STEP-SIZE CONTROL FOR THE UNIFORM APPROXIMATION OF SYSTEMS OF STOCHASTIC DIFFERENTIAL EQUATIONS WITH ADDITIVE NOISE

NORBERT HOFMANN[1], THOMAS MÜLLER-GRONBACH, AND KLAUS RITTER

ABSTRACT. We analyze the pathwise approximation for systems of stochastic differential equations. The pathwise distance between the solution and its approximation is measured globally on the unit interval in the L_∞-norm, and we study the expectation of this distance. For systems with additive noise we obtain sharp lower and upper bounds for the minimal error in the class of arbitrary methods which use discrete observations of a Brownian path. The optimal order is achieved by an Euler scheme with adaptive step-size control. We illustrate the superiority of the adaptive method compared to equidistant discretization by a simulation experiment.

1. INTRODUCTION

We study pathwise approximation for systems of stochastic differential equations. We consider numerical schemes that use the values of the driving Brownian motion at finitely many points and generate an approximation to the whole trajectory of the solution. Up to now results are available mainly for equidistant discretizations or in terms of the maximal step-size. Here we determine to what extent adaptive discretizations are superior to equidistant ones.

Let X and $\overline{X}$ denote the exact solution and an approximate solution, respectively, of a given system of equations on $T = [0,1]$. For a corresponding pair of trajectories in $\mathbb{R}^d$ we define the error as

$$\|X - \overline{X}\|_\infty = \max_{t \in T} \max_{1 \le i \le d} |X_i(t) - \overline{X}_i(t)|,$$

i.e., as the maximal distance of the components of X and $\overline{X}$ on the interval T. Averaging over all trajectories we end up with the definition

$$e(\overline{X}) = \Big(E(\|X - \overline{X}\|_\infty^q)\Big)^{1/q}, \qquad 1 \le q < \infty,$$

of the error of $\overline{X}$.

Date: March 11, 1999.

1991 *Mathematics Subject Classification.* Primary: 65U05; Secondary: 60H10.

Key words and phrases. systems of stochastic differential equations, pathwise approximation, adaption, step-size control, asymptotic optimality.

[1] Supported by the DFG:GR 876/9-2.

For $n \in \mathbb{N}$ we use $\widehat{X}^n$ to denote the Euler scheme with constant step-size $1/n$. From Faure (1992) we know that

$$e(\widehat{X}^n) \leq c \cdot (\ln n/n)^{1/2}. \tag{1}$$

See also Bouleau and Lépingle (1994, Remark 5.B.1.5).

We address the following questions: Is the upper bound in (1) sharp, and if so, what can be said about the asymptotic constants? Can we decrease the error by switching to a different discretization or approximation scheme? What is the minimal error that can be achieved by any method that uses n observations of the driving Brownian motion?

We provide answers in an asymptotic sense for systems

$$dX(t) = a(t, X(t))\, dt + \sigma(t)\, dW(t) \tag{2}$$

with additive noise. Here W denotes an m-dimensional Brownian motion, and $a : T \times \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma : T \to \mathbb{R}^{d\times m}$ satisfy certain smoothness conditions. For the best discretization the step-size $\tau_{l+1} - \tau_l$ is proportional to $1/\sigma_*^2(\tau_l)$, where

$$\sigma_*(t) = \max_{1\leq i\leq d} \left(\sum_{j=1}^m \sigma_{i,j}^2(t) \right)^{1/2}.$$

It suffices to use the Euler scheme with this discretization to get an asymptotically optimal approximation scheme. The corresponding error is of order

$$1/\sqrt{2} \cdot \|\sigma_*\|_2 \cdot (\ln n/n)^{1/2}.$$

No method that uses n observations of W can lead to better results. The error of the Euler method with step-size $1/n$ differs from the above minimal error by the factor $\|\sigma_*\|_\infty / \|\sigma_*\|_2$.

The above discretization seems reasonable for general systems

$$dX(t) = a(t, X(t))\, dt + \sigma(t, X(t))\, dW(t), \tag{3}$$

where σ_* and the step-size now depend on the state variable, too. However, we do not have analytical results in this case.

Different results hold if the error is defined by the L_2-distance of X and $\overline{X}$. In this case the best discretization takes steps proportional to $1/\sigma^*(\tau_l)$ with

$$\sigma^*(t) = \left(\sum_{i=1}^d \sum_{j=1}^m \sigma_{i,j}^2(t) \right)^{1/2}.$$

See Hofmann, Müller-Gronbach, and Ritter (1999) for the case $d = m = 1$. Furthermore the minimal errors tend to zero like $n^{-1/2}$, and the diffusion coefficients enters the asymptotic constant of the minimal errors via $\|\sigma^*\|_1$. Hence there is a strong correspondence between the underlying norm and the best discretization. See Remark 1 for comments concerning L_p-approximation.

We add that discrete norms, i.e., the distance between X and $\overline{X}$ at a finite number of points, are frequently used in the literature. See Hofmann *et al.* (1999) for a discussion. Only a few papers deal with adaptive step-size control. Our approach is related to Cambanis and Hu (1996), where the mean square error at the point $t = 1$ is studied. Different approaches towards adaptive step-size control are taken in Newton (1990), Gaines and Lyons (1997), and Mauthner (1998).

2. Main Results

The system (3) consists of the equations

$$dX_i(t) = a_i(t, X(t))\,dt + \sum_{j=1}^{m} \sigma_{i,j}(t, X(t))\,dW_j(t) \tag{4}$$

for $t \in T$ and $i = 1, \dots, d$, and it is driven by an m-dimensional Brownian motion $W = (W_1, \dots, W_m)$ on the unit interval T. The Euler scheme for (4) is based on a discretization

$$0 = \tau_0 < \cdots < \tau_n = 1 \tag{5}$$

of the unit interval. A discrete approximation for the solution X is defined by

$$\widehat{X}_i(\tau_0) = X_i(0)$$

for $i = 1, \dots, d$ with $X(0)$ denoting the vector of initial values, and

$$\widehat{X}_i(\tau_{l+1}) = \widehat{X}_i(\tau_l) + a_i(\tau_l, \widehat{X}(\tau_l)) \cdot (\tau_{l+1} - \tau_l) + \sum_{j=1}^{m} \sigma_{i,j}(\tau_l, \widehat{X}(\tau_l)) \cdot (W_j(\tau_{l+1}) - W_j(\tau_l))$$

for $l = 0, \dots, n-1$ and $i = 1, \dots, d$. A global approximation $\widehat{X}$ for X on T is obtained by piecewise linear interpolation of the data $(\tau_l, \widehat{X}(\tau_l))$.

We suggest to use the following adaptive step-size control for the Euler scheme in order to obtain small errors in L_∞-norm. Let

$$\sigma_*(t, x) = \max_{1 \le i \le d} \left(\sum_{j=1}^{m} \sigma_{i,j}^2(t, x) \right)^{1/2}.$$

Choose a basic step-size

$$h \in \,]0, 1[\,,$$

put $\tau_0 = 0$, and define

$$\tau_{l+1} = \tau_l + h / \sigma_*^2(\tau_l, \widehat{X}(\tau_l)) \tag{6}$$

as long as the right-hand side does not exceed one. Otherwise put $\tau_{l+1} = 1$. We use $\widehat{X}^h$ to denote the corresponding Euler scheme in the sequel.

By $|\cdot|_\infty$ we denote the ℓ_∞-norm on the space $\mathbb{R}^d$ and by $\|\cdot\|_p$ the L_p-norm of real-valued functions on T. We define $\|f\|_p = \max_{1\le i\le d} \|f_i\|_p$ for $\mathbb{R}^d$-valued functions f on T.

In our analysis we consider systems of equations

$$dX_i(t) = a_i(t, X(t))\, dt + \sum_{j=1}^{m} \sigma_{i,j}(t)\, dW_j(t) \tag{7}$$

with additive noise, see (2). We assume that the drift vector $a : T \times \mathbb{R}^d \to \mathbb{R}^d$ and the diffusion matrix $\sigma : T \to \mathbb{R}^{d\times m}$ have the following properties:

(A) There exist constants $K_1, K_2, K_3 > 0$ such that

$$|a(t,x) - a(t,y)|_\infty \le K_1 \cdot |x-y|_\infty\,, \qquad |a(t,0)|_\infty \le K_2$$

and

$$|a(t,x) - a(s,x)|_\infty \le K_3 \cdot (1 + |x|_\infty) \cdot |t-s|^{1/2}$$

for all $s, t \in T$ and $x, y \in \mathbb{R}$.

(B) σ is continuously differentiable with

$$\sigma_*(t) > 0$$

for all $t \in T$, and there exist constants $K_4, K_5 > 0$ such that

$$\|\sigma_{i,j}\|_\infty \le K_4\,, \qquad \|\sigma'_{i,j}\|_\infty \le K_5$$

for all $i = 1, \dots, d$ and $j = 1, \dots, m$.

Furthermore we assume that

(C) $X(0)$ is independent of W and

$$E\left(|X(0)|_\infty^q\right) \le K_6$$

for some constant $K_6 > 0$.

Conditions (A) and (C) are standard assumptions for analyzing stochastic differential systems, while (B) is slightly stronger than the standard assumption for systems with additive noise. See, e.g., Bouleau and Lépingle (1994, Chapter 5). In particular the conditions assure that

$$E\left(\|X\|_\infty^q\right) < \infty. \tag{8}$$

At first we compare equidistant discretization and our adaptive discretization (6) for the Euler scheme. We use $n(h, \sigma_*)$ to denote the total number of discretization points in (6).

Theorem 1. *Assume that (A)–(C) hold for the system of equations* (7)*. Then*

$$\lim_{h\to 0}\, (n(h,\sigma_*)/\ln n(h,\sigma_*))^{1/2} \cdot e(\widehat{X}^h) = 1/\sqrt{2} \cdot \|\sigma_*\|_2$$

for the Euler method based on the discretization (6). *The Euler approximation* $\widehat{X}^n$ *with constant step-size*

$$\tau_{k+1} - \tau_k = 1/n$$

yields

$$\lim_{n\to\infty} (n/\ln n)^{1/2} \cdot e(\widehat{X}^n) = 1/\sqrt{2} \cdot \|\sigma_*\|_\infty.$$

Taking (6) instead of an equidistant discretization therefore reduces the error roughly by the factor

$$\|\sigma_*\|_2/\|\sigma_*\|_\infty$$

for the same number of observations of W. The overall computational cost increases by just a few arithmetical operations per step, which is irrelevant in practice.

Choosing steps of length $h/\sigma_*^2(\tau_l)$ leads to the good asymptotic performance of $\widehat{X}^h$. However, we already want to control the error for larger values of h or small values of $\sigma_*(\tau_l)$. To this end we modify (6) according to

$$\tau_{l+1} = \tau_l + \min\Big(h \cdot \ln h^{-1}, h/\sigma_*^2(\tau_l, \widehat{X}^h(\tau_l))\Big), \tag{9}$$

See Remark 2 for details. A step of size $h \cdot \ln h^{-1}$ is only taken, if $h > \exp(-\sigma_*^{-2}(\tau_l)))$. Obviously the additional term is irrelevant asymptotically.

Now we show that no other discretization or other use of the data $W(\tau_1), \dots, W(\tau_n)$ and $X(0)$ can lead to better asymptotic results. An arbitrary method $\overline{X}^n$ with components $\overline{X}_i^n$ that is based on a realization of the vector of initial values $X(0)$ and on n observations of a trajectory of the m-dimensional Brownian motion W is defined by measurable mappings

$$\psi_k : \mathbb{R}^{(k-1)\cdot m+d} \to T$$

for $k = 1, \dots, n$ and

$$\phi_i : \mathbb{R}^{n\cdot m+d} \to L_\infty(T).$$

for $i = 1, \dots, d$. The mapping ψ_k determines the observation point in step k in terms of the previous evaluations. A pathwise approximation is computed according to

$$\overline{X}_i^n = \phi_i(X(0), Y_1, \dots, Y_n),$$

for $i = 1, \dots, d$, where $Y_1 = W(\psi_1(X(0)))$, and

$$Y_k = W(\psi_k(X(0), Y_1, \dots, Y_{k-1}))$$

is the observation in step $k \geq 2$. Every such method is called an n-point method in the sequel.

The quantity

$$e(n) = \inf_{\overline{X}^n} e(\overline{X}^n)$$

is the minimal error that can be obtained by n-point methods for the system of equations (7).

Theorem 2. *Assume that (A)–(C) hold for the system of equations* (7)*. The minimal errors satisfy*

$$\lim_{n\to\infty} (n/\ln n)^{1/2} \cdot e(n) = 1/\sqrt{2} \cdot \|\sigma_*\|_2.$$

Due to Theorems 1 and 2 the Euler method $\widehat{X}^h$ is asymptotically optimal for every system of equations with additive noise.

Remark 1. Our approach also covers the asymptotic analysis of errors

$$e(\overline{X}) = \left(E(\|X - \overline{X}\|_p^p)\right)^{1/p}, \qquad 1 \le p < \infty.$$

For scalar equations we propose an adaptive discretization that takes steps proportionally to $1/\sigma^\lambda(\tau_l, \widehat{X}(\tau_l))$ with $\lambda = 1/(2^{-1}+p^{-1})$. For equations with additive noise the error of the corresponding Euler scheme $\widehat{X}$ yields the asymptotic constant $(\eta_p\cdot\mu_p)^{1/p}\cdot\|\sigma\|_\lambda$ which is the best possible one. Here η_p denotes the pth moment of the absolute value of a standard Gaussian random variable and $\mu_p = \Gamma^2(p/2+1)/\Gamma(p+2)$. For the Euler approximation with constant step-size $1/n$ we get the asymptotic constant $(\eta_p\cdot\mu_p)^{1/p}\cdot\|\sigma\|_p$.

3. Simulation Experiment

We consider the scalar equation

$$dX(t) = -2\,dt + \sigma(t)\,dW(t), \qquad X(0) = 1, \tag{10}$$

driven by a one-dimensional Brownian motion W. We choose the diffusion coefficient σ according to

$$\sigma(t) = 20 \cdot \exp(-(10t-1)^2) - t.$$

The solution of (10) is not known explicitly. Instead of X we therefore use an M-point Euler approximation $\widehat{X}_M$ with M sufficiently large. We simulate $K = 10000$ trajectories of W and use

$$\varepsilon(\overline{X}) = \left(\frac{1}{K}\cdot\sum_{k=1}^{K} \|\widehat{X}_M(\omega_k) - \overline{X}(\omega_k)\|_\infty^q\right)^{1/q}$$

as an estimator for $e(\overline{X})$. The quantities $\varepsilon(\widehat{X}^h)$ and $\varepsilon(\widehat{X}^n)$ are easy to compute. We express the efficiency of the method $\widehat{X}^h$ by the ratio

$$\mathrm{eff}(h) = \varepsilon(\widehat{X}^h)/\varepsilon(\widehat{X}^n).$$

Theorem 1 tells us that $\mathrm{eff}(h)$ is approximately given by

$$\|\sigma\|_2/\|\sigma\|_\infty = 0.35\ldots$$

for sufficiently large K and small h. We also show the mean ratio $\mathrm{ex}(h)$ between steps of length $h\cdot\ln h^{-1}$ (called exceptions) and the total number of knots in the table below.

h	$\varepsilon(\widehat{X}^h)$	$n(h,\sigma)$	$\varepsilon(\widehat{X}^n)$	eff(h)	ex(h)
$5\cdot 10^{-1}$	1.246034	100	2.489485	0.50	0.0200
$1\cdot 10^{-1}$	0.651305	489	1.262255	0.52	0.0041
$5\cdot 10^{-2}$	0.485591	975	0.954914	0.51	0.0021
$1\cdot 10^{-2}$	0.242423	4864	0.487348	0.50	0.0008
$5\cdot 10^{-3}$	0.170084	9725	0.360682	0.47	0.0006
$1\cdot 10^{-3}$	0.075044	48608	0.179113	0.42	0.0003
$5\cdot 10^{-4}$	0.054014	97212	0.131334	0.41	0.0002

Simulation for (10) and $q = 2$ with different basic step-sizes h

The same equation is studied in Hofmann *et al.* (1999), where pathwise approximation with respect to the global error in the L_2-norm is analyzed. A comparison of the results shows a much slower convergence of the efficiency in the L_∞-case. The proofs and a closer look to the data reveal that this is caused by error $\varepsilon(\widehat{X}^n)$ of the Euler method with equidistant discretization. In contrast to the L_2-case the asymptotical behavior of $e(\widehat{X}^n)$ is now determined by the performance of $\widehat{X}^n$ around the points where σ_* takes its maximum. Therefore a large number of discretization points is needed around these sites, and the total number of points has to be very large.

4. Proofs

The proofs are based on the decomposition

$$X(t) = X(0) + V(t) + Y(t) - Z(t)$$

with components

$$V_i(t) = \sum_{j=1}^{m} \sigma_{i,j}(t) \cdot W_j(t),$$

$$Y_i(t) = \int_0^t a_i(s, X(s))\, ds,$$

and

$$Z_i(t) = \int_0^t \sum_{j=1}^{m} \sigma'_{i,j}(s) \cdot W_j(s)\, ds$$

for $i = 1, \dots, d$. Fix a discretization (5) and let

$$\Delta_l = \tau_l - \tau_{l-1}$$

and

$$\Delta_{\max} = \max_{1\le l\le n} \Delta_l.$$

The corresponding Euler approximation $\widehat{X}$ of X may be written as

$$\widehat{X}(\tau_l) = X(0) + \widehat{V}(\tau_l) + \widehat{Y}(\tau_l) - \widehat{Z}(\tau_l) \tag{11}$$

with components

$$\widehat{V}_i(\tau_l) = \sum_{j=1}^{m} \sigma_{i,j}(\tau_l) \cdot W_j(\tau_l),$$

$$\widehat{Y}_i(\tau_l) = \sum_{k=1}^{l} a_i(\tau_{k-1}, \widehat{X}(\tau_{k-1})) \cdot \Delta_k,$$

and

$$\widehat{Z}_i(\tau_l) = \sum_{k=1}^{l} \Big(\sum_{j=1}^{m} (\sigma_{i,j}(\tau_k) - \sigma_{i,j}(\tau_{k-1})) \cdot W_j(\tau_k) \Big)$$

for $i = 1, \dots, d$. Piecewise linear interpolation yields $\mathbb{R}^d$-valued processes $\widehat{V}$, $\widehat{Y}$, and $\widehat{Z}$ over T.

4.1. **Preliminary estimates.** In the sequel $\widetilde{X}$ denotes the piecewise linear interpolation of an arbitrary $\mathbb{R}^d$-valued process X on T. Thus

$$\widetilde{X}(t) = \Big((t - \tau_{l-1}) \cdot X(\tau_l) + (\tau_l - t) \cdot X(\tau_{l-1}) \Big) / \Delta_l$$

for $t \in [\tau_{l-1}, \tau_l]$. Moreover, we use c to denote unspecified positive constants which only depend on q, d, m, and on the constants from conditions (A)–(C).

At first we compare the processes Y and Z with their counterparts $\widehat{Y}$ and $\widehat{Z}$ from the corresponding Euler approximation.

Lemma 1. *We have*

$$\Big(E(\|Z - \widehat{Z}\|_\infty^q) \Big)^{1/q} \le c \cdot \Delta_{\max}$$

and

$$\Big(E(\|Y - \widehat{Y}\|_\infty^q) \Big)^{1/q} \le c \cdot \Delta_{\max}^{1/2}.$$

Proof. Let $t \in [\tau_{l-1}, \tau_l]$. Observing property (B) we get

$$\begin{aligned} |Z_i(t) - \widetilde{Z}_i(t)| &\le \left((\tau_l - t) \cdot |Z_i(t) - Z_i(\tau_{l-1})| + (t - \tau_{l-1}) \cdot |Z_i(t) - Z_i(\tau_l)| \right) / \Delta_l \\ &\le \int_{\tau_{l-1}}^{\tau_l} \sum_{j=1}^{m} |\sigma'_{i,j}(s) \cdot W_j(s)| \, ds \\ &\le K_5 \cdot \Delta_l \cdot \sum_{j=1}^{m} \|W_j\|_\infty. \end{aligned}$$

Therefore

$$\Big(E(\|Z_i - \widetilde{Z}_i\|_\infty^q) \Big)^{1/q} \le c \cdot \Delta_{\max}. \tag{12}$$

Clearly

$$\|\widetilde{Z}_i - \widehat{Z}_i\|_\infty = \max_{1 \le l \le n} \left| \sum_{k=1}^{l} \int_{\tau_{k-1}}^{\tau_k} \sum_{j=1}^{m} \sigma'_{i,j}(s) \cdot (W_j(s) - W_j(\tau_k)) \, ds \right|.$$

We use Lévy's inequality and (B) to obtain

$$
\begin{aligned}
\left(E(\|\widetilde{Z}_i - \widehat{Z}_i\|_\infty^q)\right)^{1/q} &\le 2^{1/q} \cdot \left(E\left(\left|\sum_{k=1}^n \int_{\tau_{k-1}}^{\tau_k} \sum_{j=1}^m \sigma'_{i,j}(s)\cdot(W_j(s)-W_j(\tau_k))\,ds\right|^q\right)\right)^{1/q} \\
&= c\cdot\left(E\left(\left|\sum_{k=1}^n \int_{\tau_{k-1}}^{\tau_k} \sum_{j=1}^m \sigma'_{i,j}(s)\cdot(W_j(s)-W_j(\tau_k))\,ds\right|^2\right)\right)^{1/2} \\
&\le c\cdot\left(\sum_{k=1}^n \int_{\tau_{k-1}}^{\tau_k}\int_{\tau_{k-1}}^{\tau_k} (\tau_k - \max(s,t))\,ds\,dt\right)^{1/2} \\
&\le c\cdot\left(\sum_{k=1}^n \Delta_k^3\right)^{1/2} \\
&\le c\cdot\Delta_{\max}. \qquad (13)
\end{aligned}
$$

Now (12), (13), and

$$
\left(E(\|Z-\widehat{Z}\|_\infty^q)\right)^{1/q} \le \sum_{i=1}^d \left(E(\|Z_i-\widehat{Z}_i\|_\infty^q)\right)^{1/q}
$$

yield the first estimate.

Due to condition (A) we have

$$
\|Y_i - \widetilde{Y}_i\|_\infty \le \max_{1\le l\le n}\int_{\tau_{l-1}}^{\tau_l} |a_i(s,X(s))|\,ds \le c\cdot\Delta_{\max}\cdot(1+\|X\|_\infty).
$$

Therefore

$$
\left(E(\|Y_i-\widetilde{Y}_i\|_\infty^q)\right)^{1/q} \le c\cdot\Delta_{\max} \qquad (14)
$$

by (8). Furthermore,

$$
\|\widetilde{Y}_i - \widehat{Y}_i\|_\infty = \max_{1\le l\le n}|Y_i(\tau_l)-\widehat{Y}_i(\tau_l)| \le \max_{1\le l\le n}|X_i(\tau_l)-\widehat{X}_i(\tau_l)| + \max_{1\le l\le n}|Z_i(\tau_l)-\widehat{Z}_i(\tau_l)|,
$$

together with (31) and the first estimate, yields

$$
\left(E(\|\widetilde{Y}_i-\widehat{Y}_i\|_\infty^q)\right)^{1/q} \le c\cdot\Delta_{\max}^{1/2} + c\cdot\Delta_{\max} \le c\cdot\Delta_{\max}^{1/2}. \qquad (15)
$$

Combine (14), (15) and

$$
\left(E(\|Y-\widehat{Y}\|_\infty^q)\right)^{1/q} \le \sum_{i=1}^d \left(E(\|Y_i-\widehat{Y}_i\|_\infty^q)\right)^{1/q}
$$

to complete the proof. □

We relate the error of the Euler method to a weighted L_∞-error of piecewise linear interpolation of W.

Lemma 2.

$$\left| e(\widehat{X}) - \left(E(\|\sigma \cdot (W - \widetilde{W})\|_\infty^q)\right)^{1/q} \right| \le c \cdot \Delta_{\max}^{1/2}.$$

Proof. Use (B) and Lemma 1. □

4.2. **Proof of the lower bounds in Theorem 1.** In the sequel we use $B_1, B_2, \dots$ to denote independent Brownian bridges on T with $B_l(0) = B_l(1) = 0$, and we put

$$B_l^* = \|B_l\|_\infty.$$

Moreover, let

$$\sigma_i(t) = \left(\sum_{j=1}^m \sigma_{i,j}(t)^2 \right)^{1/2},$$

such that

$$\sigma_*(t) = \max_{1 \le i \le d} \sigma_i(t).$$

Take $i \in \{1, \dots, d\}$ with

$$\|\sigma_i\|_\infty = \|\sigma_*\|_\infty.$$

Clearly

$$\left(E(\|\sigma \cdot (W - \widetilde{W})\|_\infty^q)\right)^{1/q} \ge \left(E(\|\sigma_i \cdot (W_1 - \widetilde{W}_1)\|_\infty^q)\right)^{1/q}. \tag{16}$$

Moreover,

$$\begin{aligned} &\left(E(\|\sigma_i \cdot (W_1 - \widetilde{W}_1)\|_\infty^q)\right)^{1/q} \\ &\quad = \left(E\left(\max_{1 \le l \le n} \left(\Delta_l^{1/2} \cdot \max_{\tau_{l-1} \le t \le \tau_l} |\sigma_i(t) \cdot B_l((t - \tau_{l-1})/\Delta_l)|\right)^q\right)\right)^{1/q}, \end{aligned}$$

and therefore

$$\begin{aligned} &\left| \left(E(\|\sigma_i \cdot (W_1 - \widetilde{W}_1)\|_\infty^q)\right)^{1/q} - \left(E\left(\max_{1 \le l \le n} \left(\Delta_l^{1/2} \cdot \sigma_i(\tau_{l-1}) \cdot B_l^*\right)^q\right)\right)^{1/q} \right| \\ &\quad \le c \cdot \Delta_{\max} \end{aligned} \tag{17}$$

by property (B). Due to Lemma 2 as well as (16) and (17) we have

$$e(\widehat{X}) \ge \left(E\left(\max_{1 \le l \le n} \left(\Delta_l^{1/2} \cdot \sigma_i(\tau_{l-1}) \cdot B_l^*\right)^q\right)\right)^{1/q} - c \cdot \Delta_{\max}^{1/2}. \tag{18}$$

Now we consider the particular discretization (6). Clearly $\sigma_*(\tau_{l-1}) \cdot \Delta_l^{1/2} = h^{1/2}$ and

$$\Delta_{\max} \le \|\sigma_*^{-2}\|_\infty \cdot h. \tag{19}$$

¿From (18) we get

$$e(\widehat{X}^h) \ge h^{1/2} \cdot \left(E\left(\max_{1 \le l \le n} \left(\sigma_i(\tau_{l-1})/\sigma_*(\tau_{l-1}) \cdot B_l^*\right)^q\right)\right)^{1/q} - c \cdot \|\sigma_*^{-2}\|_\infty \cdot h^{1/2}.$$

Moreover,

$$\lim_{h\to 0} n(h,\sigma_*)\cdot h = \|\sigma_*\|_2^2. \tag{20}$$

Corollary 1 yields

$$\begin{aligned}
&\liminf_{h\to 0} (n(h,\sigma_*)/\ln n(h,\sigma_*))^{1/2}\cdot e(\widehat{X}^h)\\
&\qquad \geq \liminf_{h\to 0} (n(h,\sigma_*)\cdot h)^{1/2}\cdot(\ln n(h,\sigma_*))^{-1/2}\\
&\qquad\qquad \cdot\left(E\left(\max_{1\le l\le n(h,\sigma_*)}\Big((\sigma_i(\tau_{l-1})/\sigma_*(\tau_{l-1}))\cdot B_l^*\Big)^q\right)\right)^{1/q}\\
&\qquad = \|\sigma_*\|_2\cdot 2^{-1/2}.
\end{aligned}$$

Next we consider the constant step-size

$$\Delta_l = 1/n.$$

We apply Corollary 1 and (18) to obtain

$$\begin{aligned}
&\liminf_{n\to\infty} (n/\ln n)^{1/2}\cdot e(\widehat{X}^n)\\
&\qquad \geq \liminf_{n\to\infty} (\ln n)^{-1/2}\cdot\left(E\left(\max_{1\le l\le n}\Big(\sigma_i((l-1)/n)\cdot B_l^*\Big)^q\right)\right)^{1/q}\\
&\qquad = \|\sigma_*\|_\infty\cdot 2^{-1/2}.
\end{aligned}$$

4.3. **Proof of the upper bounds in Theorem 1.** Due to Lemma 2 as well as (19) and (20) it suffices to study $\|\sigma\cdot(W-\widetilde{W})\|_\infty$. Put $\sigma_{d+1,1}=\sigma_*$ and $\sigma_{d+1,j}=0$ for $j=2,\dots,m$. Note that

$$\begin{aligned}
&(n/\ln n)^{q/2}\cdot E\Big(\|\sigma\cdot(W-\widetilde{W})\|_\infty^q\Big)\\
&\quad\leq \sum_{i=1}^{d+1}\int_z^\infty P\Big(\|(n/\ln n)^{1/2}\cdot\sum_{j=1}^m \sigma_{i,j}\cdot(W_j-\widetilde{W}_j)\|_\infty^q > u\Big)\,du\\
&\qquad + \int_0^z P\Big(\max_{1\le i\le d+1}\|(n/\ln n)^{1/2}\cdot\sum_{j=1}^m \sigma_{i,j}\cdot(W_j-\widetilde{W}_j)\|_\infty^q > u\Big)\,du
\end{aligned} \tag{21}$$

for every $z > 0$. Moreover,

$$P\Big(\|\sum_{j=1}^{m}\sigma_{i,j}\cdot(W_j-\widetilde{W}_j)\|_\infty > v\Big) = P\Big(\|\sigma_i\cdot(W_1-\widetilde{W}_1)\|_\infty > v\Big)$$
$$\le P\Big(\|\sigma_*\cdot(W_1-\widetilde{W}_1)\|_\infty > v\Big)$$
$$\le P\Big(\max_{1\le l\le n}\max_{\tau_{l-1}\le t\le\tau_l}\Big(\sigma_*(t)\cdot\Delta_l^{1/2}\cdot B_l^*\Big) > v\Big) \tag{22}$$

as well as

$$P\Big(\max_{1\le i\le d+1}\|\sum_{j=1}^{m}\sigma_{i,j}\cdot(W_j-\widetilde{W}_j)\|_\infty > v\Big) \ge P\Big(\|\sum_{j=1}^{m}\sigma_{d+1,j}\cdot(W_j-\widetilde{W}_j)\|_\infty > v\Big)$$
$$= P\Big(\|\sigma_*\cdot(W_1-\widetilde{W}_1)\|_\infty > v\Big)$$
$$\ge P\Big(\max_{1\le l\le n}\min_{\tau_{l-1}\le t\le\tau_l}\Big(\sigma_*(t)\cdot\Delta_l^{1/2}\cdot B_l^*\Big) > v\Big) \tag{23}$$

for every $v > 0$.

First we consider the discretization (6). Put $n = n(h, \sigma_*)$ to obtain

$$P\Big(\|(n/\ln n)^{1/2}\cdot\sum_{j=1}^{m}\sigma_{i,j}\cdot(W_j-\widetilde{W}_j)\|_\infty^q > u\Big)$$
$$\le P\Big(\max_{1\le l\le n}\Big((n\cdot h)^{1/2}\cdot\frac{\max_{\tau_{l-1}\le t\le\tau_l}\sigma_*(t)}{\sigma_*(\tau_{l-1})}\cdot(\ln n)^{-1/2}\cdot B_l^*\Big)^q > u\Big)$$

from (22). Analogously, (23) yields

$$P\Big(\max_{1\le i\le d+1}\|(n/\ln n)^{1/2}\cdot\sum_{j=1}^{m}\sigma_{i,j}\cdot(W_j-\widetilde{W}_j)\|_\infty^q > u\Big)$$
$$\ge P\Big(\max_{1\le l\le n}\Big((n\cdot h)^{1/2}\cdot\frac{\min_{\tau_{l-1}\le t\le\tau_l}\sigma_*(t)}{\sigma_*(\tau_{l-1})}\cdot(\ln n)^{-1/2}\cdot B_l^*\Big)^q > u\Big).$$

Observing (20) we take

$$z = \Big(\|\sigma_*\|_2/2^{1/2}\Big)^q$$

in (21) and apply Lemma 3 with $\alpha = \|\sigma_*\|_2$ to obtain

$$\limsup_{h\to 0}\,(n/\ln n)^{q/2}\cdot E\left(\|\sigma\cdot(W-\widetilde{W})\|_\infty^q\right) \le \Big(\|\sigma_*\|_2/2^{1/2}\Big)^q.$$

For an equidistant discretization we get

$$P\Big(\|(n/\ln n)^{1/2}\cdot\sum_{j=1}^{m}\sigma_{i,j}\cdot(W_j-\widetilde{W}_j)\|_\infty^q > u\Big)$$
$$\le P\Big(\max_{1\le l\le n}\Big(\max_{\tau_{l-1}\le t\le\tau_l}\sigma_*(t)\cdot(\ln n)^{-1/2}\cdot B_l^*\Big)^q > u\Big)$$

from (22) and

$$
\begin{aligned}
P\Big(\max_{1\le i\le d+1} \|(n/\ln n)^{1/2} \cdot \sum_{j=1}^{m} \sigma_{i,j} \cdot (W_j - \widetilde{W}_j)\|_\infty^q > u\Big) \\
\ge P\Big(\max_{1\le l\le n}\Big(\min_{\tau_{l-1}\le t\le \tau_l} \sigma_*(t) \cdot (\ln n)^{-1/2} \cdot B_l^*\Big)^q > u\Big)
\end{aligned}
$$

from (23). We take

$$
z = \left(\|\sigma_*\|_\infty / 2^{1/2}\right)^q
$$

in (21) and apply Lemma 3 with $\alpha = \|\sigma_*\|_\infty$ to obtain

$$
\limsup_{n\to\infty} \, (n/\ln n)^{q/2} \cdot E\left(\|\sigma \cdot (W - \widetilde{W})\|_\infty^q\right) \le \left(\|\sigma_*\|_\infty / 2^{1/2}\right)^q.
$$

This completes the proof of the upper bounds in Theorem 1.

Remark 2. Consider the Euler method with an arbitrary discretization (5).
¿From Lemma 2 we get

$$
e(\widehat{X}) \le c \cdot \left(\left(E(\|\sigma_* \cdot (W_1 - \widetilde{W}_1)\|_\infty^q)\right)^{1/q} + \Delta_{\max}^{1/2}\right)
$$

as a rough estimate. Assumption (B) and Corollary 1 imply

$$
\begin{aligned}
\left(E(\|\sigma_* \cdot (W_1 - \widetilde{W}_1)\|_\infty^q)\right)^{1/q} &\le \left(E\left(\max_{1\le l\le n}\left((\sigma_*(\tau_{l-1}) + c\cdot\Delta_l)\cdot \Delta_l^{1/2} \cdot B_l^*\right)^q\right)\right)^{1/q} \\
&\le c \cdot \max_{1\le l\le n} (\sigma_*(\tau_{l-1}) + c\cdot \Delta_l) \cdot \Delta_l^{1/2} \cdot (\ln n)^{1/2}.
\end{aligned}
$$

If $\Delta_l = h/\sigma_*^2(\tau_{l-1})$ then $n(h,\sigma_*) \le c/h$ due to (B), and we obtain

$$
e(\widehat{X}) \le c \cdot \left((\ln h^{-1})^{1/2} \cdot (h^{1/2} + h^{3/2} \cdot \|\sigma_*^{-3}\|_\infty) + h^{1/2} \cdot \|\sigma_*^{-1}\|_\infty\right).
$$

In fact, $e(\widehat{X})$ may get arbitrarily large for fixed h, if σ gets small at some point. Therefore we have introduced the additional term $h \cdot \ln h^{-1}$ in the definition of the step-size (9). Hereby

$$
\begin{aligned}
e(\widehat{X}) &\le c \cdot \left((\ln h^{-1})^{1/2} \cdot (h^{1/2} + (h \cdot \ln h^{-1})^{3/2}) + (h \cdot \ln h^{-1})^{1/2}\right) \\
&\le c \cdot (h \cdot \ln h^{-1})^{1/2}
\end{aligned}
$$

holds uniformly without any lower bound for σ.

Recall that all constants c do only depend on $K_1, \ldots, K_6$ and q, d and m. Therefore the bounds from Theorem 1 hold uniformly on the class of systems of equations (7) with properties (A)–(C).

4.4. **Proof of Theorem 2.** Consider an arbitrary sequence of methods $\overline{X}^n$ based on $X(0)$ and n observations of W. A formal definition of methods of this kind is given in Section 2. In view of Theorem 1 it suffices to show that

$$\liminf_{n\to\infty} (n/\ln n)^{1/2} \cdot e(\overline{X}^n) \geq 1/\sqrt{2} \cdot \|\sigma_*\|_2.$$

Let $k_n = \lceil n/(\ln n)^{1/2} \rceil$ and put

$$\tau_l^n = l/k_n \tag{24}$$

for $l = 1, \ldots, k_n$. Now define a method V^n by

$$V^n = \overline{X}^n - X(0) - \widehat{Y}^n + \widehat{Z}^n,$$

where $\widehat{Y}^n$ and $\widehat{Z}^n$ are determined by (11) with respect to the equidistant discretization (24). Note that V^n uses at most $m_n = n + k_n$ observations of W.

Observing Lemma 1 we obtain

$$e(\overline{X}^n) \geq \left(E\left(\|V - V^n\|_\infty^q\right)\right)^{1/q} - c \cdot k_n^{-1/2}$$

and therefore

$$\liminf_{n\to\infty} (n/\ln n)^{1/2} \cdot e(\overline{X}^n) \geq \liminf_{n\to\infty} (m_n/\ln m_n)^{1/2} \cdot \left(E\left(\|V - V^n\|_\infty^q\right)\right)^{1/q}. \tag{25}$$

We derive a lower bound for $E(\|V - V^n\|_\infty^q)$. Since V and $X(0)$ are independent by assumption (C), we may assume that V^n does not use the initial value $X(0)$. Since V is a Gaussian process, we may assume that V^n uses the values of W at fixed points

$$0 < \xi_1 < \cdots < \xi_{m_n} = 1. \tag{26}$$

Then we have

$$\begin{aligned} E(\|V - V^n\|_\infty^q) &\geq E\left(\max_{1\leq i\leq d} \|V_i - E(V_i \mid W(\xi_1), \ldots, W(\xi_{m_n}))\|_\infty^q\right) \\ &= E(\|\sigma \cdot (W - \widetilde{W})\|_\infty^q), \end{aligned} \tag{27}$$

where $\widetilde{W}$ denotes the piecewise linear interpolation of W with respect to the points (26).

For $l = 1, \ldots, n$ there exists $i_l \in \{1, \ldots, d\}$ such that

$$\sigma_{i_l}(\xi_{l-1}) = \sigma_*(\xi_{l-1}),$$

and therefore

$$\max_{\xi_{l-1}\leq t\leq \xi_l} |\sigma_{i_l}(t) - \sigma_*(t)| \leq c \cdot \max_{1\leq l\leq m_n} (\xi_l - \xi_{l-1}) \leq c \cdot k_n^{-1}.$$

by (24) and (B). Consequently,

$$\begin{aligned}\left(E(\|\sigma\cdot(W-\widetilde{W})\|_\infty^q)\right)^{1/q} &\geq \left(E\left(\max_{1\leq l\leq m_n}\max_{\xi_{l-1}\leq t\leq\xi_l}|\sigma_{i_l}(t)\cdot(W_1-\widetilde{W}_1)(t)|^q\right)\right)^{1/q}\\ &\geq \left(E(\|\sigma_*\cdot(W_1-\widetilde{W}_1)\|_\infty^q)\right)^{1/q} - c\cdot k_n^{-1}. \qquad (28)\end{aligned}$$

Moreover,

$$\lim_{n\to\infty}\ (m_n/\ln m_n)^{1/2}\cdot k_n^{-1} = 0. \qquad (29)$$

Combining (25) and (27)–(29) we conclude that

$$\begin{aligned}\liminf_{n\to\infty}\ (n/\ln n)^{1/2}\cdot e(\overline{X}^n) &\geq \liminf_{n\to\infty}\ (m_n/\ln m_n)^{1/2}\cdot\left(E(\|\sigma_*\cdot(W_1-\widetilde{W}_1)\|_\infty^q)\right)^{1/q}\\ &\geq \|\sigma_*\|_2\cdot 2^{-1/2}.\end{aligned}$$

Here the asymptotic lower bound for $E(\|\sigma_*\cdot(W_1-\widetilde{W}_1)\|_\infty^q)$ is valid for every discretization (26), see Müller-Gronbach and Ritter (1997) for a general result. This completes the proof of Theorem 2.

Appendix

Consider a stochastic differential equation (4) and a fixed discretization (5). An approximation to the solution X of (4) is given by

$$\breve{X}_i(t) = \breve{X}_i(\tau_{l-1}) + \int_{\tau_{l-1}}^t a_i(\tau_{l-1},\breve{X}(\tau_{l-1}))\,ds + \sum_{j=1}^m\int_{\tau_{l-1}}^t \sigma_{i,j}(\tau_{l-1},\breve{X}(\tau_{l-1}))\,dW_j(s), \qquad (30)$$

where $t\in[\tau_{l-1},\tau_l]$ and $i=1,\dots,d$. Bouleau and Lépingle (1994, p. 276) present the following upper bound for the error of $\breve{X}$ in the case of constant step-size. A straightforward modification of the proof of Theorem B.1.4 in Bouleau and Lépingle (1994) extends the result to the case of variable step-size.

Theorem 3. *Assume that a as well as σ satisfy condition (A) and let $X(0)$ satisfy (C). Then*

$$\left(E\left(\|X-\breve{X}\|_\infty^q\right)\right)^{1/q} \leq c\cdot\Delta_{\max}^{1/2}.$$

The constant c only depends on q, d, m, and on the constants from (A) and (C).

Note that (30) coincides with the Euler scheme at the discretization points τ_l and satisfies a stochastic differential equation with frozen coefficients on each interval $[\tau_{l-1},\tau_l]$. Complete knowledge of a trajectory of W is needed to generate the corresponding trajectory of $\breve{X}$. Therefore $\breve{X}$ is not an implementable numerical scheme for the global approximation of X. See Cambanis and Hu (1996) for a detailed asymptotic analysis of $\breve{X}$.

Since $\check{X}(\tau_l) = \widehat{X}(\tau_l)$ for $l = 1, \dots, n$, we immediately obtain

$$\left(E\left(\max_{1\le i\le d}\max_{1\le l\le n}|X_i(\tau_l) - \widehat{X}_i(\tau_l)|^q\right)\right)^{1/q} \le c\cdot\Delta_{\max}^{1/2} \tag{31}$$

from Theorem 3.

Consider a sequence of independent Brownian bridges B_l on the interval T with $B_l(0) = B_l(1) = 0$.

Lemma 3. *Let $\alpha, \alpha_{1,n}, \dots, \alpha_{n,n} > 0$ such that*

$$\lim_{n\to\infty}\max_{1\le l\le n}\alpha_{l,n} = \alpha$$

and

$$\liminf_{n\to\infty}\frac{\#\{l\,:\,\alpha_{l,n}\ge\alpha-\varepsilon\}}{n} > 0$$

for every $\varepsilon > 0$. Then

$$\lim_{n\to\infty}\int_{(\alpha^2/2)^{q/2}}^{\infty} P\Big(\max_{1\le l\le n}(\alpha_{l,n}\cdot(\ln n)^{-1/2}\cdot\|B_l\|_\infty)^q > u\Big)\,du = 0$$

and

$$\lim_{n\to\infty} P\Big(\max_{1\le l\le n}(\alpha_{l,n}\cdot(\ln n)^{-1/2}\cdot\|B_l\|_\infty)^q > u\Big) = 1$$

for $0 \le u < (\alpha^2/2)^{q/2}$.

See Ritter (1990) for a proof of Lemma 3 in the particular case $\alpha = \alpha_{l,n}$. The lemma immediately yields the following fact.

Corollary 1. *Under the assumptions of Lemma 3 we have*

$$\lim_{n\to\infty}(\ln n)^{-1/2}\cdot\Big(E\Big(\max_{1\le l\le n}(\alpha_{l,n}\cdot\|B_l\|_\infty)^q\Big)\Big)^{1/q} = \alpha\cdot 2^{-1/2}.$$

References

Bouleau, N. and Lépingle, D. (1994). *Numerical Methods for Stochastic Processes.* Wiley, New York.

Cambanis, S. and Hu, Y. (1996). Exact convergence rate of the Euler-Maruyama scheme, with application to sampling design. Stochastics Stochastics Rep. **59**, 211–240.

Faure, O. (1992). *Simulation du mouvement brownien et des diffusions.* Thèse ENPC, Paris.

Gaines, J. G. and Lyons, T. J. (1997). Variable step size control in the numerical solution of stochastic differential equations. SIAM J. Appl. Math. **57**, 1455–1484.

Hofmann, N., Müller-Gronbach, T., and Ritter, K. (1999), Optimal approximation of stochastic differential equations by adaptive step-size control, to appear in Math. Comp.

Mauthner, S. (1998). Step size control in the numerical solution of stochastic differential equations. Mathematische Preprintreihe der TU Darmstadt **1972**.

Müller-Gronbach, T. and Ritter, K. (1997). Uniform reconstruction of Gaussian processes. Stochastic Processes Appl. **69**, 55–70.

Newton, N. J. (1990). An efficient approximation for stochastic differential equations on the partition of symmetrical first passage times. Stochastics Stochastics Rep. **29**, 227–258.

Ritter, K. (1990), Approximation and optimization on the Wiener space, J. Complexity **6**, 337–364.

Mathematisches Institut, Universität Erlangen-Nürnberg, Bismarckstrasse 1 1/2, 91054 Erlangen, Germany
E-mail address: hofmann@mi.uni-erlangen.de

1. Mathematisches Institut, Freie Universität Berlin, Arminallee 2–6, 14195 Berlin, Germany
E-mail address: gronbach@math.fu-berlin.de

Fakultät für Mathematik und Informatik, Universität Passau, Innstr. 33, 94032 Passau, Germany
E-mail address: klaus.ritter@fmi.uni-passau.de