Numer. Math. (2014) 128:103–136
DOI 10.1007/s00211-014-0606-4

Numerische Mathematik

First order strong approximations of scalar SDEs defined in a domain

Andreas Neuenkirch · Lukasz Szpruch

Received: 3 September 2012 / Revised: 9 October 2013 / Published online: 31 January 2014

Abstract We are interested in strong approximations of one-dimensional SDEs which have non-Lipschitz coefficients and which take values in a domain. Under a set of general assumptions we derive an implicit scheme that preserves the domain of the SDEs and is strongly convergent with rate one. Moreover, we show that this general result can be applied to many SDEs we encounter in mathematical finance and bio-mathematics. We will demonstrate flexibility of our approach by analyzing classical examples of SDEs with sublinear coefficients (CIR, CEV models and Wright–Fisher diffusion) and also with superlinear coefficients (3/2-volatility, Aït-Sahalia model). Our goal is to justify an efficient Multilevel Monte Carlo method for a rich family of SDEs, which relies on good strong convergence properties.

Mathematics Subject Classification (2000) 60H10 · 65J15

1 Introduction

The goal of this paper is to derive an efficient numerical approximation for one-dimensional SDEs which take values in a domain and have non-Lipschitz drift or diffusion coefficients. Typical examples of such SDEs are the Cox–Ingersoll–Ross process (CIR), the CEV model and the Wright–Fisher diffusion, where the main difficulty is the sublinearity of the diffusion coefficients. Furthermore, the approach developed in this

A. Neuenkirch (✉)
Institut für Mathematik, Universität Mannheim, A5, 6, 68131 Mannheim, Germany
e-mail: neuenkirch@kiwi.math.uni-mannheim.de

L. Szpruch
Mathematical Institute, University of Oxford, 24-29 St Giles', Oxford OX1 3LB, UK
e-mail: Szpruch@maths.ox.ac.uk

paper can be also applied to SDEs with superlinear coefficients. Prominent examples are here the Heston-3/2 volatility process and the Aït-Shalia model. All the mentioned processes play an important role in mathematical finance or bio-mathematical applications. Our key idea is to transform the original SDE using the Lamperti transformation into a SDE with constant diffusion coefficient, see e.g. [19]. The transformed SDE is then approximated by a backward (also called drift-implicit) Euler–Maruyama scheme (BEM) and transforming back yields an approximation scheme for the original SDE. This strategy was suggested by Alfonsi in [1] for the CIR process and was found successful in a recent work [7] where the authors proved that the piecewise linearly interpolated BEM scheme strongly converges with rate one half (up to a log-term) with respect to a uniform L^p-error criterion. Here, we extend that work in several ways:

- Considering the maximum error in the discretization points, we prove that the drift-implicit Euler–Maruyama scheme for the CIR process strongly converges with rate one under slightly more restrictive conditions on the parameters of the process than in [7].
- We provide a general framework for the strong order one convergence of the BEM scheme for SDEs with constant diffusion and one-sided Lipschitz drift coefficients.
- Using this framework we present a detailed convergence analysis for several SDEs with sub- and super-linear coefficients.
- We also show that BEM for the transformed SDE is closely related to a drift-implicit Milstein scheme for the original SDE, which has been introduced in [14]. In the case of the CIR process we provide an error analysis for this scheme.

Independently of and simultaneously to the research presented in this paper, the same approach was also used by Alfonsi in [2] to derive strong order one convergence of the BEM scheme for the CIR and the CEV process. See Remark 2.9 for a discussion and a comparison.

To illustrate the main difficulties and also our main idea let us consider the CIR process

$$dy(t) = \kappa(\theta - y(t))dt + \sigma\sqrt{y(t)}dw(t) \tag{1}$$

with $2\kappa\theta \geq \sigma^2$. It is a simple implication of the Feller test that the solution of Eq. (1) is strictly positive when $2\kappa\theta \geq \sigma^2$ and $y(0) > 0$. This SDE is often used in mathematical finance for interest rate or stochastic volatility models. However scalar SDEs with square root diffusion coefficients appear not only in the financial literature but belong to the most fundamental SDEs as they are an approximation to Markov jump processes [8].

Once we attempt to simulate (1) using classical discretization methods, see e.g. [26], we face two difficulties:

①不能保证 因此一般的数值方法做不到).

- In general, these methods do not preserve positivity and therefore are not well defined when directly applied to Eq. (1).
  ②非Lipschitz连续
- The diffusion term is not Lipschitz continuous and therefore standard assumptions required for weak and strong convergence, see e.g. [26], do not hold.

Consequently, a considerable amount of research was devoted to the numerical approximation of this equation, see [1,3,4,12,22,28], to mention a few. However no strong convergence of order one results have been obtained so far up to best of our knowledge. For a comparison of different proposed schemes based on simulation studies, see [1,28].

Our approach is based on a suitable transformation of the CIR process. Applying Itô's formula to $\boxed{x(t) = \sqrt{y(t)}}$ gives

$$dx(t) = \frac{1}{2}\kappa\left(\left(\theta - \frac{\sigma^2}{4\kappa}\right)x(t)^{-1} - x(t)\right)dt + \frac{1}{2}\sigma dw(t). \tag{2}$$

Zhu [32] pointed out that a drawback of the transformed equation is that a naive Euler discretization cannot capture the erratic behavior of the $x(t)^{-1}$-term, although almost sure convergence of this method holds true, see [20]. The weakness of a naive Euler discretization is that its transition density is Gaussian and therefore its moments explode due to the $x(t)^{-1}$-term.

On the other hand, Alfonsi showed in [1] that BEM applied to (1) preserves positivity of the solution and also monotonicity with respect to the initial value. Moreover, his simulation studies indicated good convergence properties of this scheme. In this paper we follow the recent result by Dereich et al. [7], where it was shown that the piecewise linear interpolation of BEM applied to (1) strongly converges with a rate one half up to a log-term.

Given any step size $\Delta t > 0$, the BEM scheme has the form

$$\begin{aligned} X_{k+1} &= X_k + \frac{1}{2}\kappa\left(\theta_v X_{k+1}^{-1} - X_{k+1}\right)\Delta t + \frac{1}{2}\sigma\Delta w_{k+1}, \qquad k = 0, 1, \ldots, \\ X_0 &= x(0) \end{aligned} \tag{3}$$

with

$$\Delta w_{k+1} = w((k+1)\Delta t) - w(k\Delta t), \qquad k = 0, 1, \ldots$$

and

$$\theta_v = \theta - \frac{\sigma^2}{4\kappa}.$$

We will establish a strong convergence of order one for the maximum L^p-distance in the discretization points between (1) and (3), see Sect. 3. For example for the L^2-distance we obtain

$$\mathbb{E} \max_{k=0,\ldots,\lceil T/\Delta t\rceil} |X_k - x(k\Delta t)|^2 \le C_2 \cdot |\Delta t|^2 \qquad \text{for} \quad \frac{2\kappa\theta}{\sigma^2} > 3.$$

As a consequence we also obtain the same convergence order for the approximation of the original CIR process by X_k^2, $k = 0, 1, \ldots, \lceil T/\Delta t\rceil$.

In this paper we will show that the above idea naturally extends to many types of SDEs with non-Lipschitz coefficients. The combination of the Lamperti transformation and the backward Euler scheme enables us to analyse the L^p-convergence rates for many scalar SDEs encountered in practice. In particular, transforming BEM back we obtain an order one scheme for the original SDE, which is close to a Milstein-type scheme, see Sect. 4. Hence our approach turns out to be a new method for deriving numerical methods with strong order one convergence for SDEs with non-Lipschitz coefficients. Although strong convergence of backward schemes for SDEs with non-Lipschitz coefficients has already been analysed in the literature and their convergence for models as the Aït-Sahalia and the Heston-3/2 volatility was obtained, see [14,29,30], schemes with strong convergence order one have not been established yet in this setting.

Another motivation for our work are results by Giles [9,10], who showed that for optimal MLMC simulations one should use discretization schemes with strong convergence order one. Note that strong convergence of the discretization scheme used for the MLMC simulations seems to be not only a sufficient but also a necessary condition [17].

The remainder of this paper is structured as follows. In the next section, we provide a general convergence result for the BEM method applied to scalar SDEs with additive noise. Section 3 contains the results for our examples, i.e. the CIR, CEV, Aït-Sahalia, 3/2-Heston volatility and Wright–Fischer SDEs. In Sect. 4, we provide the relation of BEM and a drift-implicit Milstein scheme and give an error analysis for the case of the CIR process. The last section contains a short discussion.

2 The BEM scheme for SDEs with additive noise

2.1 Preliminaries

Let $D = (l, r)$, where $-\infty \leq l < r \leq \infty$, and let $a, b : D \to D$ be continuously differentiable functions. Moreover, let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ be a filtered probability space and $w(t),\ t \geq 0$, a standard $(\mathcal{F}_t)_{t\geq 0}$-Brownian motion. We begin with the SDE

$$dy(t) = a(y(t))dt + b(y(t))dw(t), \quad t \geq 0, \qquad y(0) \in D, \tag{4}$$

and assume that it has a unique strong solution with

$$\mathbb{P}(y(t) \in D,\ t \geq 0) = 1.$$

If $b(x) > 0$ for all $x \in D$, then we can use the Lamperti-type transformation

$$F(x) = \lambda \int^{x} \frac{1}{b(y)} dy$$

for some $\lambda > 0$. Its inverse $F^{-1} : F(D) \to D$ is well defined and Itô's Lemma with $x(t) = F(y(t))$ gives the transformed SDE

$$dx(t) = f(x(t))dt + \lambda dw(t), \quad t \geq 0, \qquad x(0) \in F(D),$$

with

$$f(x) = \lambda \left(\frac{a(F^{-1}(x))}{b(F^{-1}(x))} - \frac{1}{2} b'(F^{-1}(x)) \right), \qquad x \in F(D),$$

where $F(D) = (F(l), F(r))$. Note that the classical Lamperti transformation corresponds to $\lambda = 1$. This transformation allows to shift non-linearities from the diffusion coefficient into the drift coefficient. Then, under appropriate assumptions on f (respectively a and b), one can apply the backward Euler scheme

$$X_{k+1} = X_k + f(X_{k+1})\Delta t + \lambda \Delta w_{k+1} \tag{5}$$

and derive a strong convergence rate of order one for the maximum L^p-error in the discretization points, see Theorem 2.7 in Sect. 2.3.

2.2 Backward Euler–Maruyama scheme

In this section we focus on the numerical approximation of

$$dx(t) = f(x(t))dt + \sigma dw(t), \quad t \geq 0, \qquad x(0) = x_0 \tag{6}$$

by the backward Euler–Maruyama scheme

$$X_{k+1} = X_k + f(X_{k+1})\Delta t + \sigma \Delta w_{k+1}, \quad k = 0, 1, \ldots, \qquad X_0 = x_0. \tag{7}$$

We will work under the following assumption on the SDE itself:

Assumption 2.1 Let $-\infty \leq \alpha < \beta \leq \infty$ and assume that SDE (6) has a unique strong solution which takes values in the set $(\alpha, \beta) \subseteq \mathbb{R}$, i.e.

$$\mathbb{P}(x(t) \in (\alpha, \beta), \ t \geq 0) = 1. \tag{8}$$

For the well-definedness of the drift-implicit Euler–Maruyama scheme we need the following assumption on the drift coefficient:

Assumption 2.2 The function $f : (\alpha, \beta) \to \mathbb{R}$ is continuous. Moreover, there exists a constant $K \in \mathbb{R}$ such that

$$(x - y)(f(x) - f(y)) \leq K \, |x - y|^2 \tag{9}$$

for all $x, y \in (\alpha, \beta)$.

The Feller test, see e.g. Theorem V.5.29 in [23], gives that condition (8) is equivalent to

$$\lim_{x \to \alpha_+} v(x) = \infty, \qquad \lim_{x \to \beta_-} v(x) = \infty$$

with

$$v(x) = \frac{2}{\sigma^2} \int_{x_0}^{x} \int_{x_0}^{\zeta_2} \frac{p'(\zeta_2)}{p'(\zeta_1)} d\zeta_1 d\zeta_2$$

and the scale function

$$p(x) = \int_{x_0}^{x} \exp\left(-2 \int_{x_0}^{\xi} \frac{f(u)}{\sigma^2} \, du \right) d\xi, \qquad x \in (\alpha, \beta).$$

Note that v can be rewritten as

$$v(x) = \frac{2}{\sigma^2} \int_{x_0}^{x} \int_{x_0}^{\zeta_2} \exp\left(-2 \int_{\zeta_1}^{\zeta_2} \frac{f(u)}{\sigma^2} \, du \right) d\zeta_1 d\zeta_2.$$

From the above condition on v we can directly characterize the behavior of f in the case of a finite α or β, i.e. $\alpha > -\infty$ resp. $\beta < \infty$. In this case we have

$$\lim_{x \to \alpha_+} \int_{x}^{x_0} \exp\left(2 \int_{x}^{\zeta_1} \frac{f(u)}{\sigma^2} \, du \right) d\zeta_1 = \infty \tag{10}$$

and

$$\lim_{x \to \beta_-} \int_{x_0}^{x} \exp\left(-2 \int_{\zeta_1}^{x} \frac{f(u)}{\sigma^2} \, du \right) d\zeta_1 = \infty. \tag{11}$$

Now consider (10) and assume that $\limsup_{x \to \alpha_+} f(x) \neq \infty$. However, if this would be true, the expression in (10) would be finite due to the continuity of f. Using a similar argument for (11) we obtain

$$\limsup_{x \to \alpha_+} f(x) = \infty, \qquad \liminf_{x \to \beta_-} f(x) = -\infty. \tag{12}$$

The drift-implicit Euler scheme is well defined if

$$X_{k+1} - f(X_{k+1})\Delta t = X_k + \sigma \Delta w_{k+1}$$

has a unique solution for $k = 0, 1, \ldots$. This is guaranteed by the following result:

Lemma 2.3 *Let Assumptions 2.1 and 2.2 hold and let* $K\Delta t < 1$. *Moreover set*

$$G(x) = x - f(x)\Delta t, \quad x \in (\alpha, \beta).$$

Then for any $c \in \mathbb{R}$ there exists a unique $x \in (\alpha, \beta)$ such that $G(x) = c$.

Proof The result follows, if we can show that the function G is continuous, coercive and strictly monotone on (α, β), see [31]. However, due to Assumption 2.2 the function G is continuous on (α, β). Moreover, since

$$\begin{aligned}(x-y)(G(x)-G(y)) &= (x-y)^2 - (x-y)(f(x)-f(y))\Delta t \\ &\geq (1-K^+\Delta t)(x-y)^2 > 0 \end{aligned} \tag{13}$$

by (9) (with $K^+ = \max\{0, K\}$) the required strict monotonicity is obtained. To finish the proof, it remains to show that

$$\lim_{x\to\alpha_+} G(x) = -\infty, \qquad \lim_{x\to\beta_-} G(x) = \infty.$$

By the monotonicity of G, we have

$$\liminf_{x\to\alpha_+} G(x) = \lim_{x\to\alpha_+} G(x)$$

and

$$\limsup_{x\to\beta_-} G(x) = \lim_{x\to\beta_-} G(x).$$

Now we have to distinguish between finite and infinite boundaries α and β.

(i) $\alpha = -\infty$: Eq. (13) implies

$$(x-x_0)(G(x)-G(x_0)) \geq (1-K^+\Delta t)(x-x_0)^2$$

and hence

$$G(x) \leq G(x_0) + (1-K^+\Delta t)(x-x_0)$$

for $x < x_0$, which gives $\liminf_{x\to\alpha_+} G(x) = -\infty$.

(ii) $\alpha > -\infty$: Here $\liminf_{x\to\alpha_+} G(x) = -\infty$ follows from (12).

(iii) $\beta = \infty$: Eq. (13) implies

$$G(x) \geq G(x_0) + (1-K^+\Delta t)(x-x_0)$$

for $x > x_0$ and thus $\limsup_{x\to\beta_-} G(x) = \infty$.

(iv) $\beta < \infty$: Here $\limsup_{x\to\beta_-} G(x) = \infty$ follows again from (12).

□

Note that for $K \leq 0$ there is no restriction on Δt.

For completeness, we state here a well known discrete version of Gronwall's Lemma:

Lemma 2.4 *Let* $\Delta t > 0$ *and let* $g_n, \lambda_n \in \mathbb{R}$, $n \in \mathbb{N}$, *and* $\eta \geq 0$ *be given. Moreover, assume that* $1 - \eta \Delta t > 0$ *and* $1 + \lambda_n > 0$, $n \in \mathbb{N}$. *Then, if* $a_n \in \mathbb{R}$, $n \in \mathbb{N}$, *satisfies* $a_0 = 0$ *and*

$$a_{n+1} \leq a_n(1 + \lambda_n) + \eta a_{n+1} \Delta t + g_{n+1}, \qquad n = 0, 1, \ldots,$$

then this sequence also satisfies

$$a_n \leq \frac{1}{(1 - \eta \Delta t)^n} \sum_{j=0}^{n-1} (1 - \eta \Delta t)^j g_{j+1} \prod_{l=j+1}^{n-1} (1 + \lambda_l), \qquad n = 0, 1, \ldots.$$

Under the above assumptions we have the following moment bounds for the SDE and the BEM scheme:

Lemma 2.5 *Let* $T > 0$ *and let Assumptions 2.1 and 2.2 hold. Then we have*

$$\mathbb{E} \sup_{t \in [0,T]} |x(t)|^q < \infty$$

for all $q \geq 1$. *If additionally* $2K \Delta t < \zeta$ *for some* $\zeta < 1$, *then for all* $q \geq 1$ *there exist a constant* $C_q > 0$, *which are independent of* Δt, *such that*

$$\mathbb{E} \sup_{k=0,\ldots,\lceil T/\Delta t \rceil} |X_k|^q \leq C_q.$$

Proof The first assertion can be shown by a straightforward modification of the proof of Lemma 3.2 in [13].

For the proof of the second assertion, we will denote constants which are independent of Δt and whose particular value is not important by c regardless of their value. Now let $x^* \in (\alpha, \beta)$ and rewrite BEM as

$$X_{k+1} - x^* - f(x^*)\Delta t = X_k - x^* + (f(X_{k+1}) - f(x^*))\Delta t + \sigma \Delta w_{k+1}.$$

Multiplying with $X_{k+1} - x^*$ and using the one-sided Lipschitz condition on f yields that

$$\begin{aligned}(X_{k+1} - x^*)^2 \leq &\left((X_k - x^*) + f(x^*)\Delta t + \sigma \Delta w_{k+1}\right)(X_{k+1} - x^*) \\ &+ K^+ \Delta t (X_{k+1} - x^*)^2.\end{aligned}$$

where $K^+ = \max\{0, K\}$. Moreover, setting $u_k = X_k - x^*$ using $ab \leq \frac{1}{2}(a^2 + b^2)$ and rearranging the terms gives

$$u_{k+1}^2 \leq 2K^+ \Delta t u_{k+1}^2 + (u_k + f(x^*)\Delta t + \sigma \Delta w_{k+1})^2 .$$

Using again $ab \leq \frac{1}{2}(a^2 + b^2)$ yields

$$\begin{aligned}\left(u_k + f(x^*)\Delta t + \sigma \Delta w_{k+1}\right)^2 &\leq u_k^2(1+\Delta t) + 2\sigma u_k \Delta w_{k+1} \\ &+ 2\sigma^2|\Delta w_{k+1}|^2 + 2|f(x^*)|^2 \Delta t(1+\Delta t),\end{aligned}$$

hence we have

$$u_{k+1}^2 \leq u_k^2(1+\Delta t) + 2K^+ u_{k+1}^2 \Delta t + 2\sigma u_k \Delta w_{k+1} + 2\sigma^2 |\Delta w_{k+1}|^2 + c\Delta t.$$

Note that $1 - 2K^+\Delta t \in (0, 1]$ and

$$\sup_{\Delta t \in (0, \zeta/(2K^+))} \sup_{k=1,\ldots,\lceil T/\Delta t \rceil} \frac{(1+\Delta t)^k}{(1-2K^+\Delta t)^k} < \infty, \tag{14}$$

$$\sup_{\Delta t \in (0, \zeta/(2K^+))} \sup_{k=1,\ldots,\lceil T/\Delta t \rceil} \frac{(1-2K^+\Delta t)^k}{(1+\Delta t)^k} < \infty. \tag{15}$$

So, the above discrete version of Gronwall's Lemma with $\lambda_n = \Delta t$, $\eta = 2K^+$ and $g_{k+1} = 2u_k \Delta w_{k+1} + 2\sigma^2|\Delta w_{k+1}|^2 + c\Delta t$ now yields

$$\begin{aligned} u_k^2 &\leq c + 2\frac{(1+\Delta t)^k}{(1-2K^+\Delta t)^k} \underbrace{\sum_{j=0}^{k-1} \frac{(1-2K^+\Delta t)^j}{(1+\Delta t)^j} u_j \Delta w_{j+1}}_{=M_k^{(1)}} \\ &+ 2\frac{(1+\Delta t)^k}{(1-2K^+\Delta t)^k} \underbrace{\sum_{j=0}^{k-1} \frac{(1-2K^+\Delta t)^j}{(1+\Delta t)^j} \sigma^2 \left(|\Delta w_{j+1}|^2 - \Delta t\right)}_{=M_k^{(2)}}, \end{aligned} \tag{16}$$

from which we obtain easily by induction that

$$\sup_{k=0,\ldots,\lceil T/\Delta t \rceil} \mathbb{E} u_k^2 < \infty.$$

Using this it can be easily checked that the processes $M_k^{(i)}$, $i = 1, 2$, are square-integrable martingales with respect to the filtration $\mathcal{F}_{k\Delta t}$, $k = 0, 1, \ldots$. Hence Doob's inequality and straightforward calculations give for any $q \geq 1$ that

$$\mathbb{E} \sup_{k=0,\ldots,\lceil T/\Delta t \rceil} |M_k^{(1)}|^{2q} \leq c \cdot \sup_{k=0,\ldots,\lceil T/\Delta t \rceil} \mathbb{E}|u_k|^{2q}$$

and

$$\mathbb{E} \sup_{k=0,\ldots,\lceil T/\Delta t \rceil} |M_k^{(2)}|^{2q} < \infty.$$

So using (16), (14) and (15) we obtain

$$\mathbb{E} \sup_{k=0,\ldots,\lceil T/\Delta t\rceil} u_k^{4q} \leq c + c \cdot \sup_{k=0,\ldots,\lceil T/\Delta t\rceil} \mathbb{E} u_k^{2q}$$

for $\Delta t < \zeta/(2K^+)$ and the assertion follows now by an induction argument in $q \in \mathbb{N}$. □

2.3 Main result

Here we prove our general theorem on the strong convergence of the numerical scheme (7) to the solution of SDE (6). So far we have assumed (Assumptions 2.1 and 2.2) that the SDE solution does not leave the domain (α, β) and that the drift is one-sided Lipschitz continuous.

Assumption 2.6 Let $T > 0$ and $p \geq 2$. We assume that the drift coefficient $f : (\alpha, \beta) \to \mathbb{R}$ of SDE (6) is twice continuously differentiable and satisfies

$$\sup_{t\in[0,T]} \mathbb{E}\left|f'(x(t))\right|^p + \sup_{t\in[0,T]} \mathbb{E}\left|(f'f)(x(t)) + \frac{\sigma^2}{2} f''(x(t))\right|^p < \infty.$$

Theorem 2.7 *Let $T > 0$, $\zeta \in (0, 1)$ and Assumptions 2.1, 2.2 and 2.6 with $p \geq 2$ hold. Then, for $2K\Delta t < \zeta$ and all $q \in [1, p]$ there exists a constant $C_{p,q} > 0$ (independent of Δt) such that*

$$\mathbb{E}\left[\sup_{k=0,\ldots,\lceil T/\Delta t\rceil} |x(k\Delta t) - X_k|^q\right] \leq C_{p,q} \cdot \Delta t^q. \tag{17}$$

Proof Recall that we will denote constants which are independent of Δt and whose particular value is not important by c regardless of their value. By applying Itô's formula on $f(x(t))$ we have

$$x((k+1)\Delta t) = x(k\Delta t) + \int_{k\Delta t}^{(k+1)\Delta t} f(x((k+1)\Delta t))dt + \sigma \int_{k\Delta t}^{(k+1)\Delta t} dw(t) + R_{k+1} \tag{18}$$

where

$$\begin{aligned} R_{k+1} = &- \int_{k\Delta t}^{(k+1)\Delta t} \int_{t}^{(k+1)\Delta t} \left((f'f)(x(s)) + \frac{\sigma^2}{2} f''(x(s))\right) ds dt \\ &- \sigma \int_{k\Delta t}^{(k+1)\Delta t} \int_{t}^{(k+1)\Delta t} f'(x(s)) dw(s) dt. \end{aligned} \tag{19}$$

We can decompose R_s as $R_s = R_s^{(1)} + R_s^{(2)}$ with

$$R_s^{(1)} = -\int_{(s-1)\Delta t}^{s\Delta t} \int_t^{s\Delta t} \left((f'f)(x(u)) + \frac{\sigma^2}{2} f''(x(u)) \right) du dt,$$

$$R_s^{(2)} = -\sigma \int_{(s-1)\Delta t}^{s\Delta t} \int_t^{s\Delta t} f'(x(u)) dw(u) dt.$$

Using Eqs. (18) and (7) we have

$$x((k+1)\Delta t) - X_{k+1} = x(k\Delta t) - X_k + [f(x((k+1)\Delta t)) - f(X_{k+1})]\Delta t + R_{k+1}$$

and thus

$$\begin{aligned}&(x((k+1)\Delta t) - X_{k+1} - [f(x((k+1)\Delta t)) - f(X_{k+1})]\Delta t)^2 \\ &\quad = (x(k\Delta t) - X_k + R_{k+1})^2 .\end{aligned}$$

We arrive at

$$\begin{aligned}&|x((k+1)\Delta t) - X_{k+1}|^2 - |x(k\Delta t) - X_k|^2 \\ &\quad = 2(x((k+1)\Delta t) - X_{k+1})[f(x((k+1)\Delta t)) - f(X_{k+1})]\Delta t \\ &\qquad - [f(x((k+1)\Delta t)) - f(X_{k+1})]^2 \Delta t^2 + 2(x(k\Delta t) - X_k)R_{k+1} + R_{k+1}^2 .\end{aligned}$$

Using the one-sided Lipschitz assumption on f we obtain

$$\begin{aligned}(1 - 2K^+\Delta t)\, |x((k+1)\Delta t) - X_{k+1}|^2 \le |x(k\Delta t) - X_k|^2 + 2(x(k\Delta t) - X_k)R_{k+1} \\ + R_{k+1}^2 .\end{aligned}$$

Let us define $e_k = x(k\Delta t) - X_k$ and $\gamma_{\Delta t} = (1 - 2K^+\Delta t)$. Note that $\gamma_{\Delta t} \in (0, 1]$ and

$$\sup_{\Delta t \in (0, \zeta/(2K^+))} \ \sup_{k=1,\ldots,\lceil T/\Delta t \rceil} \gamma_{\Delta t}^{-k} < \infty. \tag{20}$$

Now, Lemma 2.4 yields

$$e_k^2 \le 2 \sum_{s=0}^{k-1} \gamma_{\Delta t}^{s-k} e_s R_{s+1} + \sum_{s=0}^{k-1} \gamma_{\Delta t}^{s-k} R_{s+1}^2 . \tag{21}$$

Since $\mathbb{E}\left[R_{s+1}^{(2)} \,\middle|\, \mathcal{F}_{s\Delta t} \right] = 0$, we have that

$$\sum_{s=0}^{k-1} \gamma_{\Delta t}^{s} e_s R_{s+1}^{(2)}, \qquad k = 0, \ldots, \lceil T/\Delta t \rceil,$$

is a martingale and the Burkholder–Davis–Gundy inequality implies that

$$\mathbb{E}\left[\sup_{k=0,\dots,\ell}\left|\sum_{s=0}^{k-1}\gamma_{\Delta t}^{s} e_s R_{s+1}^{(2)}\right|^q\right] \le c\,\mathbb{E}\left(\sum_{s=0}^{\ell-1}\gamma_{\Delta t}^{2s}\,|e_s|^2\left|R_{s+1}^{(2)}\right|^2\right)^{q/2}$$

for any $q \ge 1$ and $\ell = 0,\dots,\lceil T/\Delta t\rceil$. By (20) and $|\gamma_{\Delta t}| \le 1$ we have

$$\mathbb{E}\left[\sup_{k=0,\dots,\ell}\left|\sum_{s=0}^{k-1}\gamma_{\Delta t}^{s-k} e_s R_{s+1}^{(2)}\right|^q\right] \le c\,\mathbb{E}\left(\sum_{s=0}^{\ell-1}|e_s|^2\left|R_{s+1}^{(2)}\right|^2\right)^{q/2}$$

for any $q \ge 1$. Using this and Jensen's inequality in (21) we now arrive at

$$\begin{aligned}\mathbb{E}\left[\sup_{k=0,\dots,\ell}|e_k|^{2q}\right] &\le c\,\lceil T/\Delta t\rceil^{q/2-1}\sum_{s=0}^{\ell-1}\mathbb{E}|e_s|^q\left|R_{s+1}^{(2)}\right|^q\\ &\quad + c\,\lceil T/\Delta t\rceil^{q-1}\sum_{s=0}^{\ell-1}\mathbb{E}|e_s|^q\left|R_{s+1}^{(1)}\right|^q + c\,\lceil T/\Delta t\rceil^{q-1}\sum_{s=0}^{\ell-1}\mathbb{E}|R_{s+1}|^{2q}. \qquad (22)\end{aligned}$$

Now, Assumption 2.6, Jensen's inequality and the Burkholder–Davis–Gundy inequality give that

$$\mathbb{E}\left|R_{k+1}^{(2)}\right|^m \le c\,\Delta t^{3m/2} \qquad (23)$$

and

$$\mathbb{E}\left|R_{k+1}^{(1)}\right|^m \le c\,\Delta t^{2m} \qquad (24)$$

for all $m \le p$. Thus, the Cauchy–Schwarz inequality and Assumption 2.6 yield that

$$\mathbb{E}|e_s|^q\left|R_{s+1}^{(2)}\right|^q \le c\left(\mathbb{E}|e_s|^{2q}\right)^{1/2}\Delta t^{3q/2}.$$

Hence Young's inequality implies

$$c\,\lceil T/\Delta t\rceil^{q/2-1}\sum_{s=0}^{\ell-1}\mathbb{E}|e_s|^q\left|R_{s+1}^{(2)}\right|^q \le c\sum_{s=0}^{\ell-1}\mathbb{E}\,|e_s|^{2q}\,\Delta t + c\Delta t^{2q}.$$

Similar we also obtain

$$c\,\lceil T/\Delta t\rceil^{q-1}\sum_{s=0}^{\ell-1}\mathbb{E}|e_s|^q\left|R_{s+1}^{(1)}\right|^q \le c\sum_{s=0}^{\ell-1}\mathbb{E}\,|e_s|^{2q}\,\Delta t + c\Delta t^{2q}.$$

Since finally

$$c\,\lceil T/\Delta t\rceil^{q-1} \sum_{s=0}^{\ell-1} \mathbb{E}|R_{s+1}|^{2q} \le c\Delta t^{2q},$$

by inserting these three estimates in (22) we end up with

$$\mathbb{E}\left[\sup_{k=0,\dots,\ell} |e_k|^{2q}\right] \le c \sum_{s=0}^{\ell-1} \mathbb{E}\,|e_s|^{2q}\,\Delta t + c\Delta t^{2q}$$

and Gronwall's Lemma completes now the proof of the assertion for $2 \le 2q \le p$. The case $2q \in [1, 2)$ follows now by Lyapunov's inequality. □

The above result and Lemma 2.5 now give convergence (without a rate) in all L^q-norms:

Corollary 2.8 *Under the assumptions of Theorem 2.7 we have*

$$\lim_{\Delta t \to 0} \mathbb{E}\left[\sup_{k=0,\dots,\lceil T/\Delta t\rceil} |x(k\Delta t) - X_k|^q\right] = 0 \tag{25}$$

for all $q \ge 1$.

Remark 2.9 In [2], independently of the research in this paper, a similar result to Theorem 2.7 is established for the case $D = (\alpha, \infty)$ and drift functions $f : D \to \mathbb{R}$, which are twice continuously differentiable and satisfy the monotone condition

$$f(x) - f(y) \le K(x - y), \qquad x \ge y > \alpha$$

for some $K \in \mathbb{R}$. Using a continuous extension of BEM and the relation

$$f(x) - f(y) = \beta_{x,y}(x - y)$$

for some $\beta_{x,y} \le K$ for the error analysis, Alfonsi obtains the bound (17) under the assumption

$$\mathbb{E}\left(\int_0^T \left|f'(x(t))\right|^2 dt\right)^{p/2} + \mathbb{E}\left(\int_0^T \left|(f'f)(x(t)) + \frac{\sigma^2}{2} f''(x(t))\right| dt\right)^p < \infty \tag{26}$$

for $p \ge 1$. This result is then applied to the CIR and CEV process, i.e. Propositions 3.1 and 3.3 are obtained both for $p \ge 1$.

Since the monotone condition is equivalent to the one-sided Lipschitz condition the main result of this article and [2] is identical for $p \ge 2$ and $D = (\alpha, \infty)$. The main differences between both works are as follows:

- In our proof, we are working with the discrete squared error and exploit the one-sided Lipschitz property

Why 单侧

$$(x-y)(f(x)-f(y)) \leq K|x-y|^2, \qquad x, y \in (\alpha, \beta).$$

  Hence we need in Assumption 2.6 the condition $p \geq 2$ instead of $p \geq 1$ as in (26) and obtain in Theorem 2.7 order one convergence in the L^p-norm for $p \geq 2$ instead of $p \geq 1$. Note that efficient Multilevel Monte Carlo simulation rely on good L^2-convergence properties.
- We work on (α, β) instead of (α, ∞) and can thus also treat the Wright–Fisher SDE and similar equations. (However the proof of [2] could be modified to this case).
- We also derive bounds for the inverse moments of the BEM scheme, see the subsection below. Hence we are also able to cover SDEs like the Heston-3/2 volatility and the Aït-Sahlia model.

2.4 Boundedness of inverse moments of BEM

If the drift coefficient has an even more specific structure, see the assumption below, we can also control the inverse moments of BEM. For the Heston-3/2 volatility and also the Aït-Sahalia model this will be helpful later on as in these cases the inverse of the Lamperti transformation is of the form $(0, \infty) \ni x \mapsto x^{-\alpha} \in (0, \infty)$ for some $\alpha > 0$.

For the BEM scheme we can write

$$f(X_{k+1}) = \frac{1}{\Delta t}\left(X_{k+1} - X_k - \sigma \Delta w_{k+1}\right). \tag{27}$$

Proceeding as in the proof of Theorem 2.7 we also have

$$f(x(t_{k+1})) = \frac{1}{\Delta t}\left(x(t_{k+1}) - x(t_k) - \sigma \Delta w_{k+1} - R_{k+1}\right), \tag{28}$$

with R_{k+1} given by (19). This can be used to derive the following result:

Lemma 2.10 *Let $T > 0$ and $p \geq 2$. Moreover, let the assumptions of Theorem 2.7 hold. Then there exist constants $C_p^{(1)}, C_p^{(2)} > 0$ such that we have*

$$\sup_{k=0,\ldots,\lceil T/\Delta t \rceil} \mathbb{E}|f(X_k)|^p \leq C_p^{(1)} \left(1 + \sup_{t \in [0,T]} \mathbb{E}|f(x(t))|^p\right)$$

and

$$\mathbb{E}\left[\sup_{k=0,\ldots,\lceil T/\Delta t \rceil} |f(X_k)|^p\right] \leq C_p^{(2)} \left(1 + \mathbb{E} \sup_{t \in [0,T]} |f(x(t))|^p\right).$$

Proof We only prove the second assertion, the proof of the first assertion is similar. Using (27), (28), (17), (23) and (24) we have

$$\mathbb{E}\left[\sup_{k=0,\dots,\lceil T/\Delta t\rceil} |f(x(t_k)) - f(X_k)|^p\right] \le c.$$

The triangle inequality completes the proof. □

Later on we are interested in obtaining estimates on the inverse moments of BEM for $D = (0,\infty)$, so we introduce the following mild assumption on the function f in this case.

Assumption 2.11 Assume that the drift coefficient $f : (0,\infty) \to \mathbb{R}$ has the structure

$$f(x) = \frac{c_1}{x^{m_1}} + h(x), \qquad x \in (0,\infty)$$

where

$$|h(x)| \le c_2 \cdot (1 + |x|^{m_2}), \qquad x \in (0,\infty)$$

for some $c_1, c_2 > 0$ and $m_1, m_2 > 0$.

Since $\lim_{x\to 0^+} f(x) = \infty$ by (12), the above assumption essentially characterizes the behavior on a fractional power scale of f for $x \to 0$ and $x \to \infty$.

Remark 2.12 If we additionally impose Assumption 2.11 in Lemma 2.10, then

$$\mathbb{E}\left[\sup_{k=0,\dots,\lceil T/\Delta t\rceil} |X_k|^{-m_1 p}\right] < C_p^{(3)}\left(1 + \mathbb{E}\left[\sup_{t\in[0,T]} |x(t)|^{-m_1 p}\right]\right)$$

for some $C_p^{(3)} > 0$.

The next lemma helps to estimate the right hand side of the above inequality.

Lemma 2.13 *Let $p \ge 2$. Let the assumptions of Lemma 2.5 hold and in addition let the drift of SDE (6) satisfy Assumption 2.11. Then there exists a constant $C_p > 0$ such that we have*

$$\mathbb{E}\left[\sup_{t\in[0,T]} |x(t)|^{-p}\right] \le C_p\left(1 + \sup_{t\in[0,T]} \mathbb{E}|x(t)|^{-(p+2)}\right).$$

Proof Let $\{\alpha_n\}_{n\in\mathbb{N}}$ and $\{\beta_n\}_{n\in\mathbb{N}}$ be such that $\alpha_n \searrow 0$ and $\beta_n \nearrow \infty$, when $n \to \infty$. Let us choose n_0 such that $\alpha_{n_0} \le x(0) \le \beta_{n_0}$. Then for $n \ge n_0$, we define the stopping time $\tau_n = \{t > 0 : x(t) \notin (\alpha_n, \beta_n)\}$. By Itô's lemma we have

$$|x(t \wedge \tau_n)|^{-p} = |x(0)|^{-p} + \int_0^t \Big(g(x(s)) - p|x(s)|^{-(p+1)} h(x(s)) \Big) \mathbf{1}_{[0,\tau_n]}(s) ds$$
$$- p\sigma \int_0^t |x(s)|^{-(p+1)} \mathbf{1}_{[0,\tau_n]}(s) dw(s),$$

where $g(x) = -c_1 p|x|^{-(p+1+m_1)} + \frac{\sigma^2 p(p+1)}{2} |x|^{-(p+2)}$. Observe that for $m_1 > 1$, $\lim_{x \searrow 0} g(x) = -\infty$ and $\lim_{x \to \infty} g(x) = 0$, hence in that case there exists a $c > 0$ such that

$$\sup_{x>0} g(x) \le c.$$

If $m_1 \in (0, 1]$ then there exists a $c > 0$ such that

$$|g(x)| \le c(1 + |x|^{-(p+2)}).$$

By Assumption 2.11 and the Burkholder–Davis–Gundy inequality we now have

$$\begin{aligned}
\mathbb{E}\left[\sup_{t\in[0,T]} |x(t \wedge \tau_n)|^{-p} \right] &\le |x(0)|^{-p} \\
&\quad + \int_0^T c \Big(1 + \mathbb{E}|x(s)|^{-(p+2)} + \mathbb{E}|x(s)|^{-(p+1)+m_2} \Big) ds \\
&\quad + c\, \mathbb{E} \left(\int_0^T |x(s)|^{-2(p+1)} \mathbf{1}_{[0,\tau_n]}(s) ds \right)^{1/2} \\
&\le |x(0)|^{-p} + \int_0^T c \Big(1 + \mathbb{E}|x(s)|^{-(p+2)} \Big) ds \\
&\quad + c\, \mathbb{E} \left(\sup_{t\in[0,T]} |x(t \wedge \tau_n)|^{-p} \int_0^T |x(s)|^{-(p+2)} ds \right)^{1/2}.
\end{aligned}$$

Applying Young's inequality to the last summand of the above inequality now yields

$$\mathbb{E}\left[\sup_{t\in[0,T]} |x(t \wedge \tau_n)|^{-p} \right] \le c \left(1 + \sup_{t\in[0,T]} \mathbb{E}|x(t)|^{-(p+2)} \right).$$

and Fatou's Lemma completes the proof. □

3 Examples

In this section, we will apply our main result to several examples. We will denote the numerical method $Y_k = F^{-1}(X_k)$, $k = 0, 1, \ldots$, where X_k, $k = 0, 1, \ldots,$ is given by (5), as *Lamperti-backward Euler* (LBE) approximation of SDE (4). Note that in all examples except the Wright–Fisher equation the Lamperti transform is of the type $(0, \infty) \ni x \mapsto x^\alpha \in (0, \infty)$ with $\alpha \neq 0$. Hence the inverse transformation is $(0, \infty) \ni x \mapsto x^{1/\alpha} \in (0, \infty)$ and the error propagation in the inverse transformation can be controlled by using Lemma 2.5 (or Remark 2.12) and Lemma 2.13.

Moreover, we will say that this method is *p-strongly convergent with order one*, if

$$\mathbb{E} \sup_{k=0,\ldots,\lceil T/\Delta t \rceil} |y(t_k) - Y_k|^p \leq C_p \cdot \Delta t^p.$$

Finally, constants whose particular value is not important will be again denoted by c.

3.1 CIR process

Recall that the Cox–Ingersoll–Ross process is given by the SDE

$$dy(t) = \kappa(\theta - y(t))dt + \sigma\sqrt{y(t)}dw(t), \quad t \geq 0, \qquad y(0) > 0. \tag{29}$$

If $2\kappa\theta \geq \sigma^2$, then we have $D = (0, \infty)$ and Assumption 2.1 holds for $(\alpha, \beta) = (0, \infty)$. Moreover, recall that the transformed SDE using $F(y) = \sqrt{y}$ reads as

$$dx(t) = f(x(t))dt + \frac{1}{2}\sigma dw(t), \quad t \geq 0, \qquad x(0) = \sqrt{y(0)} \tag{30}$$

with

$$f(x) = \frac{1}{2}\kappa\left(\theta_v x^{-1} - x\right), \qquad x > 0$$

where $\theta_v = \theta - \frac{\sigma^2}{4\kappa}$ and the BEM scheme is given by

$$X_{k+1} = X_k + f(X_{k+1})\Delta t + \frac{1}{2}\sigma\Delta w_{k+1}, \qquad k = 0, 1, \ldots \tag{31}$$

with $X_0 = x(0)$. Straightforward calculations give

$$(x - y)(f(x) - f(y)) \leq -\frac{1}{2}\kappa(x - y)^2, \qquad x, y > 0,$$

so Assumption 2.2 holds with $K = -\kappa/2$. Observe also that

$$f'(x) = -\frac{1}{2}\kappa(\theta_v x^{-2} + 1)$$

and

$$(f'f)(x) + \frac{\sigma^2}{2} f''(x) = -\frac{\kappa^2}{4}(\theta_v^2 x^{-3} - x) + \frac{1}{2}\kappa\theta_v x^{-3}\sigma^2.$$

So, for Assumption 2.6 to hold we need to control

$$\sup_{0\le t\le T} \mathbb{E}[x(t)^{-3p}] = \sup_{0\le t\le T} \mathbb{E}[y(t)^{-\frac{3}{2}p}].$$

Since

$$\sup_{0\le t\le T} \mathbb{E}[y(t)^q] < \infty \quad \text{for} \quad q > -\frac{2k\theta}{\sigma^2}, \tag{32}$$

see e.g. [7], Assumption 2.6 holds if $p < \frac{4}{3}\frac{k\theta}{\sigma^2}$ and as a consequence Theorem 2.7 can be applied for $2 \le p < \frac{4}{3}\frac{k\theta}{\sigma^2}$.

In order to approximate the original CIR process observe that

$$(x(t_k)^2 - X_k^2) = (x(t_k) + X_k)(x(t_k) - X_k).$$

Let $\varepsilon > 0$ such that $p(1+\varepsilon) < \frac{4}{3}\frac{k\theta}{\sigma^2}$. Then Hölder's inequality gives

$$\begin{aligned}\mathbb{E}\left[\sup_{0\le k\le\lceil T/\Delta t\rceil} \left|x(t_k)^2 - X_k^2\right|^p\right] \le{}& \left(\mathbb{E}\left[\sup_{0\le k\le\lceil T/\Delta t\rceil} |x(t_k) + X_k|^{p\frac{1+\varepsilon}{\varepsilon}}\right]\right)^{p\frac{\varepsilon}{1+\varepsilon}} \\ &\times \left(\mathbb{E}\left[\sup_{0\le k\le\lceil T/\Delta t\rceil} |x(t_k) - X_k|^{p(1+\varepsilon)}\right]\right)^{\frac{1}{1+\varepsilon}}.\end{aligned}$$

Using Lemma 2.5 we obtain:

Proposition 3.1 *Let $T > 0$ and $2 \le p < \frac{4}{3}\frac{\kappa\theta}{\sigma^2}$. Then, the LBE approximation of the CIR process is p-strongly convergent with order one.*

In [2] the above result is established for $p \ge 1$, as already pointed out in Remark 2.9. Clearly, in the case when $2 \le \frac{4}{3}\frac{\kappa\theta}{\sigma^2}$, Proposition 3.1 holds for $1 \le p < \frac{4}{3}\frac{\kappa\theta}{\sigma^2}$.

3.2 Numerical experiment

Note that the unique solution to (31) is given by

$$X_{k+1} = \frac{1}{2+\kappa\Delta t}\left(X_k + \frac{1}{2}\sigma\Delta w_{k+1} + \sqrt{\left(X_k + \frac{1}{2}\sigma\Delta w_{k+1}\right)^2 + \kappa\theta_v\Delta t}\right)$$

with $\theta_v = \theta - \frac{\sigma^2}{4\kappa}$. Hence implicit schemes not necessarily increase the computational effort with comparison to classical explicit procedures. In our numerical experiment, we focus on the L^2-error at the endpoint $T = 1$, so we let

$$e_{\Delta t} = \mathbb{E}\,|x(T) - X_N|^2$$

where $\Delta t = T/N$ with $N \in \mathbb{N}$. For our numerical experiment we set $\theta = 0.125$, $\kappa = 2$, and $\sigma = 0.5$. This gives $\frac{2\kappa\theta}{\sigma^2} = 2$ and corresponds to the critical parameters for which Dereich et al. [7] established strong convergence of order one half up to a log-term for the linearly interpolated BEM (31) with respect to a uniform L^2-error criteria. Moreover, although theoretical results obtained in this paper impose slightly more restrictive conditions for parameters than those in Dereich et al., performed numerical experiments suggest that in practical simulations the condition $1 < p < \frac{2\kappa\theta}{\sigma^2} = 2$ suffices for uniform L^p-convergence with order one. Moreover, although an explicit solution to (30) is unknown, Theorem 1.1 in [7] guarantees that BEM strongly converges to the true solution. Therefore, it is reasonable to take BEM with a very small time step, we choose $\Delta t = 2^{-15}$, as a reference solution. We then compare it to BEM evaluated with $(2^4\Delta t, 2^5\Delta t, 2^6\Delta t, 2^7\Delta t)$ in order to estimate the rate of the L^2-convergence, for which we use a Monte Carlo procedure, i.e.

$$e_{\Delta t} \approx \frac{1}{10^3}\sum_{i=1}^{10^3}\left|x^{(i)}(T) - X_N^{(i)}\right|^2.$$

Here $x^{(i)}(T)$, $X_N^{(i)}$ are iid copies of $x(T)$, X_N. We plot $e_{\Delta t}$ against Δt on a log–log scale, i.e. if we assume that a power law relation $e_{\Delta t} = C\Delta t^q$ holds for some constant C and q, then we have $\log e_{\Delta t} = \log C + q \log \Delta t$. For our simulation, a least squares fit for $\log C$ and q yields the value 1.9332 for q with a least square residual of 0.016. Hence, our results are consistent with strong order of convergence equal to one (Fig. 1).

3.3 Heston-3/2 volatility

In [11] the inverse of a CIR process is used as a stochastic volatility process, which gives the so-called Heston-3/2 volatility

$$dy(t) = c_1 y(t)(c_2 - y(t))\,dt + c_3 y(t)^{3/2}\,dw(t), \quad t \geq 0, \qquad y(0) > 0 \tag{33}$$

where $c_1, c_2, c_3 > 0$. Using $F(y) = y^{-1/2}$ leads to

$$dx(t) = \left(\left(\frac{c_1}{2} + \frac{3c_3^2}{8}\right)x(t)^{-1} - \frac{c_1c_2}{2}x(t)\right)dt - \frac{c_3}{2}\,dw(t), \tag{34}$$

which coincides with (30) if we use a reflected Brownian motion, i.e. $-w$, which is still a Brownian motion, and

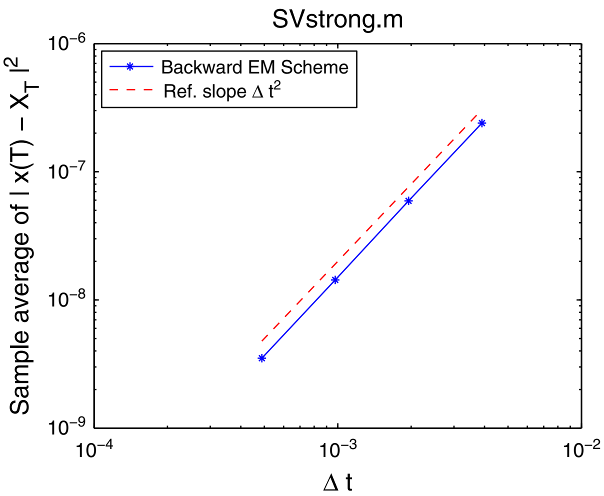

Fig. 1 Strong error plot for LBE applied to CIR.

$$\sigma = c_3, \qquad \theta = \frac{1}{c_2} + \frac{c_3^2}{c_1 c_2}, \qquad \kappa = c_1 c_2.$$

Hence we have the relation

$$\frac{\kappa\theta}{\sigma^2} = 1 + \frac{c_1}{c_3^2},$$

so Theorem 2.7 can applied here for $2 \le p < \frac{4}{3}(1 + \frac{c_1}{c_3^2})$. Note that the inverse of the CIR process has finite p-moments up to order $p < 2 + \frac{2c_1}{c_3^2}$, hence the Heston-3/2 volatility is one of the SDEs, which does not have finite p-moments for arbitrarily large $p \ge 1$.

Now, for transforming back we have to control the inverse moments of the BEM scheme for CIR. Here Remark 2.12 and 2.13 give

$$\mathbb{E} \sup_{k=0,\ldots,\lceil T/\Delta t \rceil} |X_k|^{-q} \le C_q \left(1 + \sup_{t\in[0,T]} \mathbb{E}|x(t)|^{-(q+2)}\right)$$

for $q < \frac{4}{3}(1 + \frac{c_1}{c_3^2})$. From the analysis of the CIR process we have that

$$\sup_{t\in[0,T]} \mathbb{E}|x(t)|^{-(q+2)} < \infty \quad \text{for} \quad q < 4\left(\frac{1}{2} + \frac{c_1}{c_3^2}\right),$$

and since $\frac{1}{3}(1+\frac{c_1}{c_3^2}) \leq \frac{1}{2} + \frac{c_1}{c_3^2}$ it follows that

$$\mathbb{E} \sup_{k=0,\ldots,\lceil T/\Delta t\rceil} |X_k|^{-q} < \infty \quad \text{for} \quad q < \frac{4}{3}\left(1+\frac{c_1}{c_3^2}\right).$$

To establish the convergence result for the LBM for the Heston-3/2 volatility note that

$$\left|\frac{1}{X_k^2} - \frac{1}{x(t_k)^2}\right| = \frac{X_k + x(t_k)}{X_k^2 x(t_k)^2}|X_k - x(t_k)| \leq c\left(|X_k|^{-3} + |x(t_k)|^{-3}\right)|X_k - x(t_k)|$$

and recall that

$$\mathbb{E}\left[\sup_{k=0,\ldots,\lceil T/\Delta t\rceil} |X_k - x(t_k)|^q\right] \leq C_q \cdot \Delta t^q \quad \text{for} \quad q < \frac{4}{3}\left(1+\frac{c_1}{c_3^2}\right)$$

by Theorem 2.7. Using Hölder's inequality to balance the bounds on the exponents gives

$$\begin{aligned}
&\mathbb{E}\left[\sup_{k=0,\ldots,\lceil T/\Delta t\rceil} \left|\frac{1}{X_k^2} - \frac{1}{x(t_k)^2}\right|^p\right] \\
&\leq c\left(\left(\mathbb{E} \sup_{k=0,\ldots,\lceil T/\Delta t\rceil} |X_k|^{-4p}\right)^{3/4} + \left(\mathbb{E} \sup_{k=0,\ldots,\lceil T/\Delta t\rceil} |x(t_k)|^{-4p}\right)^{3/4}\right) \\
&\quad \times \left(\mathbb{E} \sup_{k=0,\ldots,\lceil T/\Delta t\rceil} |X_k - x(t_k)|^{4p}\right)^{1/4}.
\end{aligned}$$

Choosing $p \geq 1$ such that $2 \leq 4p < \frac{4}{3}(1+\frac{c_1}{c_3^2})$ we obtain:

Proposition 3.2 *If $1 \leq p < \frac{1}{3} + \frac{1}{3}\frac{c_1}{c_3^2}$, then the LBE approximation of the Heston-3/2 process is p-strongly convergent with order one.*

3.4 CEV process

Another popular SDE in finance is the mean reverting constant elasticity of variance process ([5]) given by

$$dy(t) = \kappa(\theta - y(t))dt + \sigma y(t)^{\alpha} dw(t) \tag{35}$$

where $0.5 < \alpha < 1$, $\kappa, \theta, \sigma > 0$. By the Feller test we have $D = (0, \infty)$. Applying Itô's formula to the function $F(y(t)) = y(t)^{1-\alpha}$ we obtain that Assumption 2.1 holds with $(\alpha, \beta) = (0, \infty)$ and

$$dx(t) = f(x(t))dt + (1-\alpha)\sigma dw(t) \tag{36}$$

with

$$f(x) = (1-\alpha)\left(\kappa\theta x^{-\frac{\alpha}{1-\alpha}} - \kappa x - \frac{\alpha\sigma^2}{2}x^{-1}\right), \qquad x > 0.$$

Again we need to check the remaining assumptions. Since $\alpha > 0.5$, and consequently $\frac{1}{1-\alpha} > 2$, we have for

$$f'(x) = -\alpha\kappa\theta x^{-\frac{1}{1-\alpha}} - (1-\alpha)\kappa + (1-\alpha)\frac{\alpha\sigma^2}{2}x^{-2}, \qquad x > 0$$

that

$$\lim_{x\to 0} f'(x) = -\infty, \qquad \lim_{x\to\infty} f'(x) = -(1-\alpha)\kappa$$

and hence there exists a $c > 0$ such that

$$\sup_{x>0} f'(x) \leq c.$$

Now the mean value theorem implies

$$(x-y)(f(x)-f(y)) \leq c\,|x-y|^2, \qquad x, y > 0,$$

i.e. the drift coefficient is one-sided Lipschitz. We have moreover that

$$f''(x) = \frac{\alpha}{1-\alpha}\kappa\theta x^{-\frac{2-\alpha}{1-\alpha}} - (1-\alpha)\alpha\sigma^2 x^{-3}.$$

However, from [4] it is known that

$$\sup_{0\leq t\leq T} \mathbb{E}\,|y(t)|^p < \infty \qquad \text{for any} \quad p \in \mathbb{R},\ T > 0, \tag{37}$$

and therefore also Assumption 2.6 holds and Theorem 2.7 can be applied for any $p \geq 2$. For the back transformation note that the mean value theorem yields

$$\left| x^{\frac{1}{1-\alpha}} - y^{\frac{1}{1-\alpha}} \right| \leq \frac{1}{1-\alpha}\left(x^{\frac{\alpha}{1-\alpha}} + y^{\frac{\alpha}{1-\alpha}}\right)|x-y|.$$

Using Lemma 2.5 and the Lyapunov inequality for the case $p \in [1, 2)$ we have:

Proposition 3.3 *Let $p \geq 1$. The LBE approximation of the CEV process is p-strongly convergent with order one.*

3.5 Wright–Fisher diffusion

The Wright–Fisher SDE that originated from mathematical biology [8], and which recently is also gaining popularity in mathematical finance [15,27], reads as

$$dy(t) = (a - by(t))dt + \gamma\sqrt{|y(t)(1-y(t))|}\,dw(t), \qquad t \geq 0, \qquad y_0 \in (0,1)$$

with $a, b, \gamma > 0$. If

$$\frac{2a}{\gamma^2} \geq 1 \qquad \text{and} \qquad \frac{2(b-a)}{\gamma^2} \geq 1, \tag{38}$$

then this SDE has a unique strong solution with

$$\mathbb{P}(y(t) \in (0,1),\ t \geq 0) = 1,$$

see [24]. Using

$$F(y) = 2\arcsin(\sqrt{y}), \qquad y \in (0,1)$$

we obtain

$$dx(t) = f(x(t))dt + \gamma\,dw(t), \qquad t \geq 0, \qquad x_0 = 2\arcsin(\sqrt{y_0})$$

with

$$f(x) = \left(a - \frac{\gamma^2}{4}\right)\cot\left(\frac{x}{2}\right) - \left(b - a - \frac{\gamma^2}{4}\right)\tan\left(\frac{x}{2}\right), \qquad x \in (0,\pi).$$

Since

$$f'(x) = -\frac{1}{2}\left(a - \frac{\gamma^2}{4}\right)\left(1 + \cot^2\left(\frac{x}{2}\right)\right) - \frac{1}{2}\left(b - a - \frac{\gamma^2}{4}\right)\left(1 + \tan^2\left(\frac{x}{2}\right)\right)$$

the mean value theorem implies that Assumption 2.2 is satisfied with $K = 0$. Now note that

$$\begin{aligned} f''(x) = {} & \frac{1}{2}\left(a - \frac{\gamma^2}{4}\right)\cot\left(\frac{x}{2}\right)\left(1 + \cot^2\left(\frac{x}{2}\right)\right) \\ & -\frac{1}{2}\left(b - a - \frac{\gamma^2}{4}\right)\tan\left(\frac{x}{2}\right)\left(1 + \tan^2\left(\frac{x}{2}\right)\right) \end{aligned}$$

and

$$\begin{aligned}
&(f'f)(x(t)) + \frac{\gamma^2}{2} f''(x(t)) \\
&\quad = (f'f)\left(2\arcsin\left(\sqrt{y(t)}\right)\right) + \frac{\gamma^2}{2} f''\left(2\arcsin\left(\sqrt{y(t)}\right)\right).
\end{aligned}$$

Since

$$\cot(\arcsin(\sqrt{y(t)})) = \sqrt{\frac{1-y(t)}{y(t)}}, \qquad \tan(\arcsin(\sqrt{y(t)})) = \sqrt{\frac{y(t)}{1-y(t)}}$$

and $y(t) \in (0,1)$, we obtain

$$\left|f'(x(t))\right| \leq c\left(1 + |y(t)|^{-1} + |1-y(t)|^{-1}\right)$$

and

$$\left|(f'f)(x(t)) + \frac{\gamma^2}{2} f''(x(t))\right| \leq c\left(1 + |y(t)|^{-3/2} + |1-y(t)|^{-3/2}\right)$$

for some constant $c > 0$, depending only on $a, b, \gamma > 0$. Using Theorem 4.1 in [16] and establishing uniform convergence of the given series expressions in t using asymptotic bounds on the Jacobi polynomials we have that

$$\sup_{t\in[0,T]} \mathbb{E}|y(t)|^{q_1} < \infty$$

if $q_1 > -\frac{2a}{\gamma^2}$ and

$$\sup_{t\in[0,T]} \mathbb{E}|1-y(t)|^{q_2} < \infty$$

if $q_2 > -\frac{2(b-a)}{\gamma^2}$. So Assumption 2.6 is satisfied if

$$\frac{4}{3\gamma^2}\min\{a, b-a\} > p.$$

Now Theorem 2.7 gives

$$\mathbb{E}\left[\sup_{k=0,\ldots,\lceil T/\Delta t\rceil} |x(t_k) - X_k|^p\right] \leq C_p \cdot \Delta t^p \qquad \text{for} \qquad 2 \leq p < \frac{4}{3\gamma^2}\min\{a, b-a\}$$

and transforming back yields

$$\mathbb{E}\left[\sup_{k=0,\dots,\lceil T/\Delta t\rceil}\left|y(t_k)-\sin^2\left(\frac{X_k}{2}\right)\right|^p\right]\le C_p\cdot\Delta t^p.$$

under the same assumption on the parameters since the sine-function is bounded.

Proposition 3.4 *Let* $2\le p<\frac{4}{3\gamma^2}\min\{a,b-a\}$. *Then the LBE approximation of the Wright–Fisher process is* p*-strongly convergent with order one.*

Similarly as for the CIR process in the case when $\frac{4}{3\gamma^2}\min\{a,b-a\}>2$, the above proposition holds for $1\le p<\frac{4}{3\gamma^2}\min\{a,b-a\}$.

Recently a split-step method based on an Euler discretization of the drift and a balanced implicit method for the diffusion part has been introduced in [6]. This scheme is shown to be strongly convergent and boundary preserving without any restrictions on the parameter and numerical tests indicate the convergence order 1/2.

3.6 Aït-Sahalia model

Higham et al. analysed in [30] a backward Euler scheme for the Aït-Sahalia interest rate model

$$dy(t)=(\alpha_{-1}y(t)^{-1}-\alpha_0+\alpha_1y(t)-\alpha_2y(t)^r)dt+\sigma y(t)^\rho dw(t),\qquad(39)$$

where $\alpha_{-1},\alpha_0,\alpha_1,\alpha_2,\sigma$ are positive constants and $\rho,r>1$. In [30] it was established that

$$\mathbb{P}(y(t)\in(0,\infty),\ t\ge 0)=1.$$

Under the assumption $r+1>2\rho$ Higham et al. proved uniform L^p-convergence for any $p\ge 2$ for the backward Euler method directly applied to (39). However, their results did not reveal a rate of convergence. Here, using the Lamperti transformation approach we construct a scheme that strongly converges with rate one. We focus on the critical case with $r=2$ and $\rho=1.5$ which was not covered in [30]. Using $F(y)=y^{-1/2}$ we obtain

$$dx(t)=f(x(t))dt-\frac{1}{2}\sigma dw(t)$$

with

$$f(x)=\left(\frac{1}{2}\alpha_2+\frac{3}{8}\sigma^2\right)x^{-1}-\frac{1}{2}\alpha_1x+\frac{1}{2}\alpha_0x^3-\frac{1}{2}\alpha_{-1}x^5,\quad x>0.$$

We have for

$$f'(x) = -\left(\frac{1}{2}\alpha_2 + \frac{3}{8}\sigma^2\right)x^{-2} - \frac{1}{2}\alpha_1 + \frac{3}{2}\alpha_0 x^2 - \frac{5}{2}\alpha_{-1}x^4$$

that

$$\lim_{x\to 0} f'(x) = -\infty, \qquad \lim_{x\to\infty} f'(x) = -\infty$$

and hence there exists a $c > 0$ such that

$$\sup_{x>0} f'(x) < c.$$

Now the mean value theorem implies

$$(x-y)(f(x)-f(y)) \leq c\,|x-y|^2\,, \qquad x, y > 0,$$

i.e. the drift coefficient is one-sided Lipschitz. We have moreover that

$$f''(x) = \left(\alpha_2 + \frac{3}{4}\sigma^2\right)x^{-3} + 3\alpha_0 x - 10\,\alpha_{-1}x^3,$$

and

$$\left|(f'f)(x(t)) + \frac{\sigma^2}{8}f''(x(t))\right| \leq c\left(1 + |x(t)|^{-3} + |x(t)|^{20}\right).$$

Straightforward computations also give that

$$f(x) > g(x), \qquad x \in (0, 2x(0)]$$

where

$$g(x) := \left(\frac{1}{2}\alpha_2 + \frac{3}{8}\sigma^2\right)x^{-1} - \frac{1}{2}\alpha_2\beta x$$

with

$$\beta = \frac{\alpha_1}{\alpha_2} + 16\frac{\alpha_{-1}}{\alpha_2}x(0)^4.$$

Now a comparison result for SDEs, see e.g. Proposition V.2.18 and Exercise V.2.19 in [23], yields that almost surely

$$x(t) \geq x^{(1)}(t), \qquad t \in [0, \tau^{(1)}),$$

where

$$dx^{(1)}(t) = g(x^{(1)}(t))dt - \frac{1}{2}\sigma dw(t), \quad t \geq 0, \qquad x^{(1)}(0) = x(0)$$

and

$$\tau^{(1)} = \inf\{t \in [0, T] : x(t) > 2x(0)\}.$$

Let us define a sequence of stopping times

$$\tau^{(2i)} = \inf\{t \in [\tau^{(2i-1)}, T] : x(t) \leq 2x(0)\}, \quad i = 1, 2, \ldots,$$

$$\tau^{(2i+1)} = \inf\{t \in [\tau^{(2i)}, T] : x(t) > 2x(0)\}, \quad i = 1, 2, \ldots,$$

and an associated sequence of SDEs

$$dx^{(i)}(t) = g(x^{(i)}(t))dt - \frac{1}{2}\sigma dw(t), \quad t \geq \tau^{(2i)}, \qquad x^{(i)}(\tau^{(2i)}) = x(\tau^{(2i)}).$$

Using the comparison result for SDEs again, we have that almost surely

$$x(t) \geq x^{(i)}(t), \qquad t \in [\tau^{(2i)}, \tau^{(2i+1)}).$$

Therefore, using the CIR process as a lower bound, Assumption 2.6 holds for $q < \frac{4}{3}\left(1 + \frac{\alpha_2}{\sigma^2}\right)$ and consequently Theorem 2.7 can applied for $2 \leq q < \frac{4}{3}\left(1 + \frac{\alpha_2}{\sigma^2}\right)$. Proceeding as for the 3/2-model we have:

Proposition 3.5 *Let $1 \leq p < \frac{1}{3} + \frac{1}{3}\frac{\alpha_2}{\sigma^2}$. The LBE approximation of the Aït-Sahalia process with $r = 2$ and $\rho = 1.5$ is p-strongly convergent with order one.*

In the case $r + 1 > 2\rho$ we know from [30] that

$$\sup_{t \in [0,T]} \mathbb{E}|x(t)|^{-p} < \infty \qquad \text{for all} \qquad p \geq 1.$$

Moreover, the drift coefficient of the transformed SDE behaves at zero like the one of a transformed CEV process. A by now standard analysis gives:

Proposition 3.6 *Let $p \geq 1$. The LBE approximation of the Aït-Sahalia process with $r + 1 > 2\rho$ is p-strongly convergent with order one.*

4 A Milstein-type scheme for CIR

In this section we establish a connection between the Lamperti-backward Euler and a drift-implicit Milstein scheme for the CIR process. We will show that the order of convergence of the LBE carries over to a drift-implicit Milstein scheme, which has been proposed in [21] and [14]. While strong convergence was shown in [25], sharp convergence rates have not been established so far.

Recall that BEM for the transformed CIR process reads as

$$X_{k+1} = X_k + f(X_{k+1})\Delta t + \frac{\sigma}{2}\Delta w_{k+1}, \qquad k = 0, 1, \ldots$$

with

$$f(x) = \frac{\kappa}{2}\left(\left(\theta - \frac{\sigma^2}{4\kappa}\right)x^{-1} - x\right), \qquad x > 0.$$

Squaring yields the LBE, i.e.

$$\begin{aligned} X_{k+1}^2 &= X_k^2 + \kappa(\theta - X_{k+1}^2)\Delta t + \sigma X_k \Delta w_{k+1} \\ &\quad + \frac{\sigma^2}{4}\left((\Delta w_{k+1})^2 - \Delta t\right) - (f(X_{k+1}))^2\,\Delta t^2. \end{aligned}$$

On the other hand the drift-implicit Milstein scheme for CIR is given by

$$Z_{k+1} = Z_k + \kappa(\theta - Z_{k+1})\Delta t + \sigma\sqrt{Z_k}\Delta w_{k+1} + \frac{\sigma^2}{4}\left((\Delta w_{k+1})^2 - \Delta t\right), \quad (40)$$

hence both schemes coincide up to a term of order Δt^2. The numerical flows of the LBE and Milstein scheme are given by

$$\begin{aligned} \phi_E(x,k) &= \frac{1}{1+\kappa\Delta t}\left(x + \kappa\theta\Delta t + \sigma\sqrt{x}\Delta w_{k+1} + \frac{\sigma^2}{4}\left((\Delta w_{k+1})^2 - \Delta t\right)\right) \\ &\quad - \frac{1}{1+\kappa\Delta t}\left(f(\phi_E(x,k)\right)^2\Delta t^2 \end{aligned}$$

and

$$\phi_M(x,k) = \frac{1}{1+\kappa\Delta t}\left(x + \kappa\theta\Delta t + \sigma\sqrt{x}\Delta w_{k+1} + \frac{\sigma^2}{4}\left((\Delta w_{k+1})^2 - \Delta t\right)\right).$$

It is clear then that

$$\phi_M(x,k) \geq \phi_E(x,k)$$

for all $x > 0,\ k = 0, 1, \ldots$. From [1] we know on the other hand

$$\phi_E(x, k) \geq \phi_E(y, k) \quad \text{for} \quad x \geq y.$$

Hence we conclude

$$Z_k \geq X_k^2, \quad k = 0, 1, \ldots, \tag{41}$$

so the drift-implicit Milstein scheme dominates the Lamperti–Euler method and thus preserves positivity.

To establish the order of L^1-convergence for the drift-implicit Milstein scheme it is enough to control the difference between the Lamperti–Euler method and (40).

Lemma 4.1 *Let* $\frac{\kappa\theta}{\sigma^2} > 3/2$. *Then there exists a constant* $C > 0$ *such that*

$$\sup_{k=0,\ldots,\lceil T/\Delta t\rceil} \mathbb{E}\left|Z_k - X_k^2\right| \leq C\Delta t. \tag{42}$$

Proof Let $e_k = Z_k - X_k^2$ and note that $e_k \geq 0$ by (41). We have

$$e_{k+1} = e_k - \kappa e_{k+1}\Delta t + \sigma(\sqrt{Z_k} - X_k)\Delta w_{k+1} + (f(X_{k+1}))^2\Delta t^2.$$

Exploiting the independence of X_k and Δw_{k+1} resp. of Z_k and Δw_{k+1} it follows

$$\mathbb{E}e_{k+1} = \frac{1}{1+\kappa\Delta t}\left(\mathbb{E}e_k + \mathbb{E}(f(X_{k+1}))^2\Delta t^2\right)$$

and consequently

$$\sup_{k=0,\ldots,\lceil T/\Delta t\rceil} \mathbb{E}e_k \leq \Delta t \sum_{k=0}^{\lceil T/\Delta t\rceil-1} \mathbb{E}(f(X_{k+1}))^2\Delta t.$$

Lemma 2.10 gives that

$$\sup_{k=1,\ldots,\lceil T/\Delta t\rceil} \mathbb{E}(f(X_{k+1}))^2 \leq c \cdot \left(1 + \sup_{t\in[0,T]} \mathbb{E}|f(x(t))|^2\right)$$

which together with

$$|f(x)| \leq c \cdot \left(1 + |x| + |x|^{-1}\right)$$

and (32) shows the assertion. □

Using this result we have:

Proposition 4.2 (i) *Let* $\frac{\kappa\theta}{\sigma^2} > 3/2$. *Then, there exists a constant* $C > 0$ *such that*

$$\sup_{k=0,\dots,\lceil T/\Delta t\rceil} \mathbb{E}|y(k\Delta t) - Z_k| \leq C \cdot \Delta t. \tag{43}$$

(ii) *Let* $\frac{\kappa\theta}{\sigma^2} > 3/2$. *Then, there exists a constant* $C > 0$ *such that*

$$\mathbb{E} \sup_{k=0,\dots,\lceil T/\Delta t\rceil} |y(k\Delta t) - Z_k|^2 \leq C \cdot \Delta t. \tag{44}$$

Proof (i) This follows from the triangle inequality, Lemma 4.1 and Proposition 3.1.
(ii) Using (43) the second assertion can be shown along the lines of the proof of Proposition 5.3 in [25], where strong convergence of the drift-implicit Milstein scheme (without a convergence rate) was shown. Proceeding as in the proof of Proposition 5.3 in [25] we have

$$\begin{aligned}\mathbb{E} \sup_{k=0,\dots,\lceil T/\Delta t\rceil} |y(k\Delta t) - Z_k|^2 &\leq c \sum_{\ell=0}^{\lceil T/\Delta t\rceil-1} (1+\kappa\Delta t)^{2\ell}\mathbb{E}\,|y(\ell\Delta t) - Z_\ell|\,\Delta t \\ &\quad + c\,\mathbb{E} \sup_{k=1,\dots,\lceil T/\Delta t\rceil} \left|\sum_{\ell=0}^{k-1} \frac{1}{(1+\kappa\Delta t)^{k-\ell}}\rho_{\ell+1}\right|^2\end{aligned}$$

with

$$\begin{aligned}\rho_{k+1} = &-\kappa \int_{k\Delta t}^{(k+1)\Delta t} (y(s) - y((k+1)\Delta t))\,ds \\ &+\sigma \int_{k\Delta t}^{(k+1)\Delta t} (\sqrt{y(s)} - \sqrt{y(k\Delta t)})\,dw(s).\end{aligned}$$

So, (43) gives

$$\begin{aligned}\mathbb{E} \sup_{k=0,\dots,\lceil T/\Delta t\rceil} |y(k\Delta t) - Z_k|^2 &\leq c\Delta t \\ &\quad + c\,\mathbb{E} \sup_{k=1,\dots,\lceil T/\Delta t\rceil} \left|\sum_{\ell=0}^{k-1} \frac{1}{(1+\kappa\Delta t)^{k-\ell}}\rho_{\ell+1}\right|^2.\end{aligned}$$

For the second term straightforward computations yield that

$$\mathbb{E} \sup_{k=1,\dots,\lceil T/\Delta t\rceil} \left|\sum_{\ell=0}^{k-1} \frac{1}{(1+\kappa\Delta t)^{k-\ell}}\rho_{\ell+1}\right|^2 \leq c\Delta t,$$

which completes the proof of the proposition. □

Using the (suboptimal) second estimate of the above Proposition and Lemma 3.5 of [7] we also obtain a sharp error estimate for the piecewise linear interpolation of the drift-implicit Milstein scheme, i.e.

$$\overline{Z}_t = \frac{t_{k+1} - t}{\Delta} Z_k + \frac{t - t_k}{\Delta} Z_{k+1}, \qquad t \in [t_k, t_{k+1}],$$

in a combined $L^2 - \| \cdot \|_\infty$-norm.

Proposition 4.3 *Let* $\frac{\kappa\theta}{\sigma^2} > 3/2$. *Then, there exists a constant* $C > 0$ *such that*

$$\left(\mathbb{E} \max_{t \in [0,T]} |y(t) - \overline{Z}_t|^2 \right)^{1/2} \leq C \cdot \sqrt{|\log(\Delta t)|} \cdot \sqrt{\Delta t},$$

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

The above relation between

(a) the drift-implicit Milstein scheme applied to the original SDE

$$dy(t) = a(y(t))dt + b(y(t))dw(t)$$

and

(b) the BEM applied to the transformed SDE

$$dx(t) = f(x(t))dt + \lambda dw(t)$$

with

$$f(x) = \lambda \left(\frac{a(F^{-1}(x))}{b(F^{-1}(x))} - \frac{1}{2} b'(F^{-1}(x)) \right)$$

and

$$F(x) = \lambda \int^x \frac{1}{b(y)} dy$$

is in fact a particular case of a more general relation. Expanding LBE yields

$$\begin{aligned} F^{-1}(X_{k+1}) = F^{-1}(X_k) &+ \frac{1}{\lambda} b(F^{-1}(X_k))(X_{k+1} - X_k) \\ &+ \frac{1}{2\lambda^2} (b'b)(F^{-1}(X_k))(X_{k+1} - X_k)^2 + \ldots, \end{aligned}$$

since

$$\frac{d}{dx} F^{-1}(x) = \frac{1}{\lambda} b(F^{-1}(x)), \qquad \frac{d^2}{(dx)^2} F^{-1}(x) = \frac{1}{\lambda^2} (b'b)(F^{-1}(x)).$$

Setting $Y_k = F^{-1}(X_k)$ and using (5) we have

$$\begin{aligned} Y_{k+1} = Y_k + &\left(a(Y_{k+1}) - \frac{1}{2}(b'b)(Y_{k+1}) \right) \Delta t + b(Y_k)\Delta w_{k+1} \\ &+ \frac{1}{2}(b'b)(Y_k)|\Delta w_{k+1}|^2 \\ &+ R_k \end{aligned}$$

with

$$\begin{aligned} R_k = {}& (b(Y_k) - b(Y_{k+1}))f(X_{k+1})\Delta t + \frac{1}{\lambda}(b'b)(Y_k)\Delta w_{k+1} f(X_{k+1})\Delta t \\ &+ \frac{1}{2\lambda^2}(b'b)(Y_k) f^2(X_{k+1})\Delta t^2 + \ldots \end{aligned}$$

So dropping R_k and the other higher order terms, we end up with the numerical scheme

$$Z_{k+1} = Z_k + \left(a(Z_{k+1}) - \tfrac{1}{2} b'b(Z_{k+1}) \right) \Delta t + b(Z_k)\Delta w_{k+1} + \tfrac{1}{2} b'b(Z_k)|\Delta w_{k+1}|^2.$$

In the case of $(\alpha, \beta) = (0, \infty)$ conditions for the well-definedness, stability and strong convergence of this scheme are given in [14]. However, the convergence rate analysis for the CIR process, where we can exploit (among other things) that $F^{-1}(x) = x^2$ and also the domination property (41), seems not to carry over to the general case.

5 Conclusion and discussion

In this paper we presented a Lamperti–Euler scheme for scalar SDEs which take values in a domain $D = (l, r)$ and have non-Lipschitz drift or diffusion coefficients. Our strategy is to first use the Lamperti transformation $x(t) = F(y(t))$ (provided that the diffusion coefficient of the original SDE is strictly positive on D) and then to approximate the transformed process $x(t)$, $0 \le t \le T$, with the backward Euler scheme. Transforming back with the inverse Lamperti transformation gives an approximation scheme for the original SDE. We also pointed out a relation of this scheme to a drift-implicit Milstein scheme.

The advantages of this Lamperti–Euler method are

- the preservation of the domain of the original SDE and
- the availability of a framework to establish strong convergence rates.

In particular, we use this framework to obtain such strong convergence results for several SDEs with non-Lipschitz coefficients from both mathematical finance and bio-mathematics.

Open questions are whether Assumptions 2.2 and 2.6 can be formulated in terms of conditions on the original coefficients of the SDE, and moreover which properties of the original coefficients allow to control the error propagation when applying the inverse Lamperti transformation. Here, manageable characterizations would allow to avoid case by case studies. Clearly, if the diffusion coefficient is bounded away from zero and from above, the Lamperti transform and its inverse can be easily controlled. But this assumption would rule out all the examples from Sect. 3, where the violation of this assumption is in fact at the core of the mathematical and numerical difficulties. Further open questions are whether the implicitness of this scheme (which e.g. for the CEV process requires solving a non-linear equation) can be avoided by using a tamed Euler scheme (as in [18] for the case $D=\mathbb{R}$) and also whether the convergence rate for the Lamperti–Euler also carries over in general (and not only for the CIR process) to a drift-implicit Milstein scheme. In particular, the last point leads to the following general question: Given a certain numerical approximation for an SDE with non-Lipschitz coefficients, which perturbations do not change its convergence properties and also its qualitative properties? We will pursue all these topics in our future research.

Acknowledgments We would like to thank the referees for their valuable and insightful comments and remarks. Moreover, we would like to thank Martin Altmayer for helpful comments on an earlier version of the manuscript.

References

1. Alfonsi, A.: On the discretization schemes for the CIR (and Bessel squared) processes. Monte Carlo Methods Appl. **11**(4), 355–384 (2005)
2. Alfonsi, A.: Strong order one convergence of a drift implicit Euler scheme: application to the CIR process. Stat. Probab. Lett. **83**(2), 602–607 (2013)
3. Andersen, L.B.G.: Simple and efficient simulation of the Heston stochastic volatility model. J. Comput. Finance **11**(3), 1–42 (2008)
4. Berkaoui, A., Bossy, M., Diop, A.: Euler scheme for SDEs with non-Lipschitz diffusion coefficient: strong convergence. ESAIM Probab. Stat. **12**, 1–11 (2007)
5. Cox, J.: Notes on option pricing I: constant elasticity of variance diffusions. Unpublished manuscript. Stanford University, Stanford (1975)
6. Dangerfield, C.E., Kay, D., MacNamara, S., Burrage, K.: A boundary preserving numerical algorithm for the Wright–Fisher model with mutation. BIT Numer. Math. **52**(2), 283–304 (2012)
7. Dereich, S., Neuenkirch, A., Szpruch, L.: An Euler-type method for the strong approximation of the Cox–Ingersoll–Ross process. Proc. R. Soc. A Math. Phys. Eng. Sci. **468**(2140), 1105–1115 (2012)
8. Either, S., Kurtz, T.: Markov processes: characterization and convergences. John Wiley Sons, New York (1986)
9. Giles, M.: Monte Carlo and Quasi-Monte Carlo methods 2006. Conference Proceedings. In: Keller, A. (ed.) Improved multilevel Monte Carlo convergence using the Milstein scheme, pp. 343–358. Springer, New York (2008)
10. Giles, M.: Multilevel Monte Carlo path simulation. Oper. Res. **56**(3), 607–617 (2008)
11. Heston, S.L.: A simple new formula for options with stochastic volatility. Course notes. Washington University, St. Louis (1997)
12. Higham, D.J., Mao, X.: Convergence of Monte Carlo simulations involving the mean-reverting square root process. J. Comput. Finance **8**(3), 35–62 (2005)
13. 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 (2003)
14. Higham, D.J., Mao, X., Szpruch, L.: Convergence, non-negativity and stability of a new Milstein scheme with applications to finance. Discret. Contin. Dyn. Syst. Ser. B **18**(8), 2083–2100 (2013)
15. Howison, S., Schwarz, D.: Risk-neutral pricing of financial instruments in emission markets: a structural approach. SIAM J. Math. Finance **3**, 709–739 (2012)
16. Hurd, T.R., Kuznetsov, A.: Explicit formulas for Laplace transforms of stochastic integrals. Markov Process. Relat. Fields **14**(2), 277–290 (2008)

17. Jentzen, A., Hutzenthaler, M., Kloeden, P.E.: Divergence of the multilevel Monte Carlo Euler method for nonlinear stochastic differential equations. Ann. Appl. Prob. **23**(5), 1721–1760 (2013)
18. Jentzen, A., Hutzenthaler, M., Kloeden, P.E.: Strong convergence of an explicit numerical method for SDEs with non-globally Lipschitz continuous coefficients. Ann. Appl. Probab. **22**(4), 1611–1641 (2012)
19. Iacus, S.M.: Simulation and Inference for stochastic differential equations: with R examples. Springer, New York (2008)
20. Jentzen, A., Kloeden, P.E., Neuenkirch, A.: Pathwise approximation of stochastic differential equations on domains: higher order convergence rates without global Lipschitz coefficients. Numerische Mathematik **112**(1), 41–64 (2009)
21. Kahl, C., Günther, M., Rosberg, T.: Structure preserving stochastic integration schemes in interest rate derivative modeling. Appl. Numer. Math. **58**(3), 284–295 (2008)
22. Kahl, C., Jäckel, P.: Fast strong approximation Monte Carlo schemes for stochastic volatility models. Quant. Finance **6**(6), 513–536 (2006)
23. Karatzas, I., Shreve, S.E.: Brownian motion and stochastic calculus. Springer, New York (1991)
24. Karlin, S., Taylor, H.M.: A second course in stochastic processes. Academic Press, Dublin (1981)
25. Kloeden, P.E., Neuenkirch, A.: Recent developments in computational finance. In: Gerstner, T., Kloeden, P.E. (eds.) Convergence of numerical methods for stochastic differential equations in mathematical finance, pp. 49–81. World Scientific, Singapore (2013)
26. Kloeden, P.E., Platen, E.: Numerical solution of stochastic differential equations. Springer, New York (1992)
27. Larsen, K.S., Sørensen, M.: Diffusion models for exchange rates in a target zone. Math. Finance **17**(2), 285–306 (2007)
28. Lord, R., Koekkoek, R., van Dijk, D.J.C.: A comparison of biased simulation schemes for stochastic volatility models. Quant. Finance **10**(2), 177–194 (2009)
29. Mao, X., Szpruch, L.: Strong convergence rates for backward Euler–Maruyama method for non-linear dissipative-type stochastic differential equations with super-linear diffusion coefficients. Stochastics **85**(1), 144–171 (2013)
30. Szpruch, L., Mao, X., Higham, D.J., Pan, J.: Strongly nonlinear Aït-Sahalia-type interest rate model and its numerical approximation. BIT Numer. Math. **51**(2), 405–425 (2010)
31. Zeidler, E.: Nonlinear functional analysis and its applications. Springer, New York (1985)
32. Zhu, J.: Modular pricing of options: an application of Fourier analysis. Springer, New York (2009)