*London Mathematical Society* ISSN 1461–1570

# THE PATHWISE CONVERGENCE OF APPROXIMATION SCHEMES FOR STOCHASTIC DIFFERENTIAL EQUATIONS

P. E. KLOEDEN AND A. NEUENKIRCH

*Abstract*

We study approximation methods for stochastic differential equations and point out a simple relation between their order of convergence in the $p$th mean and their order of convergence in the pathwise sense: Convergence in the $p$th mean of order $\alpha$ for all $p \geqslant 1$ implies pathwise convergence of order $\alpha - \varepsilon$ for arbitrary $\varepsilon > 0$. We apply this result to several one-step and multi-step approximation schemes for stochastic differential equations and stochastic delay differential equations. In addition, we give some numerical examples.

## 1. *Introduction*

Approximation schemes for Itô stochastic differential equations of the form

$$dX(t) = a(X(t))\,dt + \sum_{j=1}^{m} b^j(X(t))\,dW^j(t), \qquad t \in [0,T], \tag{1}$$
$$X(0) = X_0 \in \mathbb{R}^d,$$

where $a, b^j : \mathbb{R}^d \to \mathbb{R}^d$, $j = 1,\ldots,m$, and $W^j(t)$, $t \in [0,T]$, $j = 1,\ldots,m$, are $m$ independent Brownian motions on a given probability space $(\Omega, \mathcal{F}, \mathbf{P})$, have been intensively studied in the recent years. For an overview, see, for example, [**9**] or [**12**]. The vast majority of results, however, are concerned with error criteria that measure the error of the approximation on average. For instance, in the case of the so called 'weak approximation' the error of an approximation $\overline{X}$ to $X$ is measured by the quantity

$$|\mathbf{E}\phi(X(T)) - \mathbf{E}\phi(\overline{X}(T))|$$

for (smooth) functions $\phi : \mathbb{R}^d \to \mathbb{R}$, while for the 'strong approximation' problem the $p$th mean of the difference between $X$ and $\overline{X}$ is considered; that is,

$$\left(\mathbf{E} \sup_{i=0,\ldots,n} \left|X(t_i) - \overline{X}(t_i)\right|^p\right)^{1/p}$$

for $p \geqslant 1$, where $|\cdot|$ denotes the Euclidean norm and $0 = t_0 \leqslant t_1 \leqslant \ldots \leqslant t_n = T$ are the time nodes of the discretization. In the latter case, usually the mean-square error (that is, $p = 2$) is analyzed.

---

This work was supported by the DFG project 'Pathwise numerical analysis of stochastic evolution equations'.
Received 23 January 2007, revised 11 April 2007; *published* 11 June 2007.
2000 Mathematics Subject Classification 60H35, 65C30

In general, the numerical calculation of the approximation $\overline{X}$ is actually carried out path by path; that is, the real numbers $\overline{X}(t_1, \omega), \ldots, \overline{X}(t_n, \omega)$ are calculated for a fixed $\omega \in \Omega$. In spite of this fact, only a few articles deal with the pathwise error

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

which is a random quantity but gives more information about the error of the calculated approximations $\overline{X}(t_1, \omega), \ldots, \overline{X}(t_n, \omega)$ of $X(t_1, \omega), \ldots, X(t_n, \omega)$ for a fixed $\omega \in \Omega$.

In [**16**] an upper bound for the pathwise error of the Milstein method is determined using the Doss–Sussmann approach to transform the stochastic differential equation and the Milstein scheme to a random ordinary differential equation and a corresponding approximation scheme, respectively. The pathwise approximation of random ordinary differential equations is considered in [**4**], where the Euler and Heun methods are analyzed. Moreover, it is shown that the classical convergence rates of these schemes can be retained by averaging the noise over the discretization subintervals. Gyöngy [**5**] shows that the explicit Euler–Maruyama scheme with equidistant step size $1/n$ converges pathwise with order $1/2 - \varepsilon$ for arbitrary $\varepsilon > 0$. Hence the pathwise and the mean-square rate of convergence of the Euler method almost coincide. Using an idea in the proof of [**5**], we will show here that this is not an exceptional case, but is, in fact, the rule due to the following result.

‘*If a sequence of random variables converges to zero with order of convergence $\alpha > 0$ in the pth mean for all $p \geqslant 1$, then this sequence of random variables converges also almost surely to zero with pathwise order of convergence $\alpha - \varepsilon$ for arbitrary $\varepsilon > 0$.*’

This principle applies directly to the strong Itô–Taylor approximation schemes for equation (1); see, for example, [**9**]. For instance, the Milstein scheme has consequently pathwise order of convergence of order $1 - \varepsilon$ for arbitrary $\varepsilon > 0$. The use of the above result is by no means restricted to Itô–Taylor schemes or to stochastic ordinary differential equations. As a further example we will consider a particular two-step Maruyama scheme, the stochastic Adams–Moulton-2 scheme (see [**2, 3**]), and will determine its pathwise rate of convergence. In addition, we will also consider the Euler method for stochastic delay differential equations.

The article is structured as follows. In Section 2 we state and prove our main result and consider the above-mentioned examples. Our results are then illustrated by numerical test examples in Section 3.

## 2. *Main result and examples*

The following simple lemma is the link between the convergence rates in the $p$th mean and the pathwise convergence rates.

LEMMA 2.1. *Let $\alpha > 0$ and $K(p) \in [0, \infty)$ for $p \geqslant 1$. In addition, let $Z_n$, $n \in \mathbb{N}$, be a sequence of random variables such that*

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

*for all* $p \geqslant 1$ *and all* $n \in \mathbb{N}$. *Then for all* $\varepsilon > 0$ *there exists a random variable* $\eta_\varepsilon$ *such that*

$$|Z_n| \leqslant \eta_\varepsilon \cdot n^{-\alpha+\varepsilon} \qquad \textit{almost surely}$$

*for all* $n \in \mathbb{N}$. *Moreover,* $\mathbf{E}|\eta_\varepsilon|^p < \infty$ *for all* $p \geqslant 1$.

*Proof.* Fix $\varepsilon > 0$ and $p > 1/\varepsilon$. Then for all $\delta > 0$ from the Chebyshev–Markov inequality and the assumptions of the lemma we obtain

$$\mathbf{P}(n^{\alpha-\varepsilon}|Z_n| > \delta) \leqslant \frac{\mathbf{E}|Z_n|^p}{\delta^p} n^{(\alpha-\varepsilon)p} \leqslant \frac{K(p)^p}{\delta^p} n^{-p\varepsilon}.$$

Since $p > 1/\varepsilon$ we have

$$\sum_{n=1}^{\infty} \mathbf{P}(n^{\alpha-\varepsilon}|Z_n| > \delta) < \infty$$

for all $\delta > 0$. The Borel–Cantelli lemma then implies that $Z_n \to 0$ almost surely for $n \to \infty$. Now set $\eta_\varepsilon = \sup_{n\in\mathbb{N}} n^{\alpha-\varepsilon}|Z_n|$. It follows that

$$\mathbf{E}|\eta_\varepsilon|^q = \mathbf{E} \sup_{n\in\mathbb{N}} n^{(\alpha-\varepsilon)q}|Z_n|^q \leqslant \sum_{n=1}^{\infty} n^{(\alpha-\varepsilon)q}\mathbf{E}|Z_n|^q \leqslant K(q)^q \sum_{n=1}^{\infty} n^{-q\varepsilon} < \infty$$

for $q > 1/\varepsilon$. Applying Jensen's inequality we obtain $\mathbf{E}|\eta_\varepsilon|^q < \infty$ for all $q \geqslant 1$. The assertion of the lemma now follows by

$$|Z_n| \leqslant \left(\sup_{n\in\mathbb{N}} n^{\alpha-\varepsilon}|Z_n|\right) \cdot n^{-\alpha+\varepsilon} = \eta_\varepsilon \cdot n^{-\alpha+\varepsilon}.$$

□

The pathwise rate of convergence of an approximation method $\overline{X}_n$ for equation (1) can thus be determined by calculating its convergence rate in the $p$th mean, just by applying the above lemma with $Z_n = |X(T) - \overline{X}_n(T)|$ or $Z_n = \sup_{k=0,\ldots,n} |X(t_k) - \overline{X}_n(t_k)|$. The following examples will illustrate that Lemma 2.1 is a powerful tool for Itô stochastic differential equations due to the Burkholder–Davis–Gundy inequality; see, for example, [**14**]. However, for other types of stochastic differential equations it may be more appropriate to determine the pathwise rate of convergence by direct methods. See, for example, [**13**] for stochastic differential equations driven by fractional Brownian motion.

For simplicity we will consider only equidistant discretizations $t_i = (i/n) \cdot T$, $i = 0, \ldots, n$, but the following examples can be easily generalized to non-equidistant discretizations.

## 2.1. *Itô–Taylor schemes*

The first class of approximation schemes that we consider are the Itô–Taylor schemes. For convenience, we recall their definition here.

Let

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

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 $\nu$ is the multi-index of length 0. Moreover, let $n(\alpha)$ be the number

https://doi.org/10.1112/S1461157000001388 Published online by Cambridge University Press

of entries of $\alpha$ which are equal to 0. For $\alpha = (j_1, \dots, j_l)$ and $0 \leqslant s \leqslant t \leqslant T$ we define

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

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

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

and

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

for $j \in \{1, \dots, m\}$. Here $a^k$, $b^{k,j}$ are the $k$th components of $a$ and $b^j$, respectively.

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

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

Then the Itô–Taylor scheme of order $\gamma$ is defined as

$$\begin{aligned} \overline{X}_n^\gamma(t_0) &= X_0, \\ \overline{X}_n^\gamma(t_{i+1}) &= \overline{X}_n^\gamma(t_i) + \sum_{\alpha \in \mathcal{A}_\gamma \setminus \{\nu\}} f_\alpha(\overline{X}_n^\gamma(t_i)) \cdot I_\alpha(t_i, t_{i+1}) \end{aligned}$$

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

$$f_\alpha(x) = L^{j_1} \cdots L^{j_{l-1}} b^{j_l}(x)$$

for $\alpha = (j_1, \dots, j_l)$ and $b^0 = a$.

If the coefficients of equation (1) are sufficiently regular, then it is well known that

$$\left( \mathbf{E} \sup_{i=0,\dots,n} |X(t_i) - \overline{X}_n^\gamma(t_i)|^p \right)^{1/p} \leqslant K(p) \cdot n^{-\gamma}$$

for all $p \geqslant 1$ and appropriate $K(p) \in [0, \infty)$; see, for example, [9, Chapter 10]. Thus it follows from Lemma 2.1 that the Itô–Taylor scheme of order $\gamma$ has pathwise convergence order $\gamma - \varepsilon$ for arbitrary $\varepsilon > 0$; that is,

$$\sup_{i=0,\dots,n} |X(t_i) - \overline{X}_n^\gamma(t_i)| \leqslant \eta_{\varepsilon,\gamma} \cdot n^{-\gamma+\varepsilon} \qquad \text{almost surely,}$$

where $\eta_{\varepsilon,\gamma}$ is a random variable with all moments finite. Note that for $\gamma = 0.5$ we recover the result of [5], since the strong order 0.5 Itô–Taylor scheme is the Euler–Maruyama method:

$$\begin{aligned} \overline{X}_n^E(0) &= X_0, \\ \overline{X}_n^E(t_{i+1}) &= \overline{X}_n^E(t_i) + a(\overline{X}_n^E(t_i))(t_{i+1} - t_i) + \sum_{j=1}^m b^j(\overline{X}_n^E(t_i))(W^j(t_{i+1}) - W^j(t_i)) \end{aligned}$$

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

For $\gamma = 1.0$ we obtain the Milstein scheme, which is given by

$$\overline{X}_n^M(0) = X_0,$$

$$\overline{X}_n^M(t_{i+1}) = \overline{X}_n^M(t_i) + a(\overline{X}_n^M(t_i))(t_{i+1} - t_i) + \sum_{j=1}^m b^j(\overline{X}_n^M(t_i))(W^j(t_{i+1}) - W^j(t_i)) + \sum_{j_1,j_2=1}^m L^{j_1} b^{j_2}(\overline{X}_n^M(t_i)) I_{j_1,j_2}(t_i, t_{i+1})$$

for $i = 0, \dots, n-1$. Hence the Milstein scheme has pathwise order of convergence $1 - \varepsilon$. This improves the upper bound $1/2 - \varepsilon$ given in [16].

REMARK 2.2. Note that the random variables $\eta_{\varepsilon,\gamma}$ are not explicitly known in general. In a forthcoming paper we will study the question of whether $\eta_{\varepsilon,\gamma}$ can be replaced or estimated by a random variable that depends on the computed values $\overline{X}_n(t_0), \dots, \overline{X}_n(t_n)$ and on the driving Wiener processes $W^1, \dots, W^m$ in a simple way.

REMARK 2.3. For the Euler method and the Milstein method, the asymptotic distribution of $X(T) - \overline{X}_n(T)$ is known; see [8] and [17]. For instance, in the case $m = d = 1$ we have for the Euler method that

$$\sqrt{n} \cdot \left( X(T) - \overline{X}_n^E(T) \right) \xrightarrow{\mathcal{L}} U(T)$$

for $n \to \infty$. The process $U(t)$, $t \in [0, T]$, satisfies the stochastic differential equation

$$U_t = \int_0^t a'(X(s))U(s)\, ds + \int_0^t b'(X(s))U(s)\, dW(s) - \frac{1}{\sqrt{2}} \int_0^t b'b(X(s))\, dB(s),$$

where $B(t), t \in [0, T]$, is a Brownian motion independent of $W(t)$, $t \in [0, T]$; see [8]. For the Milstein scheme one obtains in the case $m = d = 1$ that

$$n \cdot \left( X(T) - \overline{X}_n^M(T) \right) \xrightarrow{\mathcal{L}} \widetilde{U}(T)$$

for $n \to \infty$, where $\widetilde{U}(t), t \in [0, T]$, is the solution of the stochastic differential equation

$$\widetilde{U}(t) = \int_0^t a'(X(s))\widetilde{U}(s)\, ds + \int_0^t b'(X(s))\widetilde{U}(s)\, dW(s) - \frac{1}{2} \int_0^t a'a(X(s))\, ds - \frac{1}{2} \int_0^t c_0(X(s))\, ds - \frac{1}{\sqrt{12}} \int_0^t c_1(X(s))\, dB^1(s) - \frac{1}{\sqrt{6}} \int_0^t c_2(X(s))\, dB^2(s).$$

Here $B^1(t), B^2(t), t \in [0, T]$, are independent Brownian motions, which are independent of $W(t)$, $t \in [0, T]$ and $c_0, c_1$ and $c_2$ are three functions dependent only on $a$ and $b$; see [17].

Thus, the pathwise convergence rates for the Euler and Milstein schemes, obtained in [5] and this article, are sharp.

### 2.2. *Stochastic Adams–Moulton-2 scheme*

Another class of approximation methods for equation (1) are the stochastic multi-step methods, which are a generalization of the multi-step methods for deterministic


https://doi.org/10.1112/S1461157000001388 Published online by Cambridge University Press

ordinary differential equations. See, for example, [2] and [3]. For example, the linear two-step Maruyama schemes are given by

$$\sum_{l=0}^{2} \alpha_l \overline{X}_n(t_{i-l}) = \frac{T}{n} \sum_{l=0}^{2} \beta_l a(\overline{X}_n(t_{i-l})) + \sum_{l=1}^{2} \gamma_l \sum_{j=1}^{m} b^j(\overline{X}_n(t_{i-l}))(W^j(t_{i-l+1}) - W^j(t_{i-l}))$$

for $i = 2, \ldots, n$ with coefficients $\alpha_l, \beta_l, \gamma_l \in \mathbb{R}$. For the required second initial value one has to use a properly chosen approximation $\overline{X}_n(t_1)$ of $X(t_1)$. These schemes are numerically mean-square stable (see [3]), if the coefficients $\alpha_0, \alpha_1, \alpha_2$ satisfy Dahlquist's root condition, which is well known in the deterministic case (see, for example, [6]): the roots of the polynomial

$$\rho(\xi) = \alpha_0 \xi^2 + \alpha_1 \xi + \alpha_2$$

have to lie on or within the unit circle and the roots on the unit circle have to be simple. Moreover, if $\overline{X}_n(t_1)$ is an approximation to $X(t_1)$ of mean-square order 0.5 and if the coefficients of the linear two-step Maruyama scheme satisfy the consistency conditions

$$\alpha_0 + \alpha_1 + \alpha_2 = 0, \quad 2\alpha_0 + \alpha_1 = \beta_0 + \beta_1 + \beta_2, \quad \alpha_0 = \gamma_1, \quad \alpha_0 + \alpha_1 = \gamma_2, \tag{2}$$

then the mean-square order of convergence of these linear two-step Maruyama schemes is 0.5; see [3].

One particular two-step Maruyama scheme, which we will consider in detail as an illustrative example, is the Adams–Moulton-2 scheme

$$\overline{X}_n(t_{i+1}) = \overline{X}_n(t_i) + \left( \frac{5}{12} a(\overline{X}_n(t_{i+1})) + \frac{8}{12} a(\overline{X}_n(t_i)) - \frac{1}{12} a(\overline{X}_n(t_{i-1})) \right) \frac{T}{n} + \sum_{j=1}^{m} b^j(\overline{X}_n(t_i))(W^j(t_{i+1}) - W^j(t_i)), \qquad i = 1, \ldots, n-1; \tag{3}$$

see, for example, [2]. Here we provide the second initial value by a drift implicit Euler step; that is,

$$\overline{X}_n(0) = X_0, \qquad \overline{X}_n(t_1) = X_0 + \frac{T}{n} a(\overline{X}_n(t_1)) + \sum_{j=1}^{m} b^j(X_0) W^j(t_1). \tag{4}$$

Note that this Adams–Moulton-2 scheme is a drift implicit method, which is well defined for $n > N_* = 2L_a T$, where $L_a > 0$ is the Lipschitz constant of the drift coefficient $a$. Moreover, its coefficients are given by $\alpha_0 = 1, \alpha_1 = -1, \alpha_2 = 0$, $\beta_0 = \frac{5}{12}, \beta_1 = \frac{8}{12}, \beta_2 = -\frac{1}{12}$, $\gamma_0 = 1, \gamma_2 = 0$ and satisfy Dahlquist's root condition and the consistency condition (2).

THEOREM 2.4. *Let $a, b^j \in C^1(\mathbb{R}^d; \mathbb{R}^d)$ with bounded derivatives for $j = 1, \ldots, m$ and consider the approximation scheme defined by (3) and (4). Then for all $\varepsilon > 0$ there exists a random variable $\eta_\varepsilon$ with $\mathbf{E}|\eta_\varepsilon|^p < \infty$ for all $p \geqslant 1$ such that*

$$\sup_{i=0,\ldots,n} |X(t_i) - \overline{X}_n(t_i)| \leqslant \eta_\varepsilon \cdot n^{-1/2+\varepsilon} \qquad \textit{almost surely}$$

*for all $n > N_*$.*

*Proof.* All we have to show is that

$$\mathbf{E} \sup_{i=0,\ldots,n} |X(t_i) - \overline{X}_n(t_i)|^p \leqslant K(p)^p \cdot n^{-p/2} \tag{5}$$

for each $p \in \mathbb{N}$ and $n > N_*$ Then the assertion follows from Lemma 2.1 if we choose $Z_n = 0$ for $n \leqslant N_*$.

For convenience of notation we will drop the subscript $n$ in what follows. In addition, we will denote constants that depend only on $m$, $d$, $T$, $p$ and $a$, $b^j$ and their derivatives by $C$, regardless of their value.

From (3) and (4) we obtain

$$\begin{aligned}\overline{X}(t_l) = X_0 + \frac{T}{n}a(\overline{X}(t_1)) + \frac{T}{n}\sum_{i=1}^{l-1} \frac{5}{12}a(\overline{X}(t_{i+1})) + \frac{8}{12}a(\overline{X}(t_i)) - \frac{1}{12}a(\overline{X}(t_{i-1}))\\ + \sum_{i=0}^{l-1}\sum_{j=1}^{m} b^j(\overline{X}(t_i))(W^j(t_{i+1}) - W^j(t_i))\end{aligned} \tag{6}$$

for $l = 1, \ldots, n$, with the convention that $\sum_{i=1}^{0}(\ldots) = 0$. Moreover, we have

$$\begin{aligned}X(t_{i+1}) = X(t_i) + \left(\frac{5}{12}a(X(t_{i+1})) + \frac{8}{12}a(X(t_i)) - \frac{1}{12}a(X(t_{i-1}))\right)\frac{T}{n}\\ + \sum_{j=1}^{m} b^j(X(t_i))(W^j(t_{i+1}) - W^j(t_i)) + R_i^{(1)} + R_i^{(2)}\end{aligned} \tag{7}$$

with

$$\begin{aligned}R_i^{(1)} = \frac{5}{12}\int_{t_i}^{t_{i+1}} a(X(\tau)) - a(X(t_{i+1}))\, d\tau + \frac{7}{12}\int_{t_i}^{t_{i+1}} a(X(\tau)) - a(X(t_i))\, d\tau\\ - \frac{1}{12}\left(a(X(t_i)) - a(X(t_{i-1}))\right)\frac{T}{n}\end{aligned}$$

and

$$R_i^{(2)} = \sum_{j=1}^{m}\int_{t_i}^{t_{i+1}} b^j(X(\tau)) - b^j(X(t_i))\, dW^j(\tau)$$

for $i = 1, \ldots, n-1$. Iterating (7) yields

$$\begin{aligned}X(t_l) = X_0 + \frac{T}{n}a(X(t_1)) + \frac{T}{n}\sum_{i=1}^{l-1} \frac{5}{12}a(X(t_{i+1})) + \frac{8}{12}a(X(t_i)) - \frac{1}{12}a(X(t_{i-1}))\\ + \sum_{i=0}^{l-1}\sum_{j=1}^{m} b^j(X(t_i))(W^j(t_{i+1}) - W^j(t_i)) + \sum_{i=0}^{l-1} R_i^{(1)} + \sum_{i=0}^{l-1} R_i^{(2)}\end{aligned}$$

for $l = 1, \ldots, n$, where

$$R_0^{(1)} = \int_0^{t_1} a(X(\tau)) - a(X(t_1))\, d\tau.$$

Thus we have

$$
\begin{aligned}
&X(t_l) - \overline{X}(t_l) \\
&\quad = \frac{T}{n}(a(X(t_1)) - a(\overline{X}(t_1))) + \frac{5}{12}\frac{T}{n}\sum_{i=1}^{l-1} a(X(t_{i+1})) - a(\overline{X}(t_{i+1})) \\
&\qquad + \frac{8}{12}\frac{T}{n}\sum_{i=1}^{l-1} a(X(t_i)) - a(\overline{X}(t_i)) - \frac{1}{12}\frac{T}{n}\sum_{i=1}^{l-1} a(X(t_{i-1})) - a(\overline{X}(t_{i-1})) \\
&\qquad + \sum_{i=0}^{l-1}\sum_{j=1}^{m}\left[b^j(X(t_i)) - b^j(\overline{X}(t_i))\right](W^j(t_{i+1}) - W^j(t_i)) \\
&\qquad + \sum_{i=0}^{l-1} R_i^{(1)} + \sum_{i=0}^{l-1} R_i^{(2)}.
\end{aligned}
$$

Define $U_l = \sup_{i=0,\ldots,l} |X(t_i) - \overline{X}(t_i)|$ for $l = 0,\ldots,n$. Since $a$ is Lipschitz continuous due to our assumptions, we obtain that

$$
\begin{aligned}
U_l \leqslant C\frac{1}{n}\sum_{i=1}^{l} U_i + C\sup_{v=1,\ldots,l}\left|\sum_{i=0}^{v-1}\sum_{j=1}^{m}\left[b^j(X(t_i)) - b^j(\overline{X}(t_i))\right](W^j(t_{i+1}) - W^j(t_i))\right| \\
+ C\sup_{v=1,\ldots,l}\left|\sum_{i=0}^{v-1} R_i^{(1)}\right| + C\sup_{v=1,\ldots,l}\left|\sum_{i=0}^{v-1} R_i^{(2)}\right|
\end{aligned}
$$

for $l = 1,\ldots,n$. Now

$$
\left(\frac{1}{n}\sum_{i=1}^{l} U_i\right)^p \leqslant \frac{1}{n}\sum_{i=1}^{l} U_i^p
$$

by Jensen's inequality, so we have

$$
\begin{aligned}
\mathbf{E}U_l^p &\leqslant C\,\frac{1}{n}\sum_{i=1}^{l}\mathbf{E}U_i^p \\
&\quad + C\,\mathbf{E}\sup_{v=1,\ldots,l}\left|\sum_{i=0}^{v-1}\sum_{j=1}^{m}\left[b^j(X(t_i)) - b^j(\overline{X}(t_i))\right](W^j(t_{i+1}) - W^j(t_i))\right|^p \\
&\quad + C\,\mathbf{E}\sup_{v=1,\ldots,l}\left|\sum_{i=0}^{v-1} R_i^{(1)}\right|^p + C\,\mathbf{E}\sup_{v=1,\ldots,l}\left|\sum_{i=0}^{v-1} R_i^{(2)}\right|^p
\end{aligned} \tag{8}
$$

for $l = 1,\ldots,n$.

Recall that $b^{k,j}$ denotes the $k$th component of $b^j$ and set

$$
M^k(t) = \sum_{j=1}^{m}\int_0^t \sum_{i=0}^{n-1}\left[b^{k,j}(X(t_i)) - b^{k,j}(\overline{X}(t_i))\right] 1_{[t_i,t_{i+1})}(\tau)\,dW^j(\tau), \qquad t \in [0,T],
$$

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


https://doi.org/10.1112/S1461157000001388 Published online by Cambridge University Press

Hence we have

$$\sup_{v=1,\dots,l}\left|\sum_{i=0}^{v-1}\sum_{j=1}^{m}\left[b^j(X(t_i))-b^j(\overline{X}(t_i))\right](W^j(t_{i+1})-W^j(t_i))\right|^2=\sum_{k=1}^{d}\sup_{v=1,\dots,l}|M^k(t_v)|^2,$$

and, in addition,

$$\mathbf{E}\sup_{v=1,\dots,l}\left|\sum_{i=0}^{v-1}\sum_{j=1}^{m}(b^j(X(t_i))-b^j(\overline{X}(t_i))(W^j(t_{i+1})-W^j(t_i))\right|^p \leqslant C\sum_{k=1}^{d}\mathbf{E}\sup_{v=1,\dots,l}|M^k(t_v)|^p.$$

Note that the quadratic variation of $M^k$ is given by

$$\langle M^k\rangle(t)=\sum_{j=1}^{m}\int_0^t\sum_{i=0}^{n-1}|b^{k,j}(X(t_i))-b^{k,j}(\overline{X}(t_i))|^2 1_{[t_i,t_{i+1})}(\tau)\,d\tau,\qquad t\in[0,T].$$

By the Burkholder–Davis–Gundy inequality we have

$$\mathbf{E}\sup_{v=1,\dots,l}|M^k(t_v)|^p\leqslant C\,\mathbf{E}|\langle M^k\rangle(t_l)|^{p/2},$$

and, again, an application of Jensen's inequality and the assumptions on $b^j$ yield

$$\begin{aligned}\mathbf{E}\sup_{v=1,\dots,l}|M^k(t_v)|^p&\leqslant C\sum_{j=1}^{m}\int_0^{t_l}\sum_{i=0}^{n-1}\mathbf{E}\left|b^{k,j}(X(t_i))-b^{k,j}(\overline{X}(t_i))\right|^p 1_{[t_i,t_{i+1})}(\tau)\,d\tau\\&\leqslant C\sum_{j=1}^{m}\int_0^{t_l}\sum_{i=0}^{n-1}\mathbf{E}\left|X(t_i)-\overline{X}(t_i))\right|^p 1_{[t_i,t_{i+1})}(\tau)\,d\tau\\&\leqslant C\frac{1}{n}\sum_{i=1}^{l-1}\mathbf{E}U_i^p.\end{aligned}$$

Hence it follows that

$$\mathbf{E}\sup_{v=1,\dots,l}\left|\sum_{i=0}^{v-1}\sum_{j=1}^{m}(b^j(X(t_i))-b^j(\overline{X}(t_i))(W^j(t_{i+1})-W^j(t_i))\right|^p\leqslant C\frac{1}{n}\sum_{i=1}^{l-1}\mathbf{E}U_i^p$$

and inserting this in (8) yields

$$\mathbf{E}U_l^p\leqslant C\left(\frac{1}{n}\sum_{i=1}^{l}\mathbf{E}U_i^p+\mathbf{E}\sup_{v=1,\dots,l}\left|\sum_{i=0}^{v-1}R_i^{(1)}\right|^p+\mathbf{E}\sup_{v=1,\dots,l}\left|\sum_{i=0}^{v-1}R_i^{(2)}\right|^p\right)\tag{9}$$

for $l=1,\dots,n$.

For the first remainder term we obtain by Jensen's inequality

$$\mathbf{E}\sup_{v=1,\dots,l}\left|\sum_{i=0}^{v-1}R_i^{(1)}\right|^p\leqslant n^{p-1}\sum_{i=0}^{n-1}\mathbf{E}|R_i^{(1)}|^p.$$

Since

$$|R_i^{(1)}|^2 \leqslant C \sum_{k=1}^{d} \left| \int_{t_i}^{t_{i+1}} a^k(X(\tau)) - a^k(X(t_{i+1}))\, d\tau \right|^2 + C \sum_{k=1}^{d} \left| \int_{t_i}^{t_{i+1}} a^k(X(\tau)) - a^k(X(t_i))\, d\tau \right|^2 + C \sum_{k=1}^{d} \left| a^k(X(t_i)) - a^k(X(t_{i-1})) \right|^2 \frac{1}{n^2}$$

for $i = 1, \ldots, n-1$ and

$$|R_0^{(1)}|^2 \leqslant \sum_{k=1}^{d} \left| \int_0^{t_1} a^k(X(\tau)) - a^k(X(t_1))\, d\tau \right|^2,$$

we obtain again by Jensen's inequality and

$$\mathbf{E}|X(t) - X(s)|^p \leqslant C \cdot |t-s|^{p/2}, \qquad s,t \in [0,T], \tag{10}$$

that

$$\mathbf{E}|R_i^{(1)}|^p \leqslant Cn^{-3p/2}$$

for $i = 0, \ldots, n-1$. Thus we have

$$\mathbf{E} \sup_{v=1,\ldots,n} \left| \sum_{i=0}^{v-1} R_i^{(1)} \right|^p \leqslant Cn^{-p/2}. \tag{11}$$

For the second remainder term, set

$$N^k(t) = \sum_{j=1}^{m} \int_0^t \sum_{i=0}^{n-1} \left[ b^{k,j}(X(\tau)) - b^{k,j}(X(t_i)) \right] 1_{[t_i,t_{i+1})}(\tau)\, dW^j(\tau), \qquad t \in [0,T],$$

for $k = 1, \ldots, d$. We obtain

$$\mathbf{E} \sup_{v=1,\ldots,l} \left| \sum_{i=0}^{v-1} \sum_{j=1}^{m} \int_{t_i}^{t_{i+1}} b^j(X(\tau)) - b^j(X(t_i))\, dW^j(\tau) \right|^p \leqslant C \sum_{k=1}^{d} \mathbf{E} \sup_{v=1,\ldots,l} |N^k(t_v)|^p.$$

Since

$$\begin{aligned} \mathbf{E} \sup_{v=1,\ldots,l} |N_v^k|^p &\leqslant C\, \mathbf{E} \left| \sum_{j=1}^{m} \int_0^{t_l} \sum_{i=0}^{n-1} |b^{k,j}(X(\tau)) - b^{k,j}(X(t_i))|^2 1_{[t_i,t_{i+1})}(\tau)\, d\tau \right|^{p/2} \\ &\leqslant C \sum_{j=1}^{m} \int_0^{t_l} \sum_{i=0}^{n-1} \mathbf{E}|X(\tau) - X(t_i)|^p 1_{[t_i,t_{i+1})}(\tau)\, d\tau \\ &\leqslant Cn^{-p/2} \end{aligned}$$

by the Burkholder–Davis–Gundy inequality, Jensen's inequality and (10), it follows that

$$\mathbf{E} \sup_{v=1,\ldots,n} \left| \sum_{i=0}^{v-1} R_i^{(2)} \right|^p \leqslant Cn^{-p/2}. \tag{12}$$


https://doi.org/10.1112/S1461157000001388 Published online by Cambridge University Press

By inserting (11) and (12) in (9) we obtain

$$\mathbf{E}U_l^p \leqslant C\frac{1}{n}\sum_{i=1}^{l}\mathbf{E}U_i^p + Cn^{-p/2}.$$

Hence (5) follows by a discrete version of Gronwall's lemma; see, for example, [**12**, Lemma 1.3]. □

Thus for this approximation scheme too, the mean-square order of convergence and the pathwise order of convergence coincide up to an arbitrarily small $\varepsilon > 0$.

Alternatively, the above theorem can be shown by a reformulation of the Adams–Moulton-2 scheme as a perturbated one-step scheme and a combination of [**15**, Theorem 2.1] and Lemma 2.1. If $\alpha_2 = 0$, other two-step Maruyama schemes can be treated in a similar way. However, the above proofs do not apply if $\alpha_2 \neq 0$. Here a different method is required to control the error; see, for example, [**3**].

## 2.3. *Euler–Maruyama method for stochastic delay equations*

Now we will consider a different type of stochastic differential equations, namely Itô stochastic delay differential equations of the form

$$\begin{aligned} dX(t) &= a(X(t), X(t-\delta))\,dt + \sum_{j=1}^{m} b^j(X(t), X(t-\delta))\,dW^j(t), \quad && t \in [0,T], \\ X(t) &= \psi(t), && t \in [-\delta, 0], \end{aligned} \tag{13}$$

with a constant delay $\delta > 0$, initial path $\psi : [-\delta, 0] \to \mathbb{R}^d$ and $a, b^j : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}^d$ for $j = 1, \ldots, m$.

One of the simplest approximation schemes for equation (13) is the Euler method; see, for example, [**1**]. We will again consider only an equidistant discretization $t_i = (i/n)\cdot T$ for $i = 0, \ldots, n$ and, moreover, assume that the step size $T/n$ is an integral divisor of the delay $\delta > 0$; that is, $\delta = m_\delta \cdot (T/n)$ with $m_\delta \in \mathbb{N}$.

Then the Euler–Maruyama approximation of the delay equation (13) is:

$$\begin{aligned} \overline{X}_n(t_{i+1}) = \overline{X}_n(t_i) &+ a(\overline{X}_n(t_i), \overline{X}_n(t_i - \delta))\frac{T}{n} \\ &+ \sum_{j=1}^{m} b^j(\overline{X}_n(t_i), \overline{X}_n(t_i - \delta))(W^j(t_{i+1}) - W^j(t_i)) \end{aligned} \tag{14}$$

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

$$\overline{X}_n(0) = \psi(0), \qquad \overline{X}_n(t_i - \delta) = \psi(t_i - \delta), \quad t_i < \delta. \tag{15}$$

Determining the error of this approximation scheme in the $p$th mean and applying Lemma 2.1 yields the following result.

THEOREM 2.5. *Let $a, b^j \in C^1(\mathbb{R}^d \times \mathbb{R}^d; \mathbb{R}^d)$ with bounded derivatives for all $j = 1, \ldots, m$, $\psi \in C^{1/2}([-\delta, 0]; \mathbb{R}^d)$ and consider the approximation scheme defined by (14) and (15). Then for all $\varepsilon > 0$ there exists a random variable $\eta_\varepsilon$ with $\mathbf{E}|\eta_\varepsilon|^p < \infty$ for all $p \geqslant 1$ such that*

$$\sup_{i=0,\ldots,n} |X(t_i) - \overline{X}_n(t_i)| \leqslant \eta_\varepsilon \cdot n^{-1/2+\varepsilon} \qquad \textit{almost surely}$$

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

*Proof.* Under the above assumptions it is well known (see, for example, [**10**]) that the unique strong solution of the initial value problem (13) satisfies

$$\mathbf{E} \sup_{-\delta \leqslant t \leqslant T} |X(t)|^p < \infty \tag{16}$$

and

$$\mathbf{E}\,|X(t) - X(s)|^p \leqslant \widetilde{K}(p) \cdot |t-s|^{p/2} \tag{17}$$

with appropriate constants $\widetilde{K}(p) > 0$ for all $p \geqslant 1$.

Again we have to show that

$$\mathbf{E} \sup_{i=0,\dots,n} |X(t_i) - \overline{X}_n(t_i)|^p \leqslant K(p)^p \cdot n^{-p/2} \tag{18}$$

for all $p \in \mathbb{N}$. Then the assertion follows from an application of Lemma 2.1. For the proof of (18) in the case $p = 2$, see [**11**].

For convenience of notation we will again drop the subscript $n$ and we will denote constants that depend only on $m$, $d$, $T$, $p$, $\psi$ and $a$, $b^j$ and their derivatives by $C$, regardless of their value. Furthermore, we will use the notation $\Delta_i W^j = W^j(t_{i+1}) - W^j(t_i)$ for $i = 0, \dots, n-1$, $j = 1, \dots, m$.

Iterating (14) yields

$$\begin{aligned} \overline{X}(t_l) = X_0 &+ \frac{T}{n} \sum_{i=0}^{l-1} a(\overline{X}(t_i), \overline{X}(t_i - \delta)) \\ &+ \sum_{i=0}^{l-1} \sum_{j=1}^{m} b^j(\overline{X}(t_i), \overline{X}(t_i - \delta)) \Delta_i W^j \end{aligned} \tag{19}$$

for $l = 1, \dots, n$. For the exact solution we have

$$\begin{aligned} X(t_{i+1}) = X(t_i) &+ a(X(t_i), X(t_i - \delta)) \frac{T}{n} \\ &+ \sum_{j=1}^{m} b^j(X(t_i), X(t_i - \delta)) \Delta_i W^j + R_i^{(1)} + R_i^{(2)} \end{aligned} \tag{20}$$

with

$$R_i^{(1)} = \int_{t_i}^{t_{i+1}} a(X(\tau), X(\tau - \delta)) - a(X(t_i), X(t_i - \delta))\, d\tau$$

and

$$R_i^{(2)} = \sum_{j=1}^{m} \int_{t_i}^{t_{i+1}} b^j(X(\tau), X(\tau - \delta)) - b^j(X(t_i), X(t_i - \delta))\, dW^j(\tau)$$

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

From (20) we obtain

$$X(t_l) = X_0 + \frac{T}{n}\sum_{i=0}^{l-1} a(X(t_i), X(t_i-\delta)) + \sum_{i=0}^{l-1}\sum_{j=1}^{m} b^j(X(t_i), X(t_i-\delta))\Delta_i W^j + \sum_{i=0}^{l-1} R_i^{(1)} + \sum_{i=0}^{l-1} R_i^{(2)}. \tag{21}$$

Then combining (19) and (21) yields

$$X(t_l) - \overline{X}(t_l) = \frac{T}{n}\sum_{i=0}^{l-1} a(X(t_i), X(t_i-\delta) - a(\overline{X}(t_i), \overline{X}(t_i-\delta)) + \sum_{i=0}^{l-1}\sum_{j=1}^{m} \left[b^j(X(t_i), X(t_i-\delta)) - b^j(\overline{X}(t_i), \overline{X}(t_i-\delta))\right]\Delta_i W^j + \sum_{i=0}^{l-1} R_i^{(1)} + \sum_{i=0}^{l-1} R_i^{(2)}$$

for $l = 1, \ldots, n$. Thus, with $U_l = \sup_{i=0,\ldots,l} |X(t_i) - \overline{X}(t_i)|$, it follows that

$$\mathbf{E}U_l^p \leqslant C\frac{1}{n}\sum_{i=0}^{l-1}\mathbf{E}U_i^p + C\,\mathbf{E}\sup_{v=1,\ldots,l}\left|\sum_{i=0}^{v-1}\sum_{j=1}^{m}\left[b^j(X(t_i), X(t_i-\delta)) - b^j(\overline{X}(t_i), \overline{X}(t_i-\delta))\right]\Delta_i W^j\right|^p + C\,\mathbf{E}\sup_{v=1,\ldots,l}\left|\sum_{i=0}^{v-1} R_i^{(1)}\right|^p + C\,\mathbf{E}\sup_{v=1,\ldots,l}\left|\sum_{i=0}^{v-1} R_i^{(2)}\right|^p. \tag{22}$$

Now define

$$M^k(t) = \sum_{j=1}^{m}\int_0^t \sum_{i=0}^{n-1}\left[b^{k,j}(X(t_i), X(t_i-\delta)) - b^{k,j}(\overline{X}(t_i), \overline{X}(t_i-\delta))\right] 1_{[t_i,t_{i+1})}(\tau)\,dW^j(\tau)$$

for $t \in [0,T]$ and $k = 1, \ldots, d$. Clearly, we have

$$\mathbf{E}\sup_{v=1,\ldots,l}\left|\sum_{i=0}^{v-1}\sum_{j=1}^{m}\left[b^j(X(t_i), X(t_i-\delta)) - b^j(\overline{X}(t_i), \overline{X}(t_i-\delta))\right]\Delta_i W^j\right|^p \leqslant C\sum_{k=1}^{d}\mathbf{E}\sup_{v=1,\ldots,l}|M^k(t_v)|^p.$$

Since

$$\langle M^k\rangle(t) = \sum_{j=1}^{m}\int_0^t \sum_{i=0}^{n-1} |b^{k,j}(X(t_i), X(t_i-\delta)) - b^{k,j}(\overline{X}(t_i), \overline{X}(t_i-\delta))|^2\, 1_{[t_i,t_{i+1})}(\tau)\, d\tau$$

for $t \in [0, T]$, we obtain from the Burkholder–Davis–Gundy inequality, Jensen's inequality and the assumptions on $b^j$ that

$$\begin{aligned}
&\mathbf{E} \sup_{v=1,\dots,l} |M^k(t_v)|^p \\
&\quad\leqslant C\, \mathbf{E}|\langle M^k\rangle(t_l)|^{p/2} \\
&\quad\leqslant C \sum_{j=1}^{m}\int_0^{t_l} \sum_{i=0}^{n-1} \mathbf{E}|b^{k,j}(X(t_i), X(t_i-\delta)) - b^{k,j}(\overline{X}(t_i), \overline{X}(t_i-\delta))|^p 1_{[t_i,t_{i+1})}(\tau)\, d\tau \\
&\quad\leqslant C \sum_{j=1}^{m}\int_0^{t_l} \sum_{i=0}^{n-1} \left[\mathbf{E}|X(t_i) - \overline{X}(t_i)|^p + \mathbf{E}|X(t_i-\delta) - \overline{X}(t_i-\delta)|^p\right] 1_{[t_i,t_{i+1})}(\tau)\, d\tau \\
&\quad\leqslant C \frac{1}{n}\sum_{i=0}^{l-1} \mathbf{E} U_i^p.
\end{aligned}$$

Thus we have

$$\mathbf{E} \sup_{v=1,\dots,l} \left| \sum_{i=0}^{v-1}\sum_{j=1}^{m} \left[b^j(X(t_i), X(t_i-\delta)) - b^j(\overline{X}(t_i), \overline{X}(t_i-\delta))\right] \Delta_i W^j \right| \leqslant \frac{1}{n}\sum_{i=0}^{l-1} \mathbf{E} U_i^p$$

and inserting this in (22) yields

$$\mathbf{E} U_l^p \leqslant C \left( \frac{1}{n}\sum_{i=0}^{l-1} \mathbf{E} U_i^p + \mathbf{E} \sup_{v=1,\dots,l} \left|\sum_{i=0}^{v-1} R_i^{(1)}\right|^p + \mathbf{E} \sup_{v=1,\dots,l} \left|\sum_{i=0}^{v-1} R_i^{(2)}\right|^p \right) \tag{23}$$

for $l = 1, \dots, n$.

In view of (17) we have

$$\mathbf{E} \sup_{v=1,\dots,n} \left|\sum_{i=0}^{v-1} R_i^{(1)}\right|^p + \mathbf{E} \sup_{v=1,\dots,n} \left|\sum_{i=0}^{v-1} R_i^{(2)}\right|^p \leqslant C n^{-p/2}, \tag{24}$$

which can be shown to be completely analogous to (11) and (12) in the proof of Theorem 2.4. Hence (18) follows from (23), (24) and the discrete version of Gronwall's lemma; see, for example, [**12**, Lemma 1.3]. □

In [**7**] the Milstein method for stochastic delay differential equations is introduced, which has mean-square order of convergence 1.0. A similar consideration yields that this Milstein scheme also has pathwise order of convergence $1 - \varepsilon$ for arbitrary $\varepsilon > 0$.


https://doi.org/10.1112/S1461157000001388 Published online by Cambridge University Press

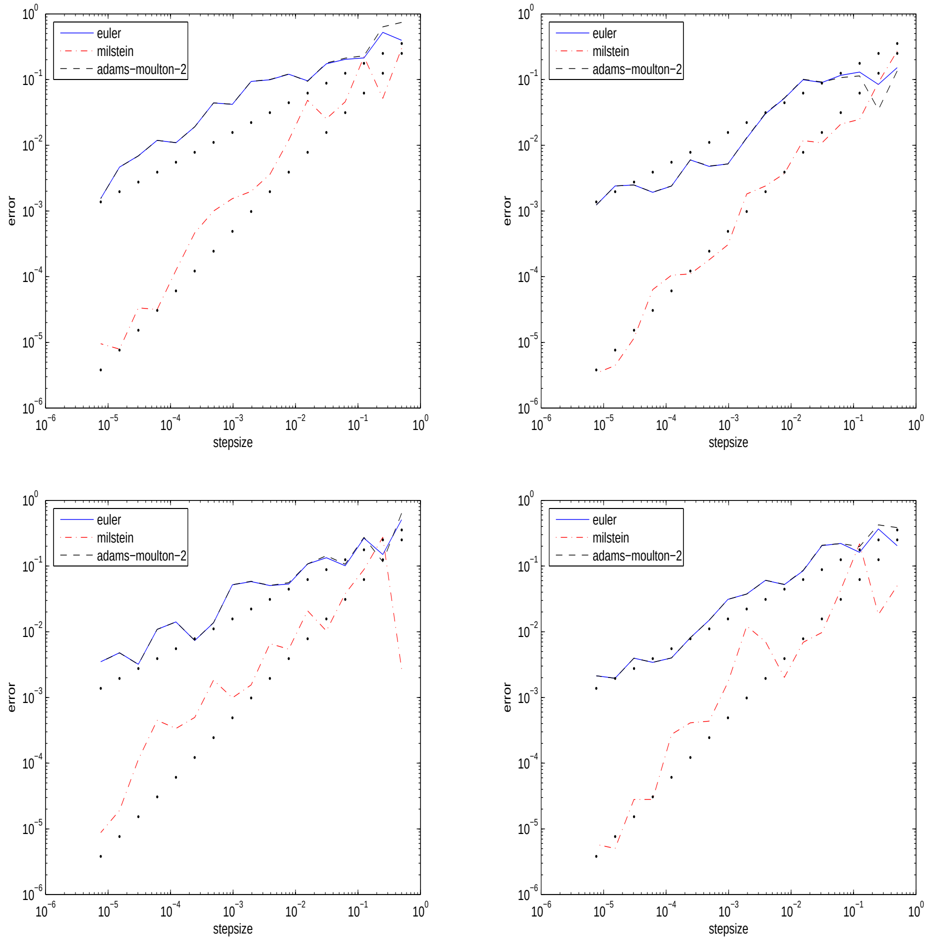


Figure 1: Example (25): pathwise maximum error vs. step size for four sample paths.

## 3. *Numerical examples*

In this section we illustrate our results with three numerical examples. The first example that we consider is the one-dimensional linear stochastic differential equation

$$dX(t) = 0.5X(t)\,dt + X(t)\,dW(t), \qquad X(0) = 1 \tag{25}$$

with exact solution

$$X(t) = \exp(W(t)).$$

Figure 1 shows the maximum error in the discretization points (that is, $\sup_{i=0,\ldots,n} |X(t_i,\omega) - \overline{X}_n(t_i,\omega)|$), which for brevity we call in the following 'pathwise maximum error', for the Euler–Maruyama (—), the Milstein (— · —) and the Adams–Moulton-2 (— —) scheme versus the step size for four different sample paths $\omega \in \Omega$.

https://doi.org/10.1112/S1461157000001388 Published online by Cambridge University Press

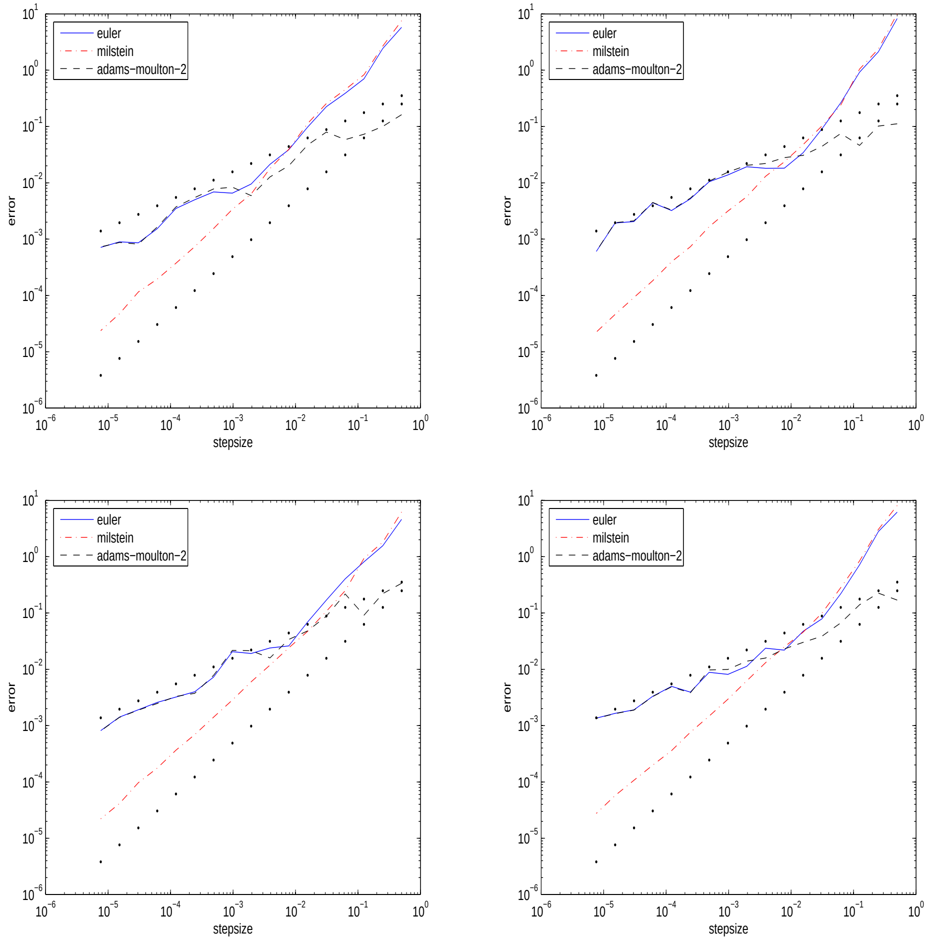


Figure 2: Example (26): pathwise maximum error vs. step size for four sample paths.

Since we use log-log-coordinates, the dotted lines correspond to the convergence orders 0.5 and 1, respectively. The errors of the Euler–Maruyama and Adams–Moulton-2 schemes differ only for large step sizes. This is quite natural, since both schemes coincide for equations without drift, and equation (25) is mainly determined by its diffusion part. Moreover, the pathwise convergence rates of all three approximation schemes are in good accordance with the theoretically predicted rates.

As second example we consider the linear equation

$$dX(t) = -5X(t)\,dt + X(t)\,dW(t), \qquad X(0) = 2, \tag{26}$$

with exact solution

$$X(t) = 2\exp(-5.5t + W(t));$$

see Figure 2. Since the behaviour of equation (26) is dominated by its drift part, the Adams–Moulton-2 scheme turns out to be superior for moderate step sizes.

https://doi.org/10.1112/S1461157000001388 Published online by Cambridge University Press

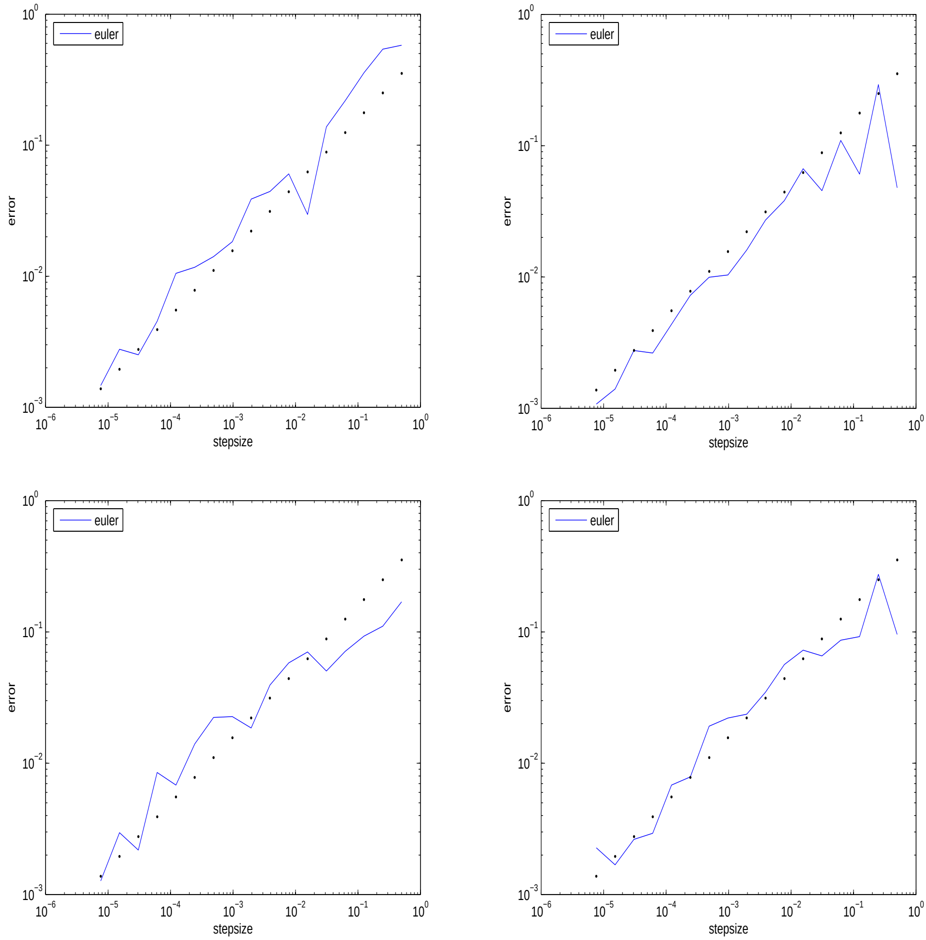


Figure 3: Example (27): pathwise maximum error vs. step size for four sample paths.

Compare, for example, [3]. The convergence rate of the Milstein scheme is in very good accordance with its predicted rate.

The third example is the linear one-dimensional stochastic delay equation

$$dX(t) = X(t-0.5)dW(t), \qquad X(t) = 1, \quad t \in [-0.5, 0]. \tag{27}$$

Here the exact solution is given by

$$X(t) = 1+W(t),\ t \in [0, 0.5], \quad X(t) = 1+W(t)+\int_{0.5}^{t} W(\tau-0.5)\,dW(\tau),\ t \in (0.5, 1],$$

which we discretize with very small step size in order to estimate the maximum error in the discretization points of the Euler–Maruyama scheme (–) for this delay equation.

Figure 3 shows the pathwise maximum error for four different sample paths. The dotted line corresponds to the convergence order 0.5. Again the pathwise convergence rate is in good accordance with the theoretically predicted rate.

https://doi.org/10.1112/S1461157000001388 Published online by Cambridge University Press

*Acknowledgements.* The authors would like to thank the referees for their helpful comments and remarks.

*References*

**1.** C. T. H. Baker and E. Buckwar, 'Numerical analysis of explicit one-step methods for stochastic delay differential equations', *LMS J. Comput. Math.* 3 (2000) 315–335, http://www.lms.ac.uk/jcm/3/lms2000-002. 245

**2.** L. Brugnano, K. Burrage and P. M. Burrage, 'Adams-type methods for the numerical solution of stochastic ordinary differential equations', *BIT* 40 (2000) 451–470. 236, 240

**3.** E. Buckwar and R. Winkler, 'Multistep methods for SDEs and their application to problems with small noise', *SIAM J. Numer. Anal.* 44 (2006) 779–803. 236, 240, 245, 251

**4.** L. Grüne and P. E. Kloeden, 'Pathwise approximation of random ordinary differential equations', *BIT* 41 (2001) 711–721. 236

**5.** I. Gyöngy, 'A note on Euler's approximations', *Potential Anal.* 8 (1998) 205–216. 236, 238, 239

**6.** E. Hairer, S. P. Nørsett and G. Wanner, *Solving ordinary differential equations. I: non stiff problems*, 2nd edn (Springer, Berlin, 1993). 240

**7.** Y. Hu, S.-E. A. Mohammed and F. Yan, 'Discrete-time approximations of stochastic delay equations: the Milstein scheme', *Ann. Probab.* 32 (2004) 265–314. 248

**8.** J. Jacod and P. Protter, 'Asymptotic error distributions for the Euler method for stochastic differential equations', *Ann. Prob.* 26 (1998) 267–307. 239

**9.** P. E. Kloeden and E. Platen, *Numerical solution of stochastic differential equations*, 3rd edn (Springer, Berlin, 1999). 235, 236, 238

**10.** X. Mao, *Stochastic differential equations and their applications* (Horwood Publishing Ltd, Chichester, 1997). 246

**11.** X. Mao, 'Numerical solutions of stochastic functional differential equations', *LMS J. Comput. Math.* 6 (2003) 141–161, http://www.lms.ac.uk/jcm/6/lms2002-027. 246

**12.** G. N. Milstein, *Numerical integration of stochastic differential equations* (Kluwer, Doordrecht, 1995). 235, 245, 248

**13.** A. Neuenkirch and I. Nourdin, 'Exact rate of convergence of some approximation schemes associated to SDEs driven by a fBm', *J. Theor. Probab.* (2007), to appear. 237

**14.** D. Nualart, *The Malliavin calculus and related topics*, 2nd edn (Springer, Berlin, 1995). 237

**15.** W. Römisch and R. Winkler, 'Stepsize control for mean-square numerical methods for stochastic differential equations with small noise', *SIAM J. Sci. Comput.* 28 (2006) 604–625. 245

**16.** D. Talay, 'Résolution trajectorielle et analyse numérique des équations différentielles stochastiques', *Stochastics* 9 (1983) 275–306. 236, 239

**17.** L. Yan, 'Asymptotic error for the Milstein scheme for SDEs driven by continuous semimartingales', *Ann. Appl. Probab.* 15 (2005) 2706–2738. 239

P. E. Kloeden kloeden@math.uni-frankfurt.de
http://www.math.uni-frankfurt.de/~numerik/kloeden/
A. Neuenkirch neuenkirch@math.uni-frankfurt.de
http://www.math.uni-frankfurt.de/~neuenkir/

Johann Wolfgang Goethe-Universität
Institut für Mathematik
Robert-Mayer-Strasse 10
60325 Frankfurt am Main
Germany



https://doi.org/10.1112/S1461157000001388 Published online by Cambridge University Press