK, PHILIPPS-UNIVERSITÄT MARBURG, HANS-MEERWEIN-STRASSE, D-35032 MARBURG, GERMANY dereich@mathematik.uni-marburg.de

ANDREAS NEUENKIRCH, FACHBEREICH MATHEMATIK, TU KAISERSLAUTERN, POSTFACH 3049, D-67663 KAISERSLAUTERN, GERMANY neuenkirch@mathematik.uni-kl.de

LUKASZ SZPRUCH, MATHEMATICAL INSTITUTE, UNIVERSITY OF OXFORD, 24-29 ST GILES', OXFORD OX1 3LB, ENGLAND Szpruch@maths.ox.ac.uk