Numer. Math. (2009) 112:41–64
DOI 10.1007/s00211-008-0200-8

Numerische Mathematik

Pathwise approximation of stochastic differential equations on domains: higher order convergence rates without global Lipschitz coefficients

A. Jentzen · P. E. Kloeden · A. Neuenkirch

Received: 9 July 2007 / Revised: 10 September 2008 / Published online: 27 November 2008

Abstract We study the approximation of stochastic differential equations on domains. For this, we introduce modified Itô–Taylor schemes, which preserve approximately the boundary domain of the equation under consideration. Assuming the existence of a unique non-exploding solution, we show that the modified Itô–Taylor scheme of order γ has pathwise convergence order $\gamma - \varepsilon$ for arbitrary $\varepsilon > 0$ as long as the coefficients of the equation are sufficiently differentiable. In particular, no global Lipschitz conditions for the coefficients and their derivatives are required. This applies for example to the so called square root diffusions.

Mathematics Subject Classification (2000) 65C30 · 60H35

1 Introduction

In this article, we study the approximation of the Itô stochastic differential equation

$$dx(t) = f(x(t))\,dt + g(x(t))\,dW(t), \qquad t \geq 0, \quad x(0) = x_0, \tag{1}$$

Partially supported by the DFG-project "Pathwise numerics and dynamics of stochastic evolution equations".

A. Jentzen · P. E. Kloeden · A. Neuenkirch (✉)
Johann Wolfgang Goethe-Universität, Institut für Mathematik,
Robert-Mayer-Straße 10, 60325 Frankfurt am Main, Germany
e-mail: neuenkirch@math.uni-frankfurt.de

A. Jentzen
e-mail: jentzen@math.uni-frankfurt.de

P. E. Kloeden
e-mail: kloeden@math.uni-frankfurt.de

which takes values in a domain $D \subset \mathbb{R}^d$. Here, $f : D \to \mathbb{R}^d$, $g : D \to \mathbb{R}^{d,m}$, $W = (W(t), t \geq 0)$ is a m-dimensional Brownian motion adapted to a filtration $\mathscr{F} = (\mathscr{F}_t, t \geq 0)$ on a probability space $(\Omega, \mathscr{A}, \mathbf{P})$, and x_0 is a $\mathscr{F}_0$-measurable random variable, which is independent of W.

In many applications, the drift- or diffusion coefficients of Eq. (1) have a simple structure, e.g. polynomial, but fail to satisfy a global Lipschitz condition.

For example, the stochastic Volterra–Lotka system

$$dx(t) = \operatorname{diag}(x_1(t), \ldots, x_d(t)) \left[(A + Bx(t))\, dt + Cx(t)\, dW(t) \right], \tag{2}$$

where $A \in \mathbb{R}^d$, $B, C \in \mathbb{R}^{d,d}$ and $(W(t), t \geq 0)$ is a one-dimensional Brownian motion, see, e.g. [17], has polynomial coefficients and scalar noise. Another example with polynomial coefficients is the stochastic Duffing–van der Pol equation (see [3])

$$\begin{aligned} dx_1(t) &= x_2(t)\, dt, \\ dx_2(t) &= (\alpha x_1(t) + \beta x_2(t) - x_1^3(t) - x_1^2(t)x_2(t) - \frac{\sigma_1^2}{2} x_1(t) - \frac{\sigma_2^2}{2} x_2(t))\, dt \\ &\quad + \sigma_1 x_1(t)\, dW^{(1)}(t) + \sigma_2 x_2(t)\, dW^{(2)}(t) + \sigma_3\, dW^{(3)}(t), \end{aligned} \tag{3}$$

where $\alpha, \beta, \sigma_1, \sigma_2, \sigma_3 \in \mathbb{R}$ and $W^{(1)}$, $W^{(2)}$ and $W^{(3)}$ are three independent scalar Brownian motions.

Moreover, in mathematical biology and financial mathematics, see, e.g. [14,21], particular interest has been given to stochastic differential equations of the so called Cox–Ingersoll–Ross type

$$dx(t) = \kappa\, (\lambda - x(t))\, dt + \theta \sqrt{|x(t)|}\, dW(t), \tag{4}$$

where $\kappa, \lambda, \theta \geq 0$. Here the diffusion coefficient neither satisfies a global Lipschitz condition nor is differentiable at $x = 0$. However, if $\kappa\lambda \geq \theta^2/2$ and $x_0 \in (0, \infty)$, then the boundary zero is unattainable. Hence in this case the solution of Eq. (4) never leaves the set $D = (0, \infty)$ and the coefficients are infinitely differentiable on D. For further examples, we refer to [2,3,14,17,21].

In this article, we will consider the approximation of Eq. (1) with respect to the pathwise maximum error in the discretization points

$$\sup_{i=1,\ldots,n} |x(t_i, \omega) - \overline{x}(t_i, \omega)|, \quad \omega \in \Omega, \tag{5}$$

where $t_1, \ldots, t_n$ are the nodes of the discretization. This error criterion is in particular appropriate for equations with a non-integrable initial value and for stochastic dynamics, since the theory of random dynamical systems is of pathwise nature. Approximation methods for stochastic differential equations with respect to pathwise error criteria have been considered, e.g. in [6,8,9,16,24]. While in [16,24] pathwise convergence rates for several approximation schemes (including the standard Itô–Taylor schemes)

under classical assumptions on the coefficients are determined, first results on the pathwise approximation under non-global Lipschitz assumptions are given in [6,9].

Assuming the existence of an appropriate Lyapunov function for Eq. (1) and that f and g are locally Lipschitz continuous, it is shown in [9] that the standard Euler method has pathwise order of convergence $1/2 - \varepsilon$ for arbitrary small $\varepsilon > 0$. This pathwise convergence rate of the explicit Euler method is also obtained in [6], in which the assumption on the Lyapunov function is replaced by the existence of a unique non-exploding solution. In [6] also the explicit Milstein method is considered and is shown to be pathwise convergent, if f, g and their first derivatives are locally Lipschitz continuous.

Here, we will consider the approximation of Eq. (1) by modified Itô–Taylor methods, assuming only that Eq. (1) has a unique solution $x = (x(t),\ t \geq 0)$, whose sample paths are contained in a domain $D \subset \mathbb{R}^d$, i.e.

$$\mathbf{P}(x(t) \in D \ \text{ for all } \ t \geq 0) = 1.$$

For this, we will modify the standard Itô–Taylor schemes by introducing auxiliary drift and diffusion functions for the case that the numerical scheme leaves the domain D. Unless $D = \mathbb{R}^d$, the modification of the standard Itô–Taylor schemes is required, since the coefficients of the equation may not be differentiable or even well defined outside D, and therefore the standard Itô–Taylor schemes are not applicable to these equations. This modification turns out to be a flexible method, which we will illustrate by several examples. In particular, since D is a domain and thus open, the boundary of D is often a reflecting barrier for Eq. (1). So the auxiliary drift and diffusion functions should reproduce this reflection property of the boundary.

Our main result is that the modified Itô–Taylor method of order γ has pathwise order of convergence $\gamma - \varepsilon$, if the coefficients of the equation are $2\gamma + 1$-times continuously differentiable on D and the auxiliary drift and diffusion coefficients are $2\gamma - 1$-times continuously differentiable outside of $\overline{D}$.

Thus, we recover the pathwise convergence rate $\gamma - \varepsilon$ of the standard Itô–Taylor method of order γ, which has been derived under classical assumptions in [16]. Moreover the case $D = \mathbb{R}^d$, in which the modified Itô–Taylor schemes and the standard Itô–Taylor schemes coincide, illustrates that the pathwise convergence rates of the standard Itô–Taylor schemes are very "robust" in the sense that they can be retained under almost minimal assumptions.

The remainder of this article is structured as follows: In the next section, we will introduce our modified Itô–Taylor schemes and we will state our main result. Moreover, several examples are given for illustration and we will comment briefly on results on the weak and strong approximation of stochastic differential equations under non-standard assumptions. Finally, in Sect. 3 we give some numerical examples, while the proofs are postponed to Sect. 4.

We will use the following notation: For $D \subset \mathbb{R}^d$ we will denote by $C^r(D; \mathbb{R}^p)$ the set of all functions from D to $\mathbb{R}^p$, which are r-times continuously differentiable. Moreover, we will use the notation $C_b^r(D; \mathbb{R}^p)$ for the set of all functions, which belong to $C^r(D; \mathbb{R}^p)$ and are bounded together with their first r-derivatives.

2 Modified Itô–Taylor schemes

Throughout this article, we will impose the following general assumption on the stochastic differential equation, which we want to approximate:

(A) Equation (1) has a pathwise unique strong solution $x = (x(t),\ t \geq 0)$ with

$$\mathbf{P}(x(t) \in D \ \text{ for all } \ t \geq 0) = 1,$$

i.e. almost all solution sample paths never leave the set D for all $t \geq 0$.

Recall that a process $x = (x(t),\ t \geq 0)$ is called a solution of Eq. (1), see, e.g. [15], if

(i) almost all sample paths of x are continuous and x is $\mathscr{F}$-adapted,

(ii) $\mathbf{P}(x(0) = x_0) = 1$,

(iii)

$$\int_0^t |f(x(s))|\,ds + \sum_{j=1}^{m} \int_0^t |g^j(x(s))|^2\,ds < \infty$$

with probability one for all $t \geq 0$ and

(iv)

$$x(t) = x_0 + \int_0^t f(x(s))\,ds + \sum_{j=1}^{m} \int_0^t g^j(x(s))\,dW^j(s), \quad t \geq 0$$

holds with probability one, where g^j denotes the jth column of the diffusion matrix g.

Thus, we allow in particular the case that the solution x of Eq. (1) is not integrable, i.e. $\mathbf{E}|x(t)| = \infty$ for some $t \geq 0$.

For convenience, we first recall the standard Itô–Taylor schemes.

2.1 Itô–Taylor schemes

These approximation schemes are defined as follows, see, e.g. [17]: Let

$$\mathscr{M} = \Big\{\alpha = (j_1, \ldots, j_l) \in \{0, 1, 2, \ldots, m\}^l : l \in \mathbb{N}\Big\} \cup \{v\}$$

be the set of all multi-indices. The length of a multi-index $\alpha = (j_1, \ldots, j_l)$ is defined as $l(\alpha) = l$ and v is the multi-index of length 0. Moreover, let $n(\alpha)$ be the number of entries of α, which are equal to 0. For $\alpha = (j_1, \ldots, j_l)$ and $0 \leq s \leq t < \infty$ we define

$$I_\alpha(s,t) = \int_s^t \cdots \int_s^{\tau_2} dW^{j_1}(\tau_1) \cdots dW^{j_l}(\tau_l)$$

with the convention that $dW^0(\tau) = d\tau$. We also introduce the operators

$$L^0 = \sum_{k=1}^{d} f^k \frac{\partial}{\partial x^k} + \frac{1}{2} \sum_{k,l=1}^{d} \sum_{j=1}^{m} g^{k,j} g^{l,j} \frac{\partial^2}{\partial x^k \partial x^l}$$

and

$$L^j = \sum_{k=1}^{d} g^{k,j} \frac{\partial}{\partial x^k}$$

for $j \in \{1, \dots, m\}$. Here f^k, $g^{k,j}$ are the kth components of f and g^j, respectively.

Finally, we define for $\gamma = 0.5, 1.0, 1.5, \dots$ the sets of multi-indices

$$\mathscr{A}_\gamma = \left\{ \alpha \in \mathscr{M} : l(\alpha) + n(\alpha) \leq 2\gamma \text{ or } l(\alpha) = n(\alpha) = \gamma + \frac{1}{2} \right\}$$

and let be

$$0 = t_0 < t_1 < \cdots < t_n = T$$

a discretization of $[0, T]$ with $T \in (0, \infty)$.

Then the Itô–Taylor scheme of order γ is given by

$$\begin{aligned} \overline{x}_n^\gamma(t_0) &= x_0, \\ \overline{x}_n^\gamma(t_{i+1}) &= \overline{x}_n^\gamma(t_i) + \sum_{\alpha \in \mathscr{A}_\gamma \setminus \{v\}} h_\alpha(\overline{x}_n^\gamma(t_i)) \cdot I_\alpha(t_i, t_{i+1}) \end{aligned}$$

for $i = 0, \dots, n-1$, where

$$h_\alpha(x) = L^{j_1} \cdots L^{j_{l-1}} g^{j_l}(x), \quad x \in \mathbb{R}^d$$

for $\alpha = (j_1, \dots, j_l)$ and $g^0 = f$.

However, these schemes may not be well defined, if the coefficients of Eq. (1) are only given on D: the approximations computed by these schemes may leave the set D, while the exact solution does not. To avoid this problem, we will introduce in the next section an appropriate modification of the coefficients of Eq. (1).

2.2 Modified Itô–Taylor schemes

Now assume that the coefficients of Eq. (1) are r-times continuously differentiable on D, i.e. $f \in C^r(D; \mathbb{R}^d)$ and $g \in C^r(D; \mathbb{R}^{d,m})$ with $r \in \mathbb{N}$. Moreover set $E = \{x \in \mathbb{R}^d : x \notin \overline{D}\}$.

To define our modified Itô–Taylor schemes, we now choose two auxiliary functions $a \in C^s(E; \mathbb{R}^d)$ and $b \in C^s(E; \mathbb{R}^{d,m})$ for $s \in \mathbb{N}$ and define

$$\widetilde{f}(x) = f(x) \cdot 1_D(x) + a(x) \cdot 1_E(x), \quad x \in D \cup E,$$
$$\widetilde{g}(x) = g(x) \cdot 1_D(x) + b(x) \cdot 1_E(x), \quad x \in D \cup E.$$

For $x \in \partial D$ set

$$\widetilde{f}(x) = \lim_{y \to x;\, y \in D} \widetilde{f}(y), \quad \widetilde{g}(x) = \lim_{y \to x;\, \in y \in D} \widetilde{g}(y),$$

if these limits exist. Otherwise define $\widetilde{f}(x) = 0$, respectively $\widetilde{g}(x) = 0$ for $x \in \partial D$. (Note that the Itô–Taylor schemes in general hit the boundary of D with probability zero, so the values of the coefficients on the boundary do not matter.)

Moreover, we also need the following "modified" derivative for a function $h : \mathbb{R}^d \to \mathbb{R}^d$. For $l = 1, \ldots, d$ set

$$\partial_{x^l} h(x) = \frac{\partial}{\partial x^l} h(x), \quad x \in D \cup E$$

and for $x \in \partial D$ define

$$\partial_{x^l} h(x) = \lim_{y \to x;\, y \in D} \partial_{x^l} h(x),$$

if this limit exists. Otherwise set $\partial_{x^l} h(x) = 0$ for $x \in \partial D$.

These modifications of the coefficients and the modified derivatives may seem rather technical, but all one has to do is the following:

(1) Keep the original coefficients on D.
(2) Define new coefficients on E, which is the complement of the closure of D.
(3) Define the coefficients and their derivatives appropriately on ∂D.

Using the above notations and conventions the operators

$$\widetilde{L}^0 = \sum_{k=1}^{d} \widetilde{f}^k \, \partial_{x^k} + \frac{1}{2} \sum_{k,l=1}^{d} \sum_{j=1}^{m} \widetilde{g}^{k,j} \widetilde{g}^{l,j} \, \partial_{x^k} \partial_{x^l}$$

and

$$\widetilde{L}^j = \sum_{k=1}^{d} \widetilde{g}^{k,j} \, \partial_{x^k}$$

for $j \in \{1, \ldots, m\}$ are well defined. Our modified Itô–Taylor method of order γ based on the auxiliary functions a and b is now given by

$$\widetilde{x}_n^\gamma(t_0) = x_0,$$
$$\widetilde{x}_n^\gamma(t_{i+1}) = \widetilde{x}_n^\gamma(t_i) + \sum_{\alpha \in \mathscr{A}_\gamma \setminus \{v\}} \widetilde{h}_\alpha(\widetilde{x}_n^\gamma(t_i)) \cdot I_\alpha(t_i, t_{i+1})$$

for $i = 0, \ldots, n-1$, where

$$\widetilde{h}_\alpha(x) = \widetilde{L}^{j_1} \cdots \widetilde{L}^{j_{l-1}} \widetilde{g}^{j_l}(x), \quad x \in \mathbb{R}^d$$

for $\alpha = (j_1, \ldots, j_l)$ and $\widetilde{g}^0 = \widetilde{f}$.

The purpose of the auxiliary drift- and diffusion coefficients is twofold:

(i) to obtain a well-defined numerical method,
(ii) to "reflect" the numerical scheme back to D, if it has left D.

In particular, one can always choose the auxiliary functions to be affine or even constant, in which case these schemes have a simple structure outside of D.

The modified Itô–Taylor method of order γ is well defined as long as the coefficients of the equation are $2\gamma - 1$-times differentiable on D and the auxiliary functions are $2\gamma - 1$-times differentiable on E. Hence the assumptions in the following theorem are almost minimal. For simplicity, we will choose an equidistant discretization of $[0, T]$, i.e. $t_i = iT/n$ for $i = 0, 1, \ldots, n$.

Theorem 1 *Let $\gamma = 0.5, 1.0, 1.5, \ldots$. Assume that*

$$f \in C^{2\gamma+1}(D; \mathbb{R}^d), \quad g \in C^{2\gamma+1}(D; \mathbb{R}^{d,m})$$

and

$$a \in C^{2\gamma-1}(E; \mathbb{R}^d), \quad b \in C^{2\gamma-1}(E; \mathbb{R}^{d,m}).$$

Moreover let $\widetilde{x}_n^\gamma$ be the modified Itô–Taylor method based on the auxiliary functions a and b. Then for all $\varepsilon > 0$ there exists a finite and non-negative random variable $\eta_{\gamma,\varepsilon}^{a,b}$ such that

$$\sup_{i=0,\ldots,n} \left| x(iT/n, \omega) - \widetilde{x}_n^\gamma(iT/n, \omega) \right| \leq \eta_{\gamma,\varepsilon}^{a,b}(\omega) \cdot n^{-\gamma+\varepsilon}$$

for almost all $\omega \in \Omega$ and all $n \in \mathbb{N}$.

Thus, the modified Itô–Taylor methods obtain pathwise convergence order $\gamma - \varepsilon$ for arbitrarily $\varepsilon > 0$, regardless which auxiliary functions are chosen. However, the constant in the error bound clearly depends on the auxiliary functions and there are "natural choices" for the auxiliary functions, which we will illustrate in the following.

2.3 Examples

(i) We consider first the Cox–Ingersoll–Ross process

$$dx(t) = \kappa\,(\lambda - x(t))\,dt + \theta\sqrt{|x(t)|}\,dW(t), \qquad t \geq 0, \quad x(0) = x_0, \tag{6}$$

with $\kappa\lambda \geq \theta^2/2$ and $x_0 > 0$. Here we have $D = (0, \infty)$ and $f, g \in C^\infty(D; \mathbb{R})$, since

$$f(x) = \kappa\,(\lambda - x)\,, \quad g(x) = \theta\sqrt{x}, \quad x \in D.$$

As auxiliary functions we can choose, e.g.

$$a(x) = b(x) = 0, \quad x \in E$$

or

$$a(x) = \kappa\,(\lambda - x)\,, \quad b(x) = 0, \quad x \in E.$$

The first set of auxiliary functions "kills" the numerical approximation as soon as it reaches a negative value. However, the second set of auxiliary functions is more appropriate, since if the scheme would take a negative value, the auxiliary functions force the numerical scheme to be positive again after the next steps, which better recovers the positivity of the exact solution. (In the case of the Euler scheme for Eq. (6) a detailed discussion on the choice of the auxiliary functions can be found in [19].)

Another possibility to retain the positivity of the numerical approximation to Eq. (6) in the case $\kappa\lambda \geq \theta^2/2$ and $x_0 > 0$ is to use the Lamperti transformation: the process $y(t) = \sqrt{x(t)}$, $t \geq 0$, satisfies the stochastic differential equation

$$dy(t) = \left(\frac{\kappa\lambda}{2} - \frac{\theta^2}{8}\right)\frac{1}{y(t)}\,dt - \frac{\kappa}{2}y(t)\,dt + \frac{\theta}{2}dW(t), \quad t \geq 0. \tag{7}$$

Here we have again

$$\mathbf{P}(y(t) \in D \ \text{ for all } \ t \geq 0) = 1, \tag{8}$$

and we thus can use a modified Itô–Taylor scheme of order γ to approximate Eq. (7). Transforming back, we obtain the strictly positive approximation

$$(\widetilde{y}_n^\gamma(iT/n, \omega))^2, \quad i = 0, \ldots, n,$$

to the original equation (6). Moreover, as a straightforward consequence of Theorem 1 we have that for all $\varepsilon > 0$ there exists a finite and non-negative random variable $\widetilde{\eta}_{\gamma,\varepsilon}^{a,b}$ such that

$$\sup_{i=0,\ldots,n} \left| x(iT/n, \omega) - (\widetilde{y}_n^\gamma(iT/n, \omega))^2 \right| \leq \widetilde{\eta}_{\gamma,\varepsilon}^{a,b}(\omega) \cdot n^{-\gamma+\varepsilon}$$

for almost all $\omega \in \Omega$ and all $n \in \mathbb{N}$. In view of (8) one should also here choose the auxiliary functions for these "squared Itô–Taylor methods" appropriately.

Due to its applications in financial mathematics particular attention has been given recently to the weak and strong (structure preserving) approximation of the Cox–Ingersoll–Ross, see, e.g. [1,4,5,10,13,19,23]. While weak convergence rates of order 1

have been obtained for several schemes [1,4], a strong convergence rate has been derived—up to our best knowledge—only for a symmetrized Euler scheme under very restrictive assumptions on the parameters λ, κ and θ, see [4].

(ii) The coefficients of the stochastic Volterra–Lotka equation

$$dx(t) = \operatorname{diag}(x_1(t), \dots, x_n(t)) \left[(A + Bx(t))\, dt + Cx(t)\, dW(t) \right], \quad t \geq 0, \tag{9}$$

are infinitely differentiable on $\mathbb{R}^d$, hence there is no need for a modification of the coefficients and one can use the standard Itô–Taylor schemes. On the other hand it was shown by Mao et al. [20] that the solution never leaves the set $D = (0, \infty)^d$, if the initial value satisfies $x_0 \in (0, \infty)^d$ and if the entries of the diffusion matrix are positive with the diagonal elements strictly positive. Thus, it is also favorable for this equation to use the introduced modified Itô–Taylor methods with $D = (0, \infty)^d$ and, e.g.

$$a^k(x) = -2|A_k| x_k 1_{\{x_k < 0\}}(x), \qquad b^{k,j}(x) = 0, \quad x \in E$$

for $k = 1, \dots, d$, $j = 1, \dots, m$, where A_k is the kth element of the vector A.

(iii) The diffusion limit of the Wright–Fisher model, see, e.g. [14], is another popular (one-dimensional) stochastic differential equation in mathematical biology:

$$dx(t) = f(x(t))\, dt + \sqrt{|x(t)(1 - x(t))|}\, dW(t), \quad t \geq 0. \tag{10}$$

The drift coefficient is typically a polynomial and this equation takes values in the interval $[0, 1]$. Here, in general (depending on the structure of f) the solution attains the boundaries $\{0, 1\}$, i.e. we have

$$\tau_{\{0,1\}} = \inf\{t \in [0, T] : x(t) \notin (0, 1)\} < T$$

with a positive probability. Thus, we can not directly use Theorem 1 to approximate Eq. (10). However, if we apply nevertheless the modified Itô–Taylor method $\widetilde{x}_n^\gamma$ with the auxiliary functions $a = 0$ and $b = 0$ to (10), then we obtain the error bound

$$\sup_{i=0,\dots,n} \left| y(iT/n, \omega) - \widetilde{x}_n^\gamma(iT/n, \omega) \right| \leq \eta_{\gamma,\varepsilon}(\omega) \cdot n^{-\gamma+\varepsilon}$$

for almost all $\omega \in \Omega$ and all $n \in \mathbb{N}$, where

$$y(t, \omega) = x(t, \omega) 1_{\{t < \tau_{\{0,1\}}(\omega)\}}, \qquad t \geq 0,\ \omega \in \Omega.$$

Thus, we can use these modified Itô–Taylor methods for the approximation of the Wright–Fisher equation up to the first hitting time of the boundary.

2.4 Weak and strong approximation under non-standard assumptions

To our best knowledge, weak and strong approximation of stochastic differential equations under non-standard assumptions has been studied in very few articles so far.

Yan [25] gives necessary and sufficient conditions for the weak convergence of the Euler method. In the one-dimensional case, he obtains also an upper bound for the convergence rate, if the drift coefficient is Lipschitz continuous and the diffusion coefficient is Hölder continuous. (Note that Theorem 1 clearly implies the weak convergence of the modified Itô–Taylor schemes.)

Weak approximation for equations with non-global Lipschitz coefficients is also considered in [22]. Here it is used that for a suitable class of test functions trajectories, which leave a sufficiently large sphere, can be ignored up to a prescribed approximation accuracy.

For strong approximation under non-Lipschitz assumptions, mainly Euler-type methods have been analyzed [11,12,18].

In [11], it is shown that the explicit Euler method converges in the mean square sense, if the drift and diffusion coefficients are locally Lipschitz continuous and the Euler method and the exact solution satisfy a moment condition. The same holds true for an adaptive Euler method given in [18], which is designed for long-time integration of stochastic differential equations.

Mean square convergence rates for Euler methods have been established under the additional assumptions that the drift coefficient has a polynomial behavior, satisfies a one-sided Lipschitz condition

$$\langle a-b,\, f(a)-f(b)\rangle \leq \mu \cdot |a-b|^2, \quad a,b \in \mathbb{R}^d \tag{11}$$

with $\mu \geq 0$, and the diffusion coefficient is globally Lipschitz. Under these assumptions, it is shown in [11] that the drift-implicit Euler method and the split-step Euler method have mean square order of convergence $1/2$. A similar result for the drift-implicit Euler method is also given in [12].

In [11,12], the implicitness helps in particular to control the moments of the approximation schemes. Thus, it would be a natural question also to analyze the pathwise convergence rates of drift-implicit Itô–Taylor methods under weak assumptions. However, drift-implicit methods may not be well defined assuming only that the coefficients of the equation are sufficiently differentiable. Consider, e.g. the one-dimensional Volterra–Lotka equation (also called stochastic Verhulst equation)

$$dx(t) = (\lambda x(t) - x^2(t))\,dt + \sigma x(t)\,dW(t), \quad t \geq 0, \quad x(0) = x_0 \tag{12}$$

with $\lambda, \sigma, x_0 > 0$. Then the drift-implicit Euler method for this equation is given by

$$\overline{x}_n(t_{i+1}) = \overline{x}_n(t_i) + (\lambda \overline{x}_n(t_{i+1}) - \overline{x}_n^2(t_{i+1}))\Delta_i + \sigma \overline{x}_n(t_i)\Delta_i W$$

with $\Delta_i = t_{i+1} - t_i$ and $\Delta_i W = W(t_{i+1}) - W(t_i)$. The corresponding quadratic equation for $\overline{x}_n(t_{i+1})$ has the following solutions:

$$\overline{x}_n(t_{i+1}) = \frac{1-\lambda\Delta_i}{2\Delta_i}\left(-1 \pm \sqrt{1+\frac{4\Delta_i\overline{x}_n(t_i)}{(1-\lambda\Delta_i)^2}(1+\sigma\Delta_i W)}\right).$$

Here, the negative solution can be discarded, since the solution of (12) is positive for all $t \geq 0$. However, the expression under the square-root can be negative, even for very small step sizes, and thus the drift-implicit Euler method for this stochastic differential equation is not well defined for non-adaptive discretizations. (Even adaptive discretizations would require very small step sizes with positive probability.) So, at least a one-sided Lipschitz condition as (11) is required for drift-implicit Itô–Taylor methods to be well defined.

3 Numerical examples

In this section we illustrate our results with several numerical examples. The first equation we consider is the Cox–Ingersoll–Ross process given by

$$dx(t) = (1-x(t))\,dt + \sqrt{x(t)}\,dW(t), \quad t \in [0,1], \quad x(0) = \frac{1}{5}. \tag{13}$$

The corresponding stochastic differential equation for $y(t) = \sqrt{x(t)}$, $t \in [0,1]$, reads as

$$dy(t) = \frac{3}{8}\frac{1}{y(t)}\,dt - \frac{1}{2}y(t)\,dt + \frac{1}{2}\,dW(t), \quad t \in [0,1], \quad y(0) = \frac{1}{\sqrt{5}}. \tag{14}$$

Figure 1 shows for four different sample paths $\omega \in \Omega$ the maximum error in the discretization points, i.e.

$$\sup_{i=0,\ldots,n} |x(t_i,\omega) - \overline{x}_n(t_i,\omega)|,$$

which we call "pathwise maximum error" in what follows, of

(i) the Euler scheme $(-)$ for (13) with auxiliary functions $a(x) = 1-x$, $x < 0$ and $b(x) = 0$, $x < 0$,
(ii) the "squared Euler scheme" $(- \cdot -)$, compare Sect. 2.3, for the reduced equation (14) with auxiliary functions $a(y) = -\frac{1}{2}y$, $y < 0$ and $b(y) = 0$, $y < 0$,
(iii) the "squared Wagner–Platen scheme" $(- \; -)$ for the reduced equation (14) with auxiliary functions $a(y) = -\frac{1}{2}y$, $y < 0$ and $b(y) = 0$, $y < 0$.

Since Eq. (13) has no known explicit solution, we have discretized it with very small step size in order to estimate the pathwise maximum error for the above approximation schemes. For this we apply scheme (iii).

In Fig. 1 we use log–log-coordinates. Thus, the dotted lines correspond to the convergence orders 0.5, 1 and 1.5, respectively. The pathwise convergence rates of all three approximation schemes are in very good accordance with the theoretically predicted rates, also for moderate stepsizes. Moreover, the singularity in the drift

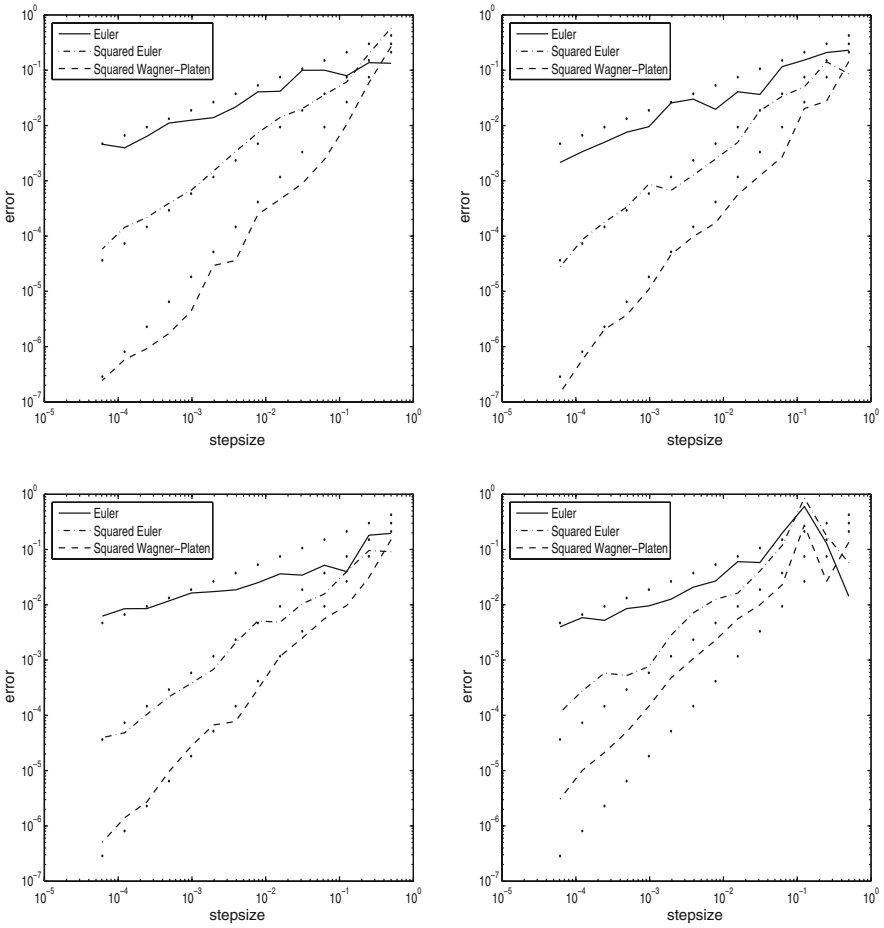

Fig. 1 Example (13): pathwise maximum error versus step size for four sample paths

coefficient of Eq. (14) turns out to be "unproblematic" in the pathwise numerical simulations.

The second example we consider is the one-dimensional Volterra–Lotka equation

$$dx(t) = x(t)\left(2 - \frac{5}{2}x(t)\right)dt + \frac{3}{2}x(t)\,dW(t), \qquad t \in [0,1], \quad x(0) = 2. \quad (15)$$

The explicit solution of this equation is given by

$$x(t) = \frac{2\exp\left(\frac{7}{8}t + \frac{3}{2}W(t)\right)}{1 + 5\int_0^t \exp\left(\frac{7}{8}s + \frac{3}{2}W(s)\right)\,ds}, \quad t \in [0,1],$$

which we discretize with very small step size in order to estimate the pathwise maximum error. Here we will concentrate on comparing different auxiliary functions: Fig. 2

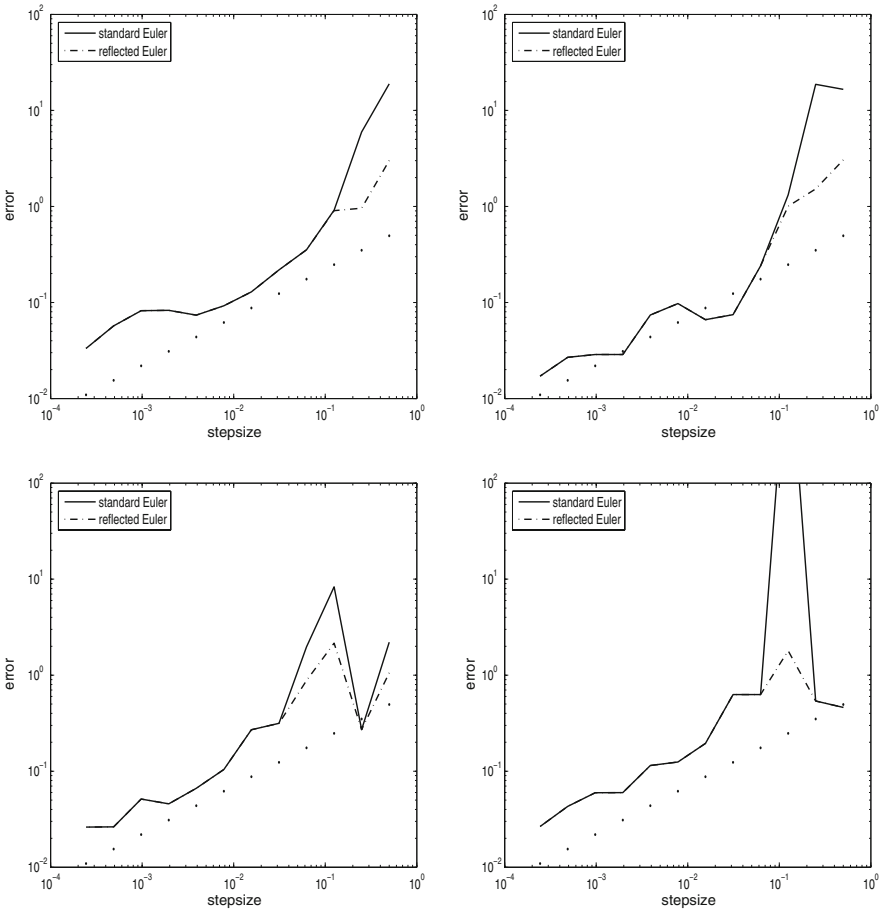

Fig. 2 Example (15): pathwise maximum error versus step size for four sample paths

shows for four different sample paths $\omega \in \Omega$ the maximum error in the discretization points for

(i) the Euler–Maruyama scheme $(-)$ with auxiliary functions $a(x) = 2x - \frac{5}{2}x^2$, $x < 0$ and $b(x) = \frac{3}{2}x$, $x < 0$ ("no auxiliary functions", i.e. the standard Euler method applied to Eq. 15),
(ii) the Euler–Maruyama scheme $(- \cdot -)$ with auxiliary functions $a(x) = -4x$, $x < 0$, $b(x) = 0$, $x < 0$ ("reflecting auxiliary functions").

The error of the standard Euler scheme with no auxiliary functions tends to explode for large step sizes, which is due to the quadratic term in Eq. (15). For smaller stepsizes the standard Euler approximation remains strictly positive and in this case the error of both schemes coincide and is in good accordance with the theoretically predicted order $0.5 - \varepsilon$. (The dotted line corresponds to the convergence order 0.5.)

The third example is the stochastic Duffing–van der Pol equation

$$\begin{aligned} dx_1(t) &= x_2(t)\,dt, \\ dx_2(t) &= (x_1(t) + 2x_2(t) - x_1^3(t) - x_1^2(t)x_2(t))\,dt \\ &\quad +0.5\,dW^{(1)}(t) + 2x_2(t)\,dW^{(2)}(t) \end{aligned} \tag{16}$$

with $x_1(0) = 1$, $x_2(0) = 1$. Here we consider the Euler scheme and the Milstein scheme using no auxiliary functions. Note that the entries of the diffusion matrix of Eq. (16) do not commute, since we have

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix} = L^2 \begin{pmatrix} 0 \\ \frac{1}{2} \end{pmatrix} \neq L^1 \begin{pmatrix} 0 \\ 2y \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Thus, we also have to approximate the iterated integrals

$$\int_{i/n}^{(i+1)/n} W_t^{(1)}\,dW_t^{(2)}, \quad i = 0, \ldots, n-1,$$

which appear in the Milstein scheme. For this we will use the Riemann sums

$$\sum_{l=0}^{n-1} W^{(1)}_{i/n+l/n^2}(W^{(2)}_{i/n+(l+1)/n^2} - W^{(2)}_{i/n+l/n^2}), \quad i = 0, \ldots, n-1,$$

see, e.g. [7] for a similar approximation of this iterated integral. (In a forthcoming article we will analyze the pathwise approximation of the iterated integrals systematically.)

To avoid the cumbersome computation of a "reference solution" using the Milstein scheme with very small stepsize, we will use the quantities

$$e(j) = \sup_{i=0,\ldots,2^j} \left| \overline{x}_{2^j}(i/2^j, \omega) - \overline{x}_{2^{j+1}}(i/2^j, \omega) \right|, \quad j = 1, 2, \ldots$$

to estimate the pathwise convergence rates.

Figure 3 shows for four different sample paths $\omega \in \Omega$ the quantities $\log_2(e(j))$ versus j for

(i) the Euler scheme (−) and
(ii) the Milstein scheme (− · −).

The dotted lines correspond to the convergence orders 0.5 and 1, respectively. Thus the estimated convergence rates are in accordance with the theoretically predicted rates.

4 Proofs

We first state some auxiliary results, which will be required.

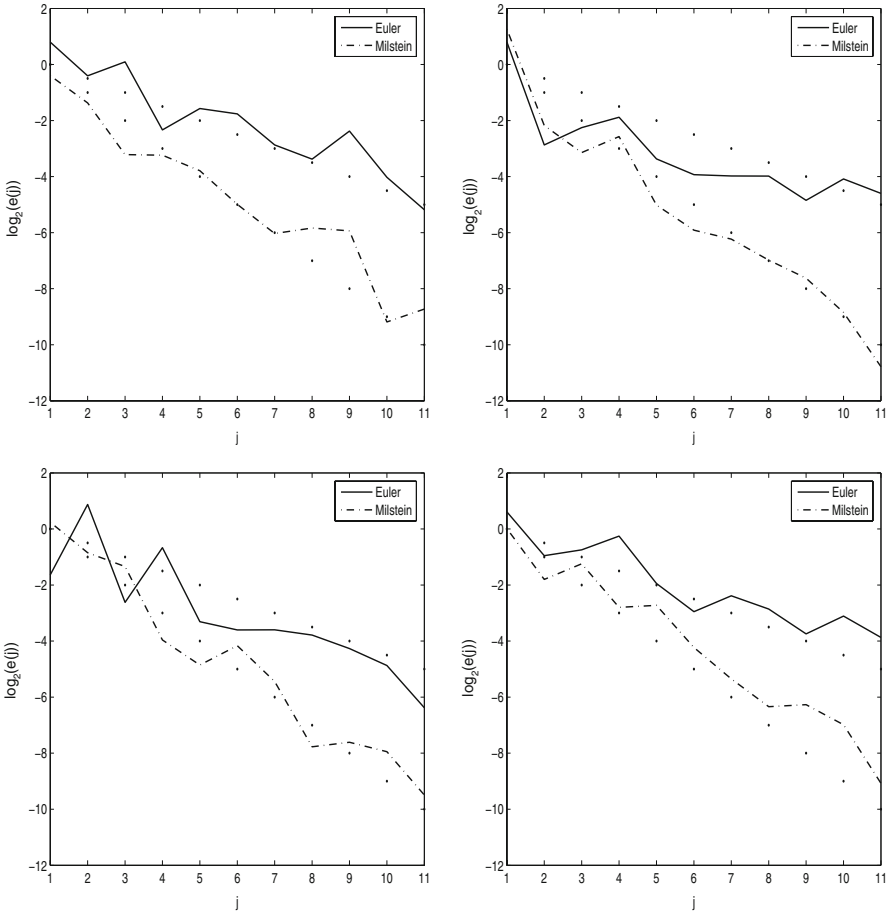

Fig. 3 Example (16): estimated convergence rates for four sample paths

4.1 Preliminaries

The following Lemma is a consequence of the Borel–Cantelli Lemma and will provide a relation between the convergence rates in the pth mean and the pathwise convergence rates. For a proof, see, e.g. [16].

Lemma 1 *Let $\alpha > 0$ and $K(p) \in [0, \infty)$ for $p \geq 1$. In addition, let Z_n, $n \in \mathbb{N}$, be a sequence of random variables such that*

$$(\mathbf{E}|Z_n|^p)^{1/p} \leq K(p) \cdot n^{-\alpha}$$

for all $p \geq 1$ and all $n \in \mathbb{N}$. Then for all $\varepsilon > 0$ there exists a finite and non-negative random variable η_ε such that

$$|Z_n| \leq \eta_\varepsilon \cdot n^{-\alpha+\varepsilon} \quad a.s.$$

for all $n \in \mathbb{N}$.

For the proof of Theorem 1 we will work with the following time-continuous versions of the Itô–Taylor schemes and of the modified Itô–Taylor schemes, respectively:

$$\overline{x}_n^\gamma(0) = x_0, \tag{17}$$
$$\overline{x}_n^\gamma(t) = \overline{x}_n^\gamma(iT/n) + \textstyle\sum_{\alpha \in \mathcal{A}_\gamma \setminus \{v\}} h_\alpha(\overline{x}_n^\gamma(iT/n)) \cdot I_\alpha(iT/n, t), \quad t \in (iT/n, (i+1)T/n],$$

and

$$\widetilde{x}_n^\gamma(0) = x_0, \tag{18}$$
$$\widetilde{x}_n^\gamma(t) = \widetilde{x}_n^\gamma(iT/n) + \textstyle\sum_{\alpha \in \mathcal{A}_\gamma \setminus \{v\}} \widetilde{h}_\alpha(\widetilde{x}_n^\gamma(iT/n)) \cdot I_\alpha(iT/n, t), \quad t \in (iT/n, (i+1)T/n],$$

for $i = 0, \ldots, 2n-1$. Note that the above schemes are based on the discretization

$$t_i = iT/n, \quad i = 0, 1, \ldots, 2n$$

of the interval $[0, 2T]$ and their restrictions to the nodes $t_i = iT/n$, $i = 0, \ldots, n-1$, coincide with the (modified) Itô–Taylor schemes introduced in Sects. 2.1 and 2.2. (The extension of the original schemes to the interval $[T, 2T]$ is required for the stopping time argument we use in the proof of Theorem 1.)

The next result we require is the well known upper bound for the error in the pth mean of the Itô–Taylor schemes under standard assumptions, see, e.g. [17].

Proposition 1 *Let* $\gamma = 0.5, 1.0, 1.5, \ldots$ *and* $f, g^j \in C_b^{2\gamma+1}(\mathbb{R}^d; \mathbb{R}^d)$ *for* $j = 1, \ldots, m$. *Moreover, assume that* $\mathbf{E}|x_0|^p < \infty$ *for all* $p \geq 1$ *and let* $\overline{x}_n^\gamma$ *be given by* (17). *Then, there exists for every* $p \geq 1$ *a constant* $K(p) > 0$ *such that*

$$\mathbf{E} \sup_{t \in [0,2T]} |x(t) - \overline{x}_n^\gamma(t)|^p \leq K(p) \cdot n^{-\gamma p}.$$

Combining Lemma 1 and Proposition 1 we obtain the following result on the pathwise convergence rates of the Itô–Taylor method given by (17).

Proposition 2 *Let* $\gamma = 0.5, 1.0, 1.5, \ldots$ *and* $f, g^j \in C_b^{2\gamma+1}(\mathbb{R}^d; \mathbb{R}^d)$ *for* $j = 1, \ldots, m$. *Moreover, assume that* $\mathbf{E}|x_0|^p < \infty$ *for all* $p \geq 1$ *and let* $\overline{x}_n^\gamma$ *be given by* (17). *Then there exists for every* $\varepsilon > 0$ *a finite and non-negative random variable* $\zeta_{\gamma,\varepsilon}$ *such that*

$$\sup_{t \in [0,2T]} \left|x(t) - \overline{x}_n^\gamma(t)\right| \leq \zeta_{\gamma,\varepsilon} \cdot n^{-\gamma+\varepsilon} \quad a.s.$$

for all $n \in \mathbb{N}$.

The proof of our main result is based on a localization procedure similar to the one used in [9]. Before we start with the proof, we give a description of its strategy.

Step 1. Here we choose an appropriate sequence of open and bounded sets $D_q \subset D$, $q \in \mathbb{N}$, such that

$$\ldots \subset D_q \subset D_{q+1} \subset \ldots$$

and

$$\cup_{q\in\mathbb{N}} D_q = D.$$

Moreover, we define the stopping times $\tau^{(q)}$ and $\tau_n^{(q)}$, which are the first exit times of the solution x from D_q and of the modified Itô–Taylor scheme $\widetilde{x}_n$ from D_q, respectively.

Step 2. Now we construct smooth coefficients f_q and g_q with compact support, which coincide with f and g on D_q ("truncated" coefficients). Moreover, we also truncate the initial value appropriately to $x_0^{(q)}$.

Step 3. For the new truncated coefficients f_q, g_q and the truncated initial value $x_0^{(q)}$ we define the corresponding "truncated" SDE, whose solution we denote by $x^{(q)}$. Moreover we also define the corresponding standard Itô–Taylor scheme $\overline{x}_n^{(q)}$ of order γ for $x^{(q)}$.
Since the drift coefficients f, f_q and the diffusion coefficients g, g_q coincide on D_q, the first exit times of the numerical schemes $\widetilde{x}_n$ and $\overline{x}_n^{(q)}$ from D_q are identical (Eq. 22). Moreover, we also show that the solution of the original SDE, i.e. x, leaves D_q earlier or at the same time than the solution $x^{(q)}$ of the truncated SDE (Eq. 24).

Step 4. Using Proposition 2, we obtain that the Itô–Taylor scheme $\overline{x}_n^{(q)}$ converges pathwise on $[0, 2T]$ to the solution $x^{(q)}$ of the truncated SDE with order $\gamma - \varepsilon$. (The coefficients of this SDE are smooth with compact support.)
Now recall that the truncated coefficients are identical to the untruncated coefficients on D_q. Thus, as a consequence of Eq. (22), i.e. the coincidence of the first exit times, the modified Itô–Taylor scheme $\widetilde{x}_n$ and the Itô–Taylor scheme $\overline{x}_n^{(q)}$ coincide on $[0, \tau_n^{(q)})$. Moreover, since the solution x of the original SDE leaves D_q earlier or at the same time than the solution $x^{(q)}$ of the truncated SDE (Eq. 24), we have that x and $x^{(q)}$ coincide on $[0, \tau^{(q)})$. Hence on the interval $[0, \tau_n^{(q)} \wedge \tau^{(q)})$ the solutions of the original and the truncated SDE coincide and also the modified Itô–Taylor scheme for the original SDE and the Itô–Taylor scheme for the truncated SDE coincide. So we can conclude that $\widetilde{x}_n$ converges pathwise to x with order $\gamma - \varepsilon$ on $[0, \tau_n^{(q)} \wedge \tau^{(q)})$, see Eq. (27).

Step 5. However, asymptotically the modified Itô–Taylor scheme $\widetilde{x}_n$ does not leave the set D_q earlier than the exact solution x (Eq. 28). This follows from the pathwise convergence of the Itô–Taylor scheme $\overline{x}_n^{(q)}$ to the solution $x^{(q)}$ of the truncated SDE and the comparison of the first exit times for the numerical schemes $\widetilde{x}_n$, $\overline{x}_n^{(q)}$ and for the SDE solutions x, $x^{(q)}$.

Step 6. As a consequence of the previous step, which gives pathwise $\tau_n^{(q)} \geq \tau^{(q)}$ for $n \to \infty$, we have that $\widetilde{x}_n$ converges pathwise to x with order $\gamma - \varepsilon$ on every compact interval contained in $[0, \tau^{(q)})$, see Eq. (29) and the following remark.

Step 7. However, $\tau^{(q)}$ is the first exit time of x from D_q. Since x never leaves D and D_q converges to D, we have pathwise $\tau^{(q)} > T$ for $q \to \infty$. Hence, the assertion of the theorem follows, i.e. the modified Itô–Taylor scheme $\widetilde{x}_n$ converges pathwise to x with order $\gamma - \varepsilon$ on $[0, T]$.

4.2 Proof of Theorem 1

Now fix $\gamma \in \{0.5, 1.0, \ldots\}$ and let $f, g_j \in C^{2\gamma+1}(D; \mathbb{R}^d)$. For convenience in notation we drop the index γ in what follows.

(Step 1) For $q \in \mathbb{N}$ set

$$D_q = \{x \in D : |x| < q \text{ and } d(x, \partial D) > 1/q\}, \tag{19}$$

where

$$d(x, \partial D) = \inf\{|x - y| : y \in \partial D\}$$

for $x \in \mathbb{R}^d$. Moreover, define the stopping times

$$\tau^{(q)} = \inf\{t \geq 0 : x(t) \notin D_q\} \wedge 2T$$

and

$$\tau_n^{(q)} = \inf\{t \geq 0 : \widetilde{x}_n(t) \notin D_q\} \wedge 2T,$$

where $\widetilde{x}_n$ is the modified Itô–Taylor method on $[0, 2T]$ of order γ given by (18) with auxiliary functions a and b.

Since the sample paths of x are continuous functions, which are contained in D by assumption (A), we have that

$$\lim_{q \to \infty} \tau^{(q)} = 2T \quad \text{a.s.}$$

(Step 2) Moreover, for all $q \in \mathbb{N}$ there exists a function $\varphi_q \in C_b^\infty(\mathbb{R}^d; \mathbb{R})$ such that $0 \leq \varphi(x) \leq 1$ and

$$\varphi_q(x) = \begin{cases} 1 & \text{for } x \in \overline{D}_q, \\ 0 & \text{for } x \in \mathbb{R}^d \setminus D_{2q}. \end{cases}$$

Using these functions, set

$$f_q = f \cdot \varphi_q, \quad g_q^1 = g^1 \cdot \varphi_q, \ \ldots, \ g_q^m = g^m \cdot \varphi_q. \tag{20}$$

Clearly, we have $f_q, g_q^j \in C_b^{2\gamma+1}(\mathbb{R}^d;\mathbb{R}^d)$. We also need to truncate the initial value. Here we set

$$x_0^{(q)}(\omega) = \begin{cases} x_0(\omega) & \text{if } |x_0(\omega)| < q, \\ \frac{q}{|x_0(\omega)|} x_0(\omega) & \text{otherwise} \end{cases}$$

for $\omega \in \Omega$.

(Step 3) Now, denote by $\overline{x}_n^{(q)}$ the Itô–Taylor approximation on $[0, 2T]$ given by (17) of order γ applied to the truncated equation

$$\begin{aligned} dx^{(q)}(t) &= f_q(x^{(q)}(t))\,dt + \sum_{j=1}^m g_q^j(x^{(q)}(t))\,dW^j(t), \quad t \geq 0, \\ x^{(q)}(0) &= x_0^{(q)}. \end{aligned} \tag{21}$$

First note that clearly

$$\tau_n^{(q)} = \inf\{t \geq 0 : \widetilde{x}_n(t) \notin D_q\} \wedge 2T = \inf\{t \geq 0 : \overline{x}_n^{(q)}(t) \notin D_q\} \wedge 2T. \tag{22}$$

We moreover have

$$x(t \wedge \tau^{(q)}) 1_{\{\tau^{(q)}>0\}} = x^{(q)}(t \wedge \tau^{(q)}) 1_{\{\tau^{(q)}>0\}}, \quad t \geq 0, \tag{23}$$

almost surely, which implies

$$\tau^{(q)} = \inf\{t \geq 0 : x(t) \notin D_q\} \wedge 2T \leq \inf\{t \geq 0 : x^{(q)}(t) \notin D_q\} \wedge 2T \tag{24}$$

almost surely. Relation (23) is without doubt known, but since we could not find a suitable reference, we give a proof of it in the next subsection.

(Step 4) From Proposition 2 it follows that for every $\varepsilon > 0$, there exists a finite and non-negative random variable $\zeta_\varepsilon^{(q)}$ such that

$$\sup_{t\in[0,2T]} |x^{(q)}(t) - \overline{x}_n^{(q)}(t)| \leq \zeta_\varepsilon^{(q)} \cdot n^{-\gamma+\varepsilon} \quad \text{a.s.} \tag{25}$$

for all $n \in \mathbb{N}$. Since clearly

$$\overline{x}_n^{(q)}(t \wedge \tau_n^{(q)}) 1_{\{\tau^{(q)}>0\}} = \widetilde{x}_n(t \wedge \tau_n^{(q)}) 1_{\{\tau^{(q)}>0\}}, \quad t \in [0, 2T], \tag{26}$$

almost surely and

$$x(t \wedge \tau^{(q)}) 1_{\{\tau^{(q)}>0\}} = x^{(q)}(t \wedge \tau^{(q)}) 1_{\{\tau^{(q)}>0\}}, \quad t \in [0, 2T],$$

almost surely, see (23), we obtain

$$\sup_{t\in[0,\tau^{(q)}\wedge\tau_n^{(q)})} |x(t) - \widetilde{x}_n(t)| \leq \zeta_\varepsilon^{(q)} \cdot n^{-\gamma+\varepsilon} \quad \text{a.s.} \tag{27}$$

for all $n \in \mathbb{N}$.

(Step 5) Note that (25) in particular implies that

$$\lim_{n\to\infty} \sup_{t\in[0,T]} |x^{(q)}(t,\omega) - \overline{x}_n^{(q)}(t,\omega)| = 0$$

for almost all $\omega \in \Omega$, which yields

$$\liminf_{n\to\infty} \tau_n^{(q)} \geq \inf\{t \geq 0 : x^{(q)}(t) \notin D_q\} \wedge 2T \quad \text{a.s.}$$

Since

$$\tau^{(q)} \leq \inf\{t \geq 0 : x^{(q)}(t) \notin D_q\} \wedge 2T \quad \text{a.s.}$$

by (24), we obtain

$$\liminf_{n\to\infty} \tau_n^{(q)} \geq \tau^{(q)} \quad \text{a.s.} \tag{28}$$

(Step 6) By (27) and (28) we have

$$\limsup_{n\to\infty} \sup_{t\in[0,\tau^{(q)}(\omega)-T/2]} n^{\gamma-\varepsilon}|x(t,\omega) - \widetilde{x}_n(t,\omega)| \leq \zeta_\varepsilon^{(q)}(\omega)$$

for almost all $\omega \in \Omega$. Hence we obtain that

$$\eta_\gamma^{(q)a,b}(\omega) := \sup_{n\in\mathbb{N}} \sup_{t\in[0,\tau^{(q)}(\omega)-T/2]} n^{\gamma-\varepsilon}|x(t,\omega) - \widetilde{x}_n(t,\omega)| < \infty \tag{29}$$

for almost all $\omega \in \Omega$. (Note that (29) holds true for any interval $[0, \tau^{(q)}(\omega) - \delta]$ with $\delta > 0$. The choice $\delta = T/2$ is for simplicity only.)

It is only here that the auxiliary functions a and b contribute to the error bound: In fact, we have the decomposition

$$\begin{aligned}\eta_\gamma^{(q)a,b}(\omega) = \sup_{n\in\mathbb{N}} \Bigg(& \sup_{t\in[0,\tau^{(q)}(\omega)-T/2]} n^{\gamma-\varepsilon}|x(t,\omega) - \widetilde{x}_n(t,\omega)| 1_{\{\tau_n^{(q)}(\omega)\leq\tau^{(q)}(\omega)\}} \\ & + \sup_{t\in[0,\tau^{(q)}(\omega)-T/2]} n^{\gamma-\varepsilon}|x(t,\omega) - \widetilde{x}_n(t,\omega)| 1_{\{\tau_n^{(q)}(\omega)>\tau^{(q)}(\omega)\}} \Bigg),\end{aligned}$$

which splits the error in the parts, where $\widetilde{x}_n$ leaves D_q earlier than x (first summand) or not (second summand). Only in the first case the auxiliary functions matter.

(Step 7) Now define

$$\Omega_q = \{\omega \in \Omega : \tau^{(q)}(\omega) \geq 3T/2\}.$$

Since $\lim_{q\to\infty}\tau^{(q)} = 2T$ almost surely, it follows

$$\mathbf{P}\left(\cup_{q\in\mathbb{N}}\Omega_q\right) = 1.$$

Now set $\eta_\varepsilon^{a,b}(\omega) = \eta_\varepsilon^{(1)a,b}(\omega)$ for $\omega\in\Omega_1$ and $\eta_\varepsilon^{a,b}(\omega) = \eta_\varepsilon^{(q)a,b}(\omega)$ for $\omega\in\Omega_q\setminus\Omega_{q-1}$, $q\geq 2$. Then, we finally obtain

$$\sup_{t\in[0,T]}|x(t)-\widetilde{x}_n(t)| \leq \eta_\varepsilon^{a,b}\cdot n^{-\gamma+\varepsilon}$$

almost surely, which shows the assertion of Theorem 1. □

4.3 Proof of identity (23)

Set $z(t) = x(t) - x^{(q)}(t)$ and apply Itô's formula to $|z(t)|^2$, $t\in[0,2T]$. We obtain

$$\begin{aligned}|z(t)|^2 = |z(0)|^2 &+ 2\int_0^t \langle z(s), f(x(s)) - f_q(x^{(q)}(s))\rangle\, ds + \sum_{j=1}^m \int_0^t |g^j(x(s)) - g_q^j(x^{(q)}(s))|^2 ds\\ &+ 2\sum_{j=1}^m \int_0^t \langle z(s), g^j(x(s)) - g_q^j(x^{(q)}(s))\rangle\, dW^j(s), \quad t\in[0,2T],\end{aligned}$$

almost surely. Now define $h(t) = z(t\wedge\tau^{(q)})1_{\{\tau^{(q)}>0\}}$, $t\in[0,2T]$, and note that

$$\sup_{t\in[0,2T]}|h(t)| \leq 3q \quad \text{a.s.}$$

due to the definition of $\tau^{(q)}$ and $x^{(q)}$. Moreover, we have

$$\begin{aligned}|h(t)|^2 = 2&\int_0^{t\wedge\tau^{(q)}} \langle z(s), f(x(s)) - f_q(x^{(q)}(s))\rangle\, 1_{\{\tau^{(q)}>0\}}\, ds\\ &+ \sum_{j=1}^m \int_0^{t\wedge\tau^{(q)}} |g^j(x(s)) - g_q^j(x^{(q)}(s))|^2\, 1_{\{\tau^{(q)}>0\}} ds\\ &+ 2\sum_{j=1}^m \int_0^{t\wedge\tau^{(q)}} \langle z(s), g^j(x(s)) - g_q^j(x^{(q)}(s))\rangle\, dW^j(s)\, 1_{\{\tau^{(q)}>0\}}, \quad t\in[0,2T],\end{aligned}$$

almost surely.

By Proposition III.2.10 and the remark on page 147 in [15] we have

$$1_{\{\tau^{(q)}>0\}} \sum_{j=1}^{m} \int_0^{t\wedge\tau^{(q)}} \langle z(s), g^j(x(s)) - g_q^j(x^{(q)}(s))\rangle \, dW^j(s)$$
$$= 1_{\{\tau^{(q)}>0\}} \sum_{j=1}^{m} \int_0^{t} \langle z(s), g^j(x(s)) - g_q^j(x^{(q)}(s))\rangle 1_{\{s\le\tau^{(q)}\}} \, dW^j(s) \quad \text{a.s.}$$

for any $t \in [0, 2T]$. Since $\tau^{(q)}$ is a stopping time, the random variable $1_{\{\tau^{(q)}>0\}}$ is $\mathscr{F}_0$-measurable, and we obtain

$$1_{\{\tau^{(q)}>0\}} \sum_{j=1}^{m} \int_0^{t} \langle z(s), g^j(x(s)) - g_q^j(x^{(q)}(s))\rangle 1_{\{s\le\tau^{(q)}\}} \, dW^j(s)$$
$$= \sum_{j=1}^{m} \int_0^{t} \langle z(s), g^j(x(s)) - g_q^j(x^{(q)}(s))\rangle 1_{\{0<s\le\tau^{(q)}\}} \, dW^j(s) \quad \text{a.s.}$$

for any $t \in [0, 2T]$. (Compare, e.g. Exercise 2.30 in [15].) Now, note that the integrand on the right hand side of the above equation is uniformly bounded. Hence it follows

$$\mathbf{E} \sum_{j=1}^{m} \int_0^{t\wedge\tau^{(q)}} \langle z(s), g^j(x(s)) - g_q^j(x^{(q)}(s))\rangle \, dW^j(s) \, 1_{\{\tau^{(q)}>0\}} = 0, \quad t \in [0, 2T]$$

by Proposition III.2.10 in [15], and we have

$$\mathbf{E}|h(t)|^2 = 2\mathbf{E} \int_0^{t\wedge\tau^{(q)}} \langle z(s), f(x(s)) - f_q(x^{(q)}(s))\rangle \, 1_{\{\tau^{(q)}>0\}} \, ds$$
$$+ \sum_{j=1}^{m} \mathbf{E} \int_0^{t\wedge\tau^{(q)}} |g^j(x(s)) - g_q^j(x^{(q)}(s))|^2 \, 1_{\{\tau^{(q)}>0\}} ds$$
$$\le 2\mathbf{E} \int_0^{t} |z(s\wedge\tau^{(q)})||f(x(s\wedge\tau^{(q)})) - f_q(x^{(q)}(s\wedge\tau^{(q)}))| \, 1_{\{\tau^{(q)}>0\}} \, ds$$
$$+ \sum_{j=1}^{m} \mathbf{E} \int_0^{t} |g^j(x(s\wedge\tau^{(q)})) - g_q^j(x^{(q)}(s\wedge\tau^{(q)}))|^2 \, 1_{\{\tau^{(q)}>0\}} ds$$

for $t \in [0, 2T]$. Recall that

$$f(x) = f_q(x), \quad g_q^1(x) = g^1(x), \ldots, g_q^m(x) = g^m(x)$$

for $x \in \overline{D}_q$. Thus it follows

$$\begin{aligned}\mathbf{E}|h(t)|^2 &\le 2\mathbf{E}\int_0^t |z(s\wedge\tau^{(q)})||f_q(x(s\wedge\tau^{(q)})) - f_q(x^{(q)}(s\wedge\tau^{(q)}))|\, 1_{\{\tau^{(q)}>0\}}\, ds \\ &\quad + \sum_{j=1}^m \mathbf{E}\int_0^t |g_q^j(x(s\wedge\tau^{(q)})) - g_q^j(x^{(q)}(s\wedge\tau^{(q)}))|^2\, 1_{\{\tau^{(q)}>0\}} ds\end{aligned}$$

for $t \in [0, 2T]$. Since $f_q, g_q^j \in C_b^{2\gamma+1}(\mathbb{R}^d; \mathbb{R}^d)$ we obtain

$$\begin{aligned}\mathbf{E}|h(t)|^2 &\le C_1 \int_0^t \mathbf{E}|z(s\wedge\tau^{(q)})|^2 1_{\{\tau^{(q)}>0\}}\, ds + C_1 \sum_{j=1}^m \int_0^t \mathbf{E}|z(s\wedge\tau^{(q)})|^2 1_{\{\tau^{(q)}>0\}}\, ds \\ &\le C_2 \int_0^t \mathbf{E}|h(s)|^2\, ds\end{aligned}$$

for $t \in [0, 2T]$ with $C_1, C_2 > 0$. Now, Gronwall's Lemma yields that

$$\mathbf{E}|h(t)|^2 = 0, \quad t \in [0, 2T],$$

and the continuity of the sample paths of x implies that

$$h(t) = 0, \quad t \in [0, 2T],$$

almost surely, which shows (23). □

Acknowledgments The authors would like to thank the referees for their valuable comments and remarks.

References

1. Alfonsi, A.: On the discretization schemes for the CIR (and Bessel squared) processes. Monte Carlo Methods Appl. **11**, 355–384 (2005)
2. Allen, E.: Modelling with Itô Stochastic Differential Equations. Springer, Dordrecht (2007)
3. Arnold, L.: Random Dynamical Systems. Springer, Berlin (1998)
4. Bossy, M., Diop, A.: An efficient discretisation scheme for one dimensional SDEs with a diffusion coefficient function of the form $|x|^\alpha$, $\alpha \in [1/2, 1)$. Working paper (2004)
5. Deelstra, G., Delbaen, F.: Convergence of discretized stochastic (interest rate) processes with stochastic drift term. Appl. Stoch. Models Data Anal. **14**, 77–84 (1998)
6. Fleury, G.: Convergence of schemes for stochastic differential equations. Prob. Eng. Mech. **21**, 35–43 (2005)

7. Gaines, J.G., Lyons, T.J.: Variable step size control in the numerical solution of stochastic differential equations. SIAM J. Appl. Math. **57**, 1455–1484 (1997)
8. Grüne, L., Kloeden, P.E.: Pathwise approximation of random ordinary differential equations. BIT **41**, 711–721 (2001)
9. Gyöngy, I.: A note on Euler's approximations. Potential Anal. **8**, 205–216 (1998)
10. Higham, D.J., Mao, X.: Convergence of Monte Carlo simulations involving the mean-reverting square root process. J. Comp. Finance **8**, 35–61 (2005)
11. Higham, D.J., Mao, X., Stuart, A.M.: Strong convergence of Euler-type methods for nonlinear stochastic differential equations. SIAM J. Numer. Anal. **40**, 1041–1063 (2002)
12. Hu, Y.: Semi-implicit Euler–Maruyama scheme for stiff stochastic equations. In: Koerezlioglu, H. (ed.) Stochastic Analysis and Related Topics V: The Silvri Workshop, Progr. Prob. 38, Boston, pp. 183–202 (1996)
13. Kahl, C., Günther, M., Rossberg, T.: Structure preserving stochastic integration schemes in interest rate derivative modeling. Appl. Numer. Math. **58**, 284–295 (2008)
14. Karlin, S., Taylor, H.M.: A Second Course in Stochastic Processes. Academic Press, New York (1981)
15. Karatzas, I., Shreve, S.E.: Brownian Motion and Stochastic Calculus, 2nd edn. Springer, New York (1991)
16. Kloeden, P.E., Neuenkirch, A.: The pathwise convergence of approximation schemes for stochastic differential equations. LMS JCM **10**, 235–253 (2007)
17. Kloeden, P.E., Platen, E.: Numerical Solution of Stochastic Differential equations, 3rd edn. Springer, Berlin (1999)
18. Lamba, H., Mattingly, J.C., Stuart, A.M.: An adaptive Euler–Maruyama scheme for SDEs: convergence and stability. IMA J. Numer. Anal. **27**, 479–506 (2007)
19. Lord, R., Koekkoek, R., van Dijk, D.J.C.: A Comparison of Biased Simulation Schemes for Stochastic Volatility Models. Working Paper (2008)
20. Mao, X., Marion, G., Renshaw, E.: Environmental Brownian noise suppresses explosions in populations dynamics. Stoch. Process. Apl. **97**, 95–110 (2002)
21. Mao, X.: Stochastic Differential Equations and their Applications. Horwood Publishing, Chichester (1997)
22. Milstein, G.N., Tretjakov, M.V.: Numerical integration of stochastic differential equations with non-globally Lipschitz coefficients. SIAM J. Numer. Anal. **43**, 1139–1154 (2005)
23. Ninomiya, S., Victoir, N.: Weak approximation of stochastic differential equations and application to derivative pricing. Appl. Math. Finance **15**, 107–121 (2008)
24. Talay, D.: Résolution trajectorielle et analyse numérique des équations différentielles stochastiques. Stochastics **9**, 275–306 (1983)
25. Yan, L.: The Euler scheme with irregular coefficients. Ann. Probab. **30**, 1172–1194 (2002)