The Annals of Applied Probability
2002, Vol. 12, No. 2, 664–690

THE OPTIMAL UNIFORM APPROXIMATION OF SYSTEMS OF STOCHASTIC DIFFERENTIAL EQUATIONS

BY THOMAS MÜLLER-GRONBACH

Technische Universität Darmstadt

We analyze numerical methods for the pathwise approximation of a system of stochastic differential equations. As a measure of performance we consider the qth mean of the maximum distance between the solution and its approximation on the whole unit interval. We introduce an adaptive discretization that takes into account the local smoothness of every trajectory of the solution. The resulting adaptive Euler approximation performs asymptotically optimal in the class of all numerical methods that are based on a finite number of observations of the driving Brownian motion.

1. Introduction. We consider a d-dimensional system of stochastic differential equations

$$dX(t) = a\big(t, X(t)\big)\,dt + \sigma\big(t, X(t)\big)\,dW(t), \qquad t \in [0,1], \tag{1}$$

with initial value $X(0)$. Here W denotes an m-dimensional Brownian motion, and $a\colon [0,1] \times \mathbb{R}^d \to \mathbb{R}^d$ and $\sigma\colon [0,1] \times \mathbb{R}^d \to \mathbb{R}^{d\times m}$ satisfy certain smoothness conditions.

In most cases an explicit solution of (1) will not be available such that an approximation must be used. In this paper we analyze numerical methods that are based on the initial value $X(0)$ and finitely many sequential observations

$$W(\tau_1), \dots, W(\tau_\nu)$$

of the driving Brownian motion W. Except for measurability conditions we do not impose any further restrictions. The kth site τ_k may depend on the previous evaluations $X(0), W(\tau_1), \dots, W(\tau_{k-1})$ and the total number ν of observations of W may be determined by a stopping rule. Finally, the resulting discrete data may be used in any way to generate an approximation.

We aim to find an approximation $\overline{X}$ that is pathwise close to the whole corresponding d-dimensional trajectory of the solution X. As a natural measure of performance we consider the maximum distance

$$\|X - \overline{X}\|_\infty = \max_{t\in[0,1]} \max_{1\le i\le d} |X_i(t) - \overline{X}_i(t)|,$$

Received May 2000; revised October 2000.

AMS 2000 subject classifications. Primary 65U05; secondary 60H10.

Key words and phrases. Systems of stochastic differential equations, pathwise uniform approximation, asymptotic optimality, adaptive method.

of the d components of X and $\overline{X}$ on the interval [0, 1], and we define the error of $\overline{X}$ by averaging over all trajectories, that is,

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

As a rough measure for the computational cost of $\overline{X}$ we use the expected number $n(\overline{X})$ of observations of W. We adress the following question: What is the minimum cost necessary to achieve an error $e_q(\overline{X}) \le \varepsilon$?

Up to now, answers to this question are partial with respect to the numerical methods or the system (1) of stochastic differential equations. Faure (1992) determines an upper bound with an unspecified constant for the error of a piecewise interpolated equidistant Euler scheme. Hofmann, Müller-Gronbach and Ritter (2000b) introduce an approximation that performs asymptotically optimal for systems with additive noise, that is, $\sigma(t, x) = \sigma(t)$.

In the present paper we provide an answer for the general case. We introduce an adaptive discretization that takes into account the local smoothness of the solution. Under standard regularity conditions, the smoothness of each component X_i of the solution at the point $(t, X(t))$ is determined by

$$E\big((X_i(t+\delta) - X_i(t))^2 \mid X(t)\big) = \sum_{j=1}^{m} \sigma_{i,j}^2\big(t, X(t)\big) \cdot \delta + o(\delta).$$

Hence $\sum_{j=1}^m \sigma_{i,j}^2(t, X(t))$ serves as a conditional Hölder constant for X_i. Due to our error criterion it is reasonable to evaluate W more often in regions where the maximum conditional Hölder constant

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

is large and vice versa. Roughly speaking, we apply this rule by taking the step-size $\tau_{k+1} - \tau_k$ proportional to the squared inverse of the current value of $\sigma^*(\tau_k, X(\tau_k))$ with a proportionality constant that depends on the average size of σ^* on the interval [0, 1]. It suffices to use the interpolated Euler Scheme with this discretization to obtain an asymptotically optimal approximation $\widehat{X}_{q,n}^{**}$ with input parameter n. The corresponding error satisfies

$$\begin{aligned}
&\lim_{n\to\infty} \big(n(\widehat{X}_{q,n}^{**}) / \ln n(\widehat{X}_{q,n}^{**})\big)^{1/2} \cdot e_q(\widehat{X}_{q,n}^{**}) \\
&\qquad = \left(E\left(\int_0^1 (\sigma^*(t, X(t))^2 \, dt \right)^{q/(q+2)} \right)^{(q+2)/2q} \Big/ \sqrt{2};
\end{aligned}$$

see Theorem 1.

No other method that uses N observations of W on the average can lead to better results. In fact, we show that

$$\liminf_{N\to\infty} (N/\ln N)^{1/2} \cdot e_q(\overline{X}_N) \geq \left(E\left(\int_0^1 (\sigma^*(t, X(t))^2\, dt\right)^{q/(q+2)}\right)^{(q+2)/2q} \Big/ \sqrt{2}$$

for every sequence of methods $\overline{X}_N$ such that

$$n(\overline{X}_N) \leq N;$$

see Theorem 3.

In Section 2 we state our assumptions on the system (1). The conditions imposed on the drift and diffusion coefficients as well as on the initial value are standard assumptions for analyzing stochastic differential systems; see, for example, Bouleau and Lépingle [(1994), Chapter 5].

The adaptive method $\widehat{X}^{**}_{q,n}$ is defined in Section 3. This method adjusts the number ν of evaluations of a given trajectory of the driving Brownian motion W to the corresponding trajectory of the solution X. Hence no a priori bound on the computation time is available for the user. Alternatively, we introduce a modified version of this method that uses the same number of evaluations for every trajectory.

In Section 4 we present the exact asymptotic performance of both the adaptive methods and the interpolated Euler Scheme with equidistant step-size. All methods yield the same order of error but the corresponding asymptotic constants may heavily differ. This is illustrated by the case of a geometric Brownian motion in (1) which results in linear and exponential dependence on the squared volatility for the respective constants of $\widehat{X}^{**}_{q,n}$ and the equidistant Euler Scheme.

Lower bounds for arbitrary numerical methods that are based on a finite number of evaluations of W are given in Section 5. In particular these results show that for being optimal it is inevitable to adjust the number ν of observations to the current trajectory of the solution.

Proofs are postponed to Section 6. Finally, in the Appendix, we present an explicit formula for moments of the maximum of a geometric Brownian motion which is needed for the calculation of some asymptotic constants.

The central role of conditional Hölder constants has already been observed and successfully exploited by Hofmann, Müller-Gronbach and Ritter (2000c) for L_2-approximation of a scalar equation (1). In this case, basically, an interpolated Milstein Scheme with step-size proportional to the inverse of $|\sigma|$ performs asymptotically optimal. See Remark 4 for a more detailed comparison.

We further mention that pathwise approximation of a stochastic differential equation is strongly connected to the problem of reconstruction of a stochastic process based on observations of the process itself. See Traub, Wasilkowski and Woźniakowski (1988), Maiorov and Wasilkowski (1996), Müller-Gronbach and

Ritter (1997, 1998) and the references therein for corresponding results with respect to uniform and L_2-approximation.

Frequently, discrete norms, that is, the distance between the solution and its approximation at a finite number of points, are studied in the literature. See Hofmann, Müller-Gronbach and Ritter (2000a) for a discussion on this topic and Kloeden and Platen (1995), Milstein (1995) and Talay (1995) for results and references.

2. Assumptions. We use $|\cdot|_\infty$ to denote the maximum-norm of a matrix or a vector. By $\|\cdot\|_p$ we denote the L_p-norm of real-valued functions on $[0,1]$. Furthermore, we define $\|f\|_p = \max_{1\le i\le d}\|f_i\|_p$ for $\mathbb{R}^d$-valued functions f on $[0,1]$.

Throughout this paper we assume that the drift coefficient a, the diffusion coefficient σ and the initial value $X(0)$ have the following properties.

(A) There exists a constant $K>0$ such that $f=a$ and $f=\sigma$ satisfy

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

for all $s,t\in[0,1]$ and $x,y\in\mathbb{R}^d$.

(B) The initial value $X(0)$ is independent of W and

$$E\,|X(0)|_\infty^{\max(2,q)} < \infty.$$

Given the above properties, a pathwise unique strong solution of the equation (1) with initial value $X(0)$ exists. In particular the conditions assure that

$$E\,\|X\|_\infty^{\max(2,q)} < \infty.$$

3. The adaptive method. The adaptive method basically works in two steps. First, we use a rough approximation at equidistant discrete points to estimate the drift and diffusion coefficients as well as the maximum Hölder constant of the current trajectory of the solution. The latter determines the number and distribution of additional observation sites for the Brownian path to be used. At the second stage we piecewise freeze the drift and diffusion coefficients and we take the additional observations to refine the approximation.

We use the Euler Scheme to compute approximations at discrete points. For every discretization

$$0=\tau_0<\cdots<\tau_k=1$$

this scheme is defined by

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

and

$$\widehat{X}(\tau_{\ell+1}) = \widehat{X}(\tau_\ell) + a\big(\tau_\ell, \widehat{X}(\tau_\ell)\big) \cdot (\tau_{\ell+1} - \tau_\ell) + \sigma\big(\tau_\ell, \widehat{X}(\tau_\ell)\big) \cdot \big(W(\tau_{\ell+1}) - W(\tau_\ell)\big),$$

where $\ell = 0, \ldots, k-1$. A global approximation $\widehat{X}$ for X on $[0, 1]$ is obtained by piecewise linear interpolation of the data $(\tau_\ell, \widehat{X}(\tau_\ell))$. Obviously $\widehat{X}$ depends on W only through its values at the discretization points.

3.1. *The adaptive method $\widehat{X}^{**}_{q,n}$ with variable number of discretization points.* For $n \in \mathbb{N}$ choose $k_n \in \mathbb{N}$ and compute the Euler approximation

$$x_\ell = \widehat{X}^{**}_{q,n}(\tau_{\ell,0})$$

at the equidistant points

$$\tau_{\ell,0} = \ell / k_n, \qquad \ell = 0, \ldots, k_n - 1.$$

The corresponding estimates of the maximum conditional Hölder constant are

$$\sigma^*_\ell = \sigma^*(\tau_{\ell,0}, x_\ell), \qquad \ell = 0, \ldots, k_n - 1.$$

Put

$$S = \left(1/k_n \cdot \sum_{\ell=0}^{k_n-1} (\sigma^*_\ell)^2\right)^{1/2}$$

and determine the number of additional knots to be placed equidistantly in the subinterval $]\tau_{i,0}, \tau_{i+1,0}[$ by

$$\mu_{q,i} = \begin{cases} \left\lfloor \left((\sigma^*_i)^2 \Big/ \sum_{\ell=0}^{k_n-1} (\sigma^*_\ell)^2 \right) \cdot n \cdot S^{2q/(q+2)} \right\rfloor & \text{if } S > 0, \\ 0 & \text{otherwise.} \end{cases}$$

The resulting adaptive discretization of the subinterval $[\tau_{i,0}, \tau_{i+1,0}]$ is thus given by

$$\tau_{i,j} = \tau_{i,0} + j/(k_n \cdot (\mu_{q,i} + 1)), \qquad j = 0, \ldots, \mu_{q,i} + 1.$$

Next, put

$$a_i = a(\tau_{i,0}, x_i), \qquad \sigma_i = \sigma(\tau_{i,0}, x_i),$$

and use the Euler method

$$\widehat{X}^{**}_{q,n}(\tau_{i,j+1}) = \widehat{X}^{**}_{q,n}(\tau_{i,j}) + a_i \cdot (\tau_{i,j+1} - \tau_{i,j}) + \sigma_i \cdot \big(W(\tau_{i,j+1}) - W(\tau_{i,j})\big),$$

on every subinterval $[\tau_{i,0}, \tau_{i+1,0}]$ without updating the drift and diffusion coefficient. Finally, piecewise linear interpolation yields the global approximation on $[0, 1]$.

Clearly, the distribution of the majority of knots used by $\widehat{X}^{**}_{q,n}$ should be adapted to the particular trajectory. On the other hand, the quality of adaption crucially depends on the estimates of the maximum Hölder constant. We therefore adjust the size k_n of the initial equidistant discretization by

$$\lim_{n\to\infty} k_n/n = 0$$

and

$$\lim_{n\to\infty} (k_n/n)\cdot \ln n = \infty.$$

Note that the total number of observations of W that are used by $\widehat{X}^{**}_{q,n}$ is roughly determined by the L_2-average

$$\|\sigma^*\|_2 = \left(\int_0^1 \bigl(\sigma^*(t, X(t))\bigr)^2\, dt\right)^{1/2}$$

of the maximum conditional Hölder constant, which may heavily depend on the particular trajectory. Consequently, there is no a priori bound on the computation time available for the user. This can be overcome by using the following version $\widehat{X}^*_n$ of the adaptive method. However, note that the property of bounded computation time is not free of charge, see Theorem 1, the second equality in Theorem 3 as well as Example 1.

3.2. *The adaptive method $\widehat{X}^*_n$ with at most n discretization points.* In contrast to $\widehat{X}^{**}_{q,n}$, the method $\widehat{X}^*_n$ does not depend on the parameter q. It coincides with $\widehat{X}^{**}_{q,n}$ up to the fact that the numbers $\mu_{q,i}$ of evaluation points in the subintervals $]\tau_{i,0}, \tau_{i+1,0}[$ are replaced by

$$\mu_i = \begin{cases} \left\lfloor \left((\sigma_i^*)^2 \Big/ \sum_{\ell=0}^{k_n-1} (\sigma_\ell^*)^2\right)\cdot (n-k_n)\right\rfloor & \text{if } S > 0,\\ \lfloor (n-k_n)/k_n \rfloor & \text{otherwise,}\end{cases}$$

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

Clearly, $\widehat{X}^*_n$ uses at most n observations of W for every trajectory. Nevertheless, the distribution of the observation sites is still adapted to the trajectory. On the subinterval $]\tau_{i,0}, \tau_{i+1,0}[$ the method $\widehat{X}^*_n$ takes steps of size roughly given by $(1/n)\cdot \|\sigma^*\|_2^2/(\sigma^*(\tau_{i,0}, X(\tau_{i,0})))^2$.

4. Performance of the adaptive method. To every system (1) and every $q \ge 1$ we associate the constants

$$C_q^{**} = \bigl(E\,\|\sigma^*\|_2^{2q/(q+2)}\bigr)^{(q+2)/2q},$$
$$C_q^{*} = (E\,\|\sigma^*\|_2^{q})^{1/q}.$$

Let $n(\widehat{X}^{**}_{q,n})$ denote the expected number of observations of W that are used by $\widehat{X}^{**}_{q,n}$ and recall that $\widehat{X}^*_n$ uses at most n observations of W for every trajectory.

THEOREM 1. *The adaptive methods* $\widehat{X}^{**}_{q,n}$ *and* $\widehat{X}^{*}_{n}$ *satisfy*

$$\lim_{n\to\infty} \big(n(\widehat{X}^{**}_{q,n})/\ln n(\widehat{X}^{**}_{q,n})\big)^{1/2} \cdot e_q(\widehat{X}^{**}_{q,n}) = C^{**}_q/\sqrt{2}$$

and

$$\lim_{n\to\infty} (n/\ln n)^{1/2} \cdot e_q(\widehat{X}^{*}_{n}) = C^{*}_q/\sqrt{2}$$

for every system (1).

In a first step, we justify the use of our adaptive method by showing its superiority to the piecewise interpolated Euler Scheme $\widehat{X}^{\mathrm{e}}_n$ with step-size $1/n$. We stress that much stronger optimality properties of our method will be established in Section 5.

The constant associated with $\widehat{X}^{\mathrm{e}}_n$ is given by the qth mean

$$C^{\mathrm{e}}_q = (E\,\|\sigma^*\|^q_\infty)^{1/q}$$

of the global maximum of the maximum conditional Hölder constant σ^*.

THEOREM 2. *The interpolated equidistant Euler Scheme* $\widehat{X}^{\mathrm{e}}_n$ *satisfies*

$$\lim_{n\to\infty} (n/\ln n)^{1/2} \cdot e_q(\widehat{X}^{\mathrm{e}}_{n}) = C^{\mathrm{e}}_q/\sqrt{2}$$

for every system (1).

Clearly, the order of convergence is the same for all of the above methods. Note, however, that

$$C^{**}_q \le C^{*}_q \le C^{\mathrm{e}}_q$$

with strict inequality in most cases. See Remark 1 for a characterization of equality.

Furthermore, we have

$$\sup_{1\le q\le 2} C^{**}_q = C^*_1 = \inf_{q\ge 1} C^*_q,$$

$$\sup_{q\ge 1} C^{**}_q = C^*_2 = \inf_{q\ge 2} C^*_q.$$

EXAMPLE 1. Consider the linear equation

$$dX(t) = bX(t)\,dW(t)$$

with $b > 0$ and initial value $X(0) = 1$ in the case $d = m = 1$. The solution is the geometric Brownian motion

$$X(t) = \exp\big(-b^2/2 \cdot t + b \cdot W(t)\big)$$

with drift zero. For $q = 2$ we obtain

$$C_2^{\mathrm{e}} = b \cdot \big(3 \cdot \exp(b^2) \cdot \Phi(3b/2) - \Phi(b/2)\big)^{1/2}$$

from Theorem 5. Here Φ denotes the standard normal distribution function. Straightforward calculations yield

$$C_2^* = \big(\exp(b^2) - 1\big)^{1/2}.$$

Finally, using Theorem 5 again, we have

$$\begin{aligned} C_2^{**} &= E(\|b \cdot X\|_2) \\ &\le b \cdot E(\|X\|_\infty^{1/2} \cdot \|X\|_1^{1/2}) \\ &\le b \cdot \big(E(\|X\|_\infty)\big)^{1/2} \cdot \big(E(\|X\|_1)\big)^{1/2} \\ &= (b \cdot C_1^{\mathrm{e}})^{1/2} \\ &\le b \cdot (b+2) \cdot 2^{-1/2}. \end{aligned}$$

For $\widehat{X}_n^{\mathrm{e}}$ and $\widehat{X}_n^*$ the asymptotic constant depends exponentially on the parameter b^2. For $\widehat{X}_{2,n}^{**}$ we only have a linear dependence on b^2.

5. General methods and lower bounds. The adaptive methods introduced in section 3 use a realization of the initial value $X(0)$ and a finite number of observations of a trajectory of the driving Brownian motion W. We now present lower bounds that hold for arbitrary methods of the above type and arbitrary systems.

Fix a and σ, and consider the corresponding system (1). We follow the notation in Hofmann, Müller-Gronbach and Ritter (2000c). Formally, a general method is then defined by mappings

$$\begin{aligned} \psi_k &: \mathbb{R}^{(k-1)\cdot m+d} \to [0,1], \\ \chi_k &: \mathbb{R}^{k\cdot m+d} \to \{\mathrm{STOP}, \mathrm{GO}\}, \\ \phi_k &: \mathbb{R}^{k\cdot m+d} \to \big(L_\infty([0,1])\big)^d \end{aligned}$$

for $k \in \mathbb{N}$. The sequential observation of a trajectory w of the Brownian motion W starts at the knot $\psi_1(x)$, which may depend on the realization x of the initial value. After k steps we have observed the data

$$\Psi_k(x, w) = (x', y_1', \dots, y_k')',$$

where

$$y_1 = w\big(\psi_1(x)\big), \quad \dots, \quad y_k = w\big(\psi_k\big((x', y_1', \dots, y_{k-1}')'\big)\big).$$

After each step we decide to stop or to further evaluate w according to the value of

$$\chi_k\big(\Psi_k(x, w)\big).$$

The total number of observations of w is then given by

$$\nu(x,w)=\min\{k\in\mathbb{N}:\chi_k(\Psi_k(x,w))=\mathrm{STOP}\}.$$

If $\nu(x,w)<\infty$ then the data

$$\Psi(x,w)=\Psi_{\nu(x,w)}(x,w)$$

are used to construct the approximation $\phi_{\nu(x,w)}(\Psi(x,w))$.

Formally, we assume measurability of the mappings ψ_k, χ_k, and ϕ_k. Furthermore, for obvious reasons, we restrict to the case $\nu(X(0),W)<\infty$ with probability one. The resulting method is given by

$$\overline{X}=\phi_{\nu(X(0),W)}(\Psi(X(0),W)).$$

Note that

$$D(X(0),W)=\{\psi_1(X(0)),\dots,\psi_{\nu(X(0),W)}(\Psi_{\nu(X(0),W)-1}(X(0),W))\}$$

is the set of observation sites used by $\overline{X}$. As previously we analyze the error $e_q(\overline{X})$ with respect to the average number

$$n(\overline{X})=E\,\#D(X(0),W)$$

of evaluations of W. Here # denotes the cardinality of a set.

Let $\mathbb{X}^{**}$ denote the class of all methods of the above form, and note that $\widehat{X}^{**}_{q,n}\in\mathbb{X}^{**}$ for the adaptive method from Section 3.1.

Put

$$\mathbb{X}^{**}_N=\{\overline{X}\in\mathbb{X}^{**}:n(\overline{X})\le N\}$$

for $N\in\mathbb{N}$, and let

$$e^{**}_q(N)=\inf\{e_q(\overline{X}):\overline{X}\in\mathbb{X}^{**}_N\}$$

denote the minimal error that can be obtained by methods that use at most N sequential observations of W on the average.

As a subclass $\mathbb{X}^*\subset\mathbb{X}^{**}$ we consider all methods that use the same number of observations for all trajectories. Formally, the mappings χ_k are constant and $\nu=\min\{k\in\mathbb{N}:\chi_k=\mathrm{STOP}\}$. Put

$$\mathbb{X}^*_N=\{\overline{X}\in\mathbb{X}^*:n(\overline{X})\le N\}$$

as well as

$$e^*_q(N)=\inf\{e_q(\overline{X}):\overline{X}\in\mathbb{X}^*_N\}.$$

Recall that the adaptive method $\widehat{X}^*_n$ from Section 3.2 uses at most n observations for each trajectory. Hence $\widehat{X}^*_n\in\mathbb{X}^*_n$ by a suitable definition of the mappings ψ_k, χ_k, and ϕ_k.

THEOREM 3. *The minimal errors satisfy*

$$\lim_{N\to\infty} (N/\ln N)^{1/2} \cdot e_q^{**}(N) = C_q^{**}/\sqrt{2},$$
$$\lim_{N\to\infty} (N/\ln N)^{1/2} \cdot e_q^{*}(N) = C_q^{*}/\sqrt{2}$$

for every equation (1) *and each* $q \geq 1$.

Note that the estimates from Theorem 1 match with the corresponding estimates from Theorem 3.

COROLLARY 1. *For each* $q \geq 1$ *the methods* $\widehat{X}_{q,n}^{**}$ *and* $\widehat{X}_n^{*}$ *are asymptotically optimal in the respective classes* $\mathbb{X}_N^{**}$ *with* $N = \lceil n(\widehat{X}_{q,n}^{**})\rceil$ *and* $\mathbb{X}_n^{*}$.

REMARK 1. Let

$$\alpha^*(t) = \bigl(E\bigl((\sigma^*\bigl(t, X(t)\bigr)^2\bigr)\bigr)^{1/2}$$

for each $t \in [0, 1]$. Note that in the one dimensional case, $d = 1$, α^* describes the smoothness of X only locally in time.

Due to the Markov property of X we have $C_q^{**} = C_q^{*}$ iff, with probability one,

$$\forall\, t \in [0,1]: \quad \sigma^*\bigl(t, X(t)\bigr) = \alpha^*(t). \tag{2}$$

In particular (2) holds for systems (1) with additive noise.

Next, it is easy to see that $C^* = C^{\mathrm{e}}$ iff, with probability one,

$$\forall\, t \in [0,1]: \quad \sigma^*\bigl(t, X(t)\bigr) = \sigma^*\bigl(0, X(0)\bigr). \tag{3}$$

In the one-dimensional and univariate case, $d = m = 1$, (3) is equivalent to the solution X being of the form

$$X(t) = X(0) + \int_0^t a\bigl(t, X(t)\bigr)\, dt \pm \sigma\bigl(0, X(0)\bigr) \cdot W(t).$$

Thus, the local smoothness of each trajectory does not vary in time.

REMARK 2. The optimal method $\widehat{X}_{q,n}^{**}$ crucially depends on the parameter q. Whether or not there is a single method $\widehat{X}_n^{**}$ which is optimal for all $q \geq 1$ remains an open question. However, by Lemma 9 and the Hölder inequality we have

$$\begin{aligned}
&\limsup_{n\to\infty} \bigl(n(\widehat{X}_{1,n}^{**})/\ln n(\widehat{X}_{1,n}^{**})\bigr)^{1/2} \cdot e_q(\widehat{X}_{1,n}^{**}) \\
&\quad \leq \bigl(E\,\|\sigma^*\|_2^{2/3}\bigr)^{1/2} \cdot \bigl(E\,\|\sigma^*\|_2^{2q/3}\bigr)^{1/q}/\sqrt{2} \\
&\quad \leq \bigl(E\,\|\sigma^*\|_2^{2q/3}\bigr)^{3/2q}/\sqrt{2} \\
&\quad \leq C_q^*/\sqrt{2}
\end{aligned}$$

for each $q \geq 1$.

Note that the last two inequalities are strict except for the specific case corresponding to (2). The method $\widehat{X}^{**}_{1,n}$ is therefore superior to the method $\widehat{X}^{*}_{n}$ for every $q \geq 1$.

REMARK 3. Roughly speaking, the method $\widehat{X}^{**}_{q,n}$ uses step-sizes proportional to $1/(\sigma^*(\tau_\ell, \widehat{X}(\tau_\ell)))^2$ with a path dependent proportionality constant determined by S; see Section 3.1. The latter seems to be essential for its superior asymptotic performance. For instance, define the method $\widetilde{X}_n$ in the same way as $\widehat{X}^{**}_{q,n}$ with the numbers $\mu_{q,i}$ replaced by

$$\mu_i = \lfloor (\sigma_i^*)^2 \cdot n \rfloor$$

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

The ratios of the number of knots assigned to the subintervals by $\widetilde{X}_n$ and $\widehat{X}^{**}_{q,n}$ basically coincide,

$$\mu_i/\mu_j = \mu_{q,i}/\mu_{q,j},$$

but the method $\widetilde{X}_n$ does not adjust the total number of knots in an appropriate way.

In fact, the methods of proof in Section 6 carry over to show that

$$\lim_{n\to\infty} \big(n(\widetilde{X}_n)/\ln n(\widetilde{X}_n)\big)^{1/2} \cdot e_q(\widetilde{X}_n) = C_2^*/\sqrt{2}.$$

Hence, the method $\widetilde{X}_n$ is inferior to $\widehat{X}^{**}_{q,n}$, and it performs even worse than $\widehat{X}^{*}_{n}$ if $q < 2$.

Note further that, in contrast to $\widehat{X}^{**}_{q,n}$, the method $\widetilde{X}_n$ progresses from the left to the right. It remains an open question whether there exists a method of the latter kind that performs asymptotically optimal.

REMARK 4. The concept of choosing the discretization according to the size of conditional Hölder constants also works for other optimality criteria. Hofmann, Müller-Gronbach and Ritter (2000a, c) analyze numerical methods for pathwise approximation of a scalar equation (1) with respect to the averaged L_2-error

$$e(\overline{X}) = (E\ \|X - \overline{X}\|_2^2)^{1/2}.$$

They show that the minimum error in the class of all methods that use at most N observations of the driving Brownian motion W on the average,

$$e^{**}(N) = \inf\{e(\overline{X}) : \overline{X} \in \mathbb{X}^{**}_N\}$$

satisfies

$$\lim_{N\to\infty} N^{1/2} \cdot e^{**}(N) = E\|\sigma\|_1/\sqrt{6}.$$

Moreover, using a step-size proportional to $1/|\sigma|$ leads to an asymptotically optimal method. We conjecture that these results may be extended to the communitive system case by employing the discrete L_2-average

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

of the respective conditional Hölder constants.

6. Proofs. Instead of estimating $X-\overline{X}$ directly, we use an approximation $\check{X}$ that is based on the whole trajectory of W and analyze $X-\check{X}$ as well as $\check{X}-\overline{X}$ separately. For a discretization

$$0=t_0<\cdots<t_k=1, \tag{4}$$

the process $\check{X}$ is given by $\check{X}(0)=X(0)$ and

$$\check{X}(t)=\check{X}(t_\ell)+a\big(t_\ell,\check{X}(t_\ell)\big)\cdot(t-t_\ell)+\sigma\big(t_\ell,\check{X}(t_\ell)\big)\cdot\big(W(t)-W(t_\ell)\big) \tag{5}$$

for $t\in[t_\ell,t_{\ell+1}]$. Clearly, $\check{X}$ is not an implementable numerical scheme for the global approximation of X. Due to Theorem 4, if the discretization (4) is equidistant and k is chosen suitably with respect to $n(\overline{X})$ then $(E\,\|\check{X}-\overline{X}\|_\infty^q)^{1/q}$ is the dominating term asymptotically.

Throughout the following we let c denote unspecified constants that only depend on d, m, q and the constants from conditions (A) and (B).

Let

$$\Delta_\ell=t_{\ell+1}-t_\ell$$

for $\ell=0,\dots,k-1$, and put

$$\Delta_{\max}=\max_{0\le\ell\le k-1}\Delta_\ell.$$

Due to Hofmann, Müller-Gronbach and Ritter (2000b) the following upper bound holds for the error of $\check{X}$.

THEOREM 4. *Assume that a as well as σ satisfy condition* (A) *and let $X(0)$ satisfy* (B). *Then*

$$(E\,\|X-\check{X}\|_\infty^q)^{1/q}\le c\cdot\Delta_{\max}^{1/2}.$$

Observe further that, conditioned on $W(t_1),\dots,W(t_k)$, each component $\check{X}_i$ of $\check{X}$ is a weighted Brownian bridge with a linear trend on each subinterval $[t_\ell,t_{\ell+1}]$. Basically, error bounds for the uniform approximation of $\check{X}$ may thus be obtained by employing results on the uniform approximation of weighted Brownian bridges.

6.1. *Approximation of weighted Brownian bridges.* Throughout the following we use $B_1, B_2, \ldots$ to denote a sequence of independent Brownian bridges on the interval [0, 1] with $B_i(0) = B_i(1) = 0$. We put

$$\mathcal{G}_q(u; \alpha_1, \ldots, \alpha_n) = P\Big(\max_{1 \le i \le n} (\alpha_i \cdot \|B_i\|_\infty)^q > u\Big), \qquad u \ge 0$$

and

$$\mathcal{M}_q(\alpha_1, \ldots, \alpha_n) = E \max_{1 \le i \le n} (\alpha_i \cdot \|B_i\|_\infty)^q$$

for nonnegative numbers $\alpha_1, \ldots, \alpha_n$. Furthermore, we use the notation

$$\mathcal{G}_q(u; n) = \mathcal{G}_q(u; 1, \ldots, 1), \qquad \mathcal{M}_q(n) = \mathcal{M}_q(1, \ldots, 1).$$

LEMMA 1. *For all numbers $\alpha_1, \ldots, \alpha_n > 0$ it holds*

$$\big(\mathcal{M}_q(\alpha_1, \ldots, \alpha_n)\big)^{1/q} \ge \left(1/n \cdot \sum_{i=1}^{n} \alpha_i^2\right)^{1/2} \cdot \big(\mathcal{M}_q(n)\big)^{1/q}.$$

LEMMA 2. *Let $\alpha, \alpha_{1,n}, \ldots, \alpha_{n,n} > 0$ such that*

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

and

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

for every $\varepsilon > 0$. Then

$$\int_0^\infty \mathcal{G}_q\big(u \cdot (\ln n)^{q/2}; \alpha_{1,n}, \ldots, \alpha_{n,n}\big)\, du \le c,$$

$$\lim_{n \to \infty} \int_{(\alpha^2/2)^{q/2}}^\infty \mathcal{G}_q\big(u \cdot (\ln n)^{q/2}; \alpha_{1,n}, \ldots, \alpha_{n,n}\big)\, du = 0$$

and

$$\lim_{n \to \infty} \mathcal{G}_q\big(u \cdot (\ln n)^{q/2}; \alpha_{1,n}, \ldots, \alpha_{n,n}\big) = 1$$

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

See Ritter (1990) for proofs of Lemmas 1 and 2 in the particular case $\alpha_{i,n} = \alpha$. The latter immediately yields the following fact.

COROLLARY 2. *Under the assumptions of Lemma 2 we have*

$$\lim_{n \to \infty} (\ln n)^{-1/2} \cdot \big(\mathcal{M}_q(\alpha_{1,n}, \ldots, \alpha_{n,n})\big)^{1/q} = \alpha \cdot 2^{-1/2}.$$

Next, consider the discretization (4) and let

$$B_{1,\ell}, \ldots, B_{m,\ell}$$

denote Brownian bridges on the interval $[t_\ell, t_{\ell+1}]$ so that $B_{1,0}, \ldots, B_{m,k-1}$ are independent. Fix numbers

$$\beta_{i,j,\ell}, \qquad j = 1, \ldots, m,\ i = 1, \ldots, d,\ \ell = 0, \ldots, k-1,$$

and let

$$\beta \geq \max_{1 \leq i \leq d} \max_{0 \leq \ell \leq k-1} \left(\Delta_\ell \cdot \sum_{j=1}^m \beta_{i,j,\ell}^2 \right)^{1/2}.$$

LEMMA 3. *For every $z \geq 0$ it holds*

$$E \max_{1 \leq i \leq d} \max_{0 \leq \ell \leq k-1} \sup_{t_\ell \leq t \leq t_{\ell+1}} \left| \sum_{j=1}^m \beta_{i,j,\ell} \cdot B_{j,\ell}(t) \right|^q \leq z + d \cdot \int_z^\infty \mathcal{G}_q(u/\beta^q; k)\, du.$$

PROOF. We have

$$\begin{aligned} &E \max_{1 \leq i \leq d} \max_{0 \leq \ell \leq k-1} \sup_{t_\ell \leq t \leq t_{\ell+1}} \left| \sum_{j=1}^m \beta_{i,j,\ell} \cdot B_{j,\ell}(t) \right|^q \\ &\quad = \int_0^\infty P\left(\max_{1 \leq i \leq d} \max_{0 \leq \ell \leq k-1} \sup_{t_\ell \leq t \leq t_{\ell+1}} \left| \sum_{j=1}^m \beta_{i,j,\ell} \cdot B_{j,\ell}(t) \right|^q > u \right) du \\ &\quad \leq z + \sum_{i=1}^d \int_z^\infty P\left(\max_{0 \leq \ell \leq k-1} \sup_{t_\ell \leq t \leq t_{\ell+1}} \left| \sum_{j=1}^m \beta_{i,j,\ell} \cdot B_{j,\ell}(t) \right|^q > u \right) du \\ &\quad = z + \sum_{i=1}^d \int_z^\infty P\left(\max_{0 \leq \ell \leq k-1} \sup_{t_\ell \leq t \leq t_{\ell+1}} \left| \left(\sum_{j=1}^m \beta_{i,j,\ell}^2 \right)^{1/2} \cdot B_{1,\ell}(t) \right|^q > u \right) du \end{aligned}$$

by the independence of $B_{1,0}, \ldots, B_{m,k-1}$.

Renormalize each interval $[t_\ell, t_{\ell+1}]$ to $[0, 1]$ to obtain

$$\begin{aligned} &\int_z^\infty P\left(\max_{0 \leq \ell \leq k-1} \sup_{t_\ell \leq t \leq t_{\ell+1}} \left| \left(\sum_{j=1}^m \beta_{i,j,\ell}^2 \right)^{1/2} \cdot B_{1,\ell}(t) \right|^q > u \right) du \\ &\qquad = \int_z^\infty P\left(\max_{0 \leq \ell \leq k-1} \sup_{0 \leq t \leq 1} \left| \left(\sum_{j=1}^m \beta_{i,j,\ell}^2 \right)^{1/2} \cdot \Delta_\ell^{1/2} \cdot B_\ell(t) \right|^q > u \right) du \\ &\qquad \leq \int_z^\infty P\left(\beta^q \cdot \max_{0 \leq \ell \leq k-1} \|B_\ell\|_\infty^q > u \right) du, \end{aligned}$$

which completes the proof. $\square$

6.2. *Proofs of the lower bounds in Theorems* 3 *and* 2. Fix a discretization (4) and put

$$U_\ell = \big(t_\ell, X(t_\ell)\big), \qquad \check{U}_\ell = \big(t_\ell, \check{X}(t_\ell)\big),$$

as well as

$$\mathcal{A} = \left(\sum_{\ell=0}^{k-1} \big(\sigma^*(U_\ell)\big)^2 \cdot \Delta_\ell\right)^{1/2}, \qquad \check{\mathcal{A}} = \left(\sum_{\ell=0}^{k-1} \big(\sigma^*(\check{U}_\ell)\big)^2 \cdot \Delta_\ell\right)^{1/2}.$$

LEMMA 4. *For every discretization* (4) *it holds that*

$$\Big(E \max_{0\le\ell\le k-1} |\sigma^*(U_\ell) - \sigma^*(\check{U}_\ell)|^q\Big)^{1/q} \le c \cdot \Delta_{\max}^{1/2}, \tag{6}$$

$$(E|\mathcal{A} - \check{\mathcal{A}}|^q)^{1/q} \le c \cdot \Delta_{\max}^{1/2} \tag{7}$$

and

$$\mathcal{A} \le c \cdot (1 + \|X\|_\infty). \tag{8}$$

PROOF. Due to (A) we have

$$\begin{aligned}
|\sigma^*(U_\ell) - \sigma^*(\check{U}_\ell)| &\le d \cdot \max_{1\le i\le d} \left| \left(\sum_{j=1}^m \sigma_{i,j}^2(U_\ell)\right)^{1/2} - \left(\sum_{j=1}^m \sigma_{i,j}^2(\check{U}_\ell)\right)^{1/2} \right| \\
&\le d \cdot \max_{1\le i\le d} \left(\sum_{j=1}^m \big(\sigma_{i,j}(U_\ell) - \sigma_{i,j}(\check{U}_\ell)\big)^2\right)^{1/2} \\
&\le c \cdot |X(t_\ell) - \check{X}(t_\ell)|_\infty \\
&\le c \cdot \|X - \check{X}\|_\infty.
\end{aligned} \tag{9}$$

Hence

$$\Big(E\Big(\max_{0\le\ell\le k-1} |\sigma^*(U_\ell) - \sigma^*(\check{U}_\ell)|^q\Big)\Big)^{1/q} \le c \cdot (E\,\|X - \check{X}\|_\infty^q)^{1/q} \le c \cdot \Delta_{\max}^{1/2}$$

by Theorem 4, which proves (6).

Furthermore, it holds

$$\begin{aligned}
|\mathcal{A} - \check{\mathcal{A}}| &\le \left(\sum_{\ell=0}^{k-1} \big(\sigma^*(U_\ell) - \sigma^*(\check{U}_\ell)\big)^2 \cdot \Delta_\ell\right)^{1/2} \\
&\le \max_{0\le\ell\le k-1} |\sigma^*(U_\ell) - \sigma^*(\check{U}_\ell)|,
\end{aligned}$$

such that the second inequality is a consequence of (6).

Finally, by (A),

$$\begin{aligned}\mathcal{A} &\le \max_{0\le\ell\le k-1} \sigma^*(U_\ell)\\ &= \max_{0\le\ell\le k-1}\max_{1\le i\le d}\left(\sum_{j=1}^m \sigma_{i,j}^2(U_\ell)\right)^{1/2}\\ &\le c\cdot \max_{0\le\ell\le k-1}\max_{1\le i\le d}\big(m\cdot\big(1+|X(t_\ell)|_\infty\big)^2\big)^{1/2}\\ &\le c\cdot(1+\|X\|_\infty),\end{aligned}$$

which completes the proof. $\square$

Consider an arbitrary method $\overline{X}\in\mathbb{X}^{**}$ and recall from Section 4 that

$$D\big(X(0),W\big)=\big\{\psi_1\big(X(0)\big),\dots,\psi_{\nu(X(0),W)}\big(\Psi_{\nu(X(0),W)-1}\big(X(0),W\big)\big)\big\}$$

is the set of observation sites used by $\overline{X}$. Let

$$d_\ell=\#\big(D\big(X(0),W\big)\cap]t_\ell,t_{\ell+1}[\big)+1,$$
$$L=\big\{\ell\in\{0,\dots,k-1\}:\sigma^*(\check{U}_\ell)\ne 0\big\},$$

and put

$$\delta=\max\bigg(1,\sum_{\ell\in L}d_\ell\bigg).$$

Note that d_ℓ as well as δ are measurable functions of Ψ.

LEMMA 5. *If $\overline{X}$ uses the knots $t_0,\dots,t_k$ then*

$$e_q(\overline{X})\ge\big(E\big((\check{\mathcal{A}}^q/\delta^{q/2})\cdot\mathcal{M}_q(\delta)\big)\big)^{1/q}-c\cdot\Delta_{\max}^{1/2}.$$

PROOF. Observing Theorem 4, it suffices to show that

$$E\,\|\check{X}-\overline{X}\|_\infty^q\ge E\big((\check{\mathcal{A}}^q/\delta^{q/2})\cdot\mathcal{M}_q(\delta)\big). \tag{10}$$

Since $\check{U}_\ell$ is a function of Ψ, we have

$$\check{X}(t)-E\big(\check{X}(t)|\Psi\big)=\sigma(\check{U}_\ell)\cdot\big(W(t)-E\big(W(t)|\Psi\big)\big)$$

for $t\in[t_\ell,t_{\ell+1}]$. Conditioned on Ψ, the discretization $D(X(0),W)$ is fixed and the process $W-E(W|\Psi)$ consists of independent, m-dimensional Brownian bridges corresponding to each subinterval. In particular it follows that, conditioned on Ψ, the distribution of the process $\check{X}-E(\check{X}|\Psi)$ is symmetric with respect to the zero function. Hence

$$E(\|\check{X}-\overline{X}\|_\infty^q|\Psi)\ge E\big(\|\check{X}-E(\check{X}|\Psi)\|_\infty^q\big|\Psi\big); \tag{11}$$

see Traub, Wasilkowski and Woźniakowski (1988).

Let

$$i(t,x)=\min\left\{i:\big(\sigma^*(t,x)\big)^2=\sum_{j=1}^m \sigma_{i,j}^2(t,x)\right\}.$$

Then

$$\begin{aligned}
&E\big(\|\check X-E(\check X|\Psi)\|_\infty^q|\Psi\big)\\
&\qquad=E\left(\max_{1\le i\le d}\max_{0\le \ell\le k-1}\sup_{t_\ell\le t\le t_{\ell+1}}\left|\sum_{j=1}^m \sigma_{i,j}(\check U_\ell)\cdot\big(W_j(t)-E\big(W_j(t)|\Psi\big)\big)\right|^q|\Psi\right)\\
(12)\qquad&\qquad\ge E\left(\max_{0\le \ell\le k-1}\sup_{t_\ell\le t\le t_{\ell+1}}\left|\sum_{j=1}^m \sigma_{i(\check U_\ell),j}(\check U_\ell)\cdot\big(W_j(t)-E\big(W_j(t)|\Psi\big)\big)\right|^q|\Psi\right)\\
&\qquad=E\Big(\max_{0\le \ell\le k-1}\sup_{t_\ell\le t\le t_{\ell+1}}\big|\sigma^*(\check U_\ell)\cdot\big(W_1(t)-E\big(W_1(t)|\Psi\big)\big)\big|^q\;|\Psi\Big).
\end{aligned}$$

Recall that d_ℓ is the number of subintervals in $[t_\ell,t_{\ell+1}]$. Renormalize each subinterval to [0, 1], and apply Lemma 1 to obtain

$$\begin{aligned}
E(\|\check X-\overline{X}\|_\infty^q|\Psi)&\ge E\Big(\max_{\ell\in L}\sup_{t_\ell\le t\le t_{\ell+1}}\big|\sigma^*(\check U_\ell)\cdot\big(W_1(t)-E\big(W_1(t)|\Psi\big)\big)\big|^q|\Psi\Big)\\
&\ge(\check{\mathcal A}^q/\delta^{q/2})\cdot E\Big(\max_{1\le i\le\delta}\|B_i\|_\infty^q\Big)
\end{aligned}$$

from (11) and (12).

Clearly, (10) follows by taking expectations. □

We now turn to a sequence of methods $\overline{X}_N\in\mathbb{X}^{**}$. Choose a sequence of positive integers k_N such that

$$\lim_{N\to\infty} k_N/N=\lim_{N\to\infty} N/(k_N\cdot\ln N)=0. \tag{13}$$

Since $k_N=o(N)$ we may assume that $\overline{X}_N$ uses in particular the knots

$$t_\ell=\ell/k_N,\qquad \ell=0,\dots,k_N. \tag{14}$$

We use this discretization in (4), and we denote the related quantities canonically by $\check X_N,\mathcal{A}_N,\delta_N,\dots.$

Due to (6) we have

$$\lim_{N\to\infty}\max_{1\le\ell\le k_N}|\sigma^*(U_{N,\ell})-\sigma^*(\check U_{N,\ell})|=0 \tag{15}$$

in probability.

Furthermore,

$$P\Big(\lim_{N\to\infty} \mathcal{A}_N = \|\sigma^*\|_2\Big) = 1, \tag{16}$$

by the continuity of X and σ^*. Observing (7) we thus conclude that

$$\lim_{N\to\infty} \check{\mathcal{A}}_N = \|\sigma^*\|_2 \tag{17}$$

in probability.

Applying a subsequence argument, we may assume that (15) and (17) hold with probability 1. Since

$$\delta_N = \max\Bigg(1, \sum_{\ell\in L_N} d_{N,\ell}\Bigg) \geq \#L_N,$$

it follows in particular from (15) that

$$P\Big(\lim_{N\to\infty} \delta_N = \infty, \|\sigma^*\|_\infty > 0\Big) = P\big(\|\sigma^*\|_\infty > 0\big). \tag{18}$$

We first analyze the classes $\mathbb{X}^{**}$ and $\mathbb{X}^*$.

LEMMA 6. *If $\overline{X}_N \in \mathbb{X}^{**}_N$ for every N then*

$$\liminf_{N\to\infty} (N/\ln N)^{1/2} \cdot \big(E\big((\check{\mathcal{A}}_N^q/\delta_N^{q/2}) \cdot \mathcal{M}_q(\delta_N)\big)\big)^{1/q} \geq C_q^{**}/\sqrt{2}.$$

PROOF. Let ν_N denote the number of knots used by $\overline{X}_N$ and note that $\nu_N \geq k_N$. By monotonicity and concavity of the function

$$x \mapsto x/\ln x, \qquad x > e^2,$$

we have

$$N/\ln N \geq E\nu_N/\ln(E\,\nu_N) \geq E(\nu_N/\ln\nu_N) \geq E\big((\delta_N/\ln\delta_N)\cdot 1_{\{\delta_N>e^2\}}\big)$$

for N sufficiently large. Here 1_A denotes the characteristic function of a set A.

Applying the Hölder inequality we conclude that

$$\begin{aligned}
&(N/\ln N)^{1/2} \cdot \big(E\big((\check{\mathcal{A}}_N^q/\delta_N^{q/2}) \cdot \mathcal{M}_q(\delta_N)\big)\big)^{1/q} \\
&\quad \geq \big(E\big((\delta_N/\ln\delta_N)\cdot 1_{\{\delta_N>e^2\}}\big)\big)^{1/2} \cdot \big(E\big((\check{\mathcal{A}}_N^q/\delta_N^{q/2}) \cdot \mathcal{M}_q(\delta_N)\big)\big)^{1/q} \\
&\quad = \Big(\big(E\big((\delta_N/\ln\delta_N)\cdot 1_{\{\delta_N>e^2\}}\big)\big)^{q/q+2} \\
&\qquad\quad \times \big(E\big((\check{\mathcal{A}}_N^q/\delta_N^{q/2}) \cdot \mathcal{M}_q(\delta_N)\big)\big)^{2/q+2}\Big)^{(q+2)/2q} \\
&\quad \geq \Big(E\big(\check{\mathcal{A}}_N^{2q/q+2} \cdot (1/\ln\delta_N)^{q/q+2} \cdot \big(\mathcal{M}_q(\delta_N)\big)^{2/q+2} \cdot 1_{\{\delta_N>e^2\}}\big)\Big)^{(q+2)/2q}
\end{aligned}$$

$$\geq \Big(E\Big(\check{\mathcal{A}}_N^{2q/q+2} \cdot \big((1/\ln\delta_N)^{1/2} \cdot \big(\mathcal{M}_q(\delta_N)\big)^{1/q}\big)^{2q/(q+2)} \\ \times 1_{\{\delta_N > e^2, \|\sigma^*\|_\infty > 0\}}\Big)\Big)^{(q+2)/2q}.$$

Furthermore, with probability 1,

$$\check{\mathcal{A}}_N^{2q/q+2} \cdot \Big((1/\ln\delta_N)^{1/2} \cdot \big(\mathcal{M}_q(\delta_N)\big)^{1/q}\Big)^{2q/(q+2)} \cdot 1_{\{\delta_N > e^2, \|\sigma^*\|_\infty > 0\}} \\ \xrightarrow{N} (\|\sigma^*\|_2/\sqrt{2})^{2q/(q+2)}$$

by (17), (18) and Corollary 2.

Now use Fatou's Lemma to complete the proof. □

LEMMA 7. *If $\overline{X}_N \in \mathbb{X}_N^*$ for every N then*

$$\liminf_{N\to\infty} (N/\ln N)^{1/2} \cdot \Big(E\big((\check{\mathcal{A}}_N^q/\delta_N^{q/2}) \cdot \mathcal{M}_q(\delta_N)\big)\Big)^{1/q} \geq C_q^*/\sqrt{2}.$$

PROOF. Note that $\delta_N \leq N$ by definition of $\mathbb{X}_N^*$. Thus

$$(N/\ln N)^{1/2} \cdot \big(E\big((\check{\mathcal{A}}_N^q/\delta_N^{q/2}) \cdot \mathcal{M}_q(\delta_N)\big)\big)^{1/q} \\ \geq \Big(E\big((\delta_N/\ln\delta_N)^{q/2} \cdot (\check{\mathcal{A}}_N^q/\delta_N^{q/2}) \cdot \mathcal{M}_q(\delta_N) \cdot 1_{\{\delta_N > e^2, \|\sigma^*\|_\infty > 0\}}\big)\Big)^{1/q} \\ = \Big(E\big(\check{\mathcal{A}}_N^q \cdot \big((1/\delta_N)^{1/2} \cdot \big(\mathcal{M}_q(\delta_N)\big)^{1/q}\big)^q \cdot 1_{\{\delta_N > e^2, \|\sigma^*\|_\infty > 0\}}\big)\Big)^{1/q},$$

and the result follows as above in the proof of Lemma 6. □

Combine Lemma 5 with Lemmas 6–7 and observe

$$\lim_{N\to\infty} (N/\ln N)^{1/2} \cdot \Delta_{N,\max}^{1/2} = \lim_{N\to\infty} \big(N/(k_N \cdot \ln N)\big)^{1/2} = 0,$$

to obtain the lower bounds in Theorem 3. Moreover, if N is chosen appropriately, these lower bounds yield the lower bounds in Theorem 1.

Finally, we turn to the piecewise linear interpolated Euler Scheme $\widehat{X}_N^{\mathrm{e}}$. Recall that $\widehat{X}_N^{\mathrm{e}}$ evaluates every Brownian trajectory at the equidistant knots

$$t_{N,\ell} = \ell/N, \qquad \ell = 0, \ldots, N, \tag{19}$$

only.

LEMMA 8. *The equidistant Euler Scheme $\widehat{X}_N^{\mathrm{e}}$ satisfies*

$$\liminf_{N\to\infty} (N/\ln N)^{1/2} \cdot e_q(\widehat{X}_N^{\mathrm{e}}) \geq C_q^{\mathrm{e}}/\sqrt{2}.$$

PROOF. By Theorem 4 we obtain

$$\limsup_{N\to\infty} (N/\ln N)^{1/2} \cdot (E\,\|X - \check{X}_N\|_\infty^q)^{1/q} \le c \cdot \limsup_{N\to\infty} 1/(\ln N)^{1/2} = 0 \tag{20}$$

for the approximation $\check{X}_N$ corresponding to (19).

From (11) and (12) we get

$$\begin{aligned} &E\big(\|\check{X}_N - \widehat{X}_N^{\mathrm{e}}\|_\infty^q \,|\, \Psi_N\big) \\ &\quad \ge E\Big(\max_{0\le\ell\le N-1} \sup_{t_{N,\ell}\le t\le t_{N,\ell+1}} \big|\sigma^*(\check{U}_{N,\ell}) \cdot \big(W_1(t) - E\big(W_1(t)|\Psi_N\big)\big)\big|^q \,\big|\Psi_N\Big). \end{aligned}$$

Note that $W_1 - E(W_1|\Psi_N) = W_1 - E(W_1|W_1(t_{N,1}), \dots, W_1(t_{N,N}))$ consists of independent Brownian bridges corresponding to each subinterval $[t_{N,\ell}, t_{N,\ell+1}]$. Renormalize each subinterval to [0, 1] to obtain

$$E(\|\check{X}_N - \widehat{X}_N^{\mathrm{e}}\|_\infty^q|\Psi_N) \ge (1/N^{q/2}) \cdot \mathcal{M}_q\big(\sigma^*(\check{U}_{N,0}), \dots, \sigma^*(\check{U}_{N,N-1})\big).$$

As above we may assume that (15) holds with probability 1. Thus

$$P\Big(\lim_{N\to\infty} \max_{0\le\ell\le N-1} \sigma^*(\check{U}_{N,\ell}) = \|\sigma^*\|_\infty\Big) = 1$$

and

$$P\Big(\liminf_{N\to\infty} (1/N) \cdot \#\big\{\ell : \sigma^*(\check{U}_{N,\ell}) \ge \|\sigma^*\|_\infty - \varepsilon\big\} > 0\Big) = 1$$

for every $\varepsilon > 0$.

Hence, by Corollary 2,

$$P\Big(\lim_{N\to\infty} (\ln N)^{-q/2} \cdot \mathcal{M}_q\big(\sigma^*(\check{U}_{N,0}), \dots, \sigma^*(\check{U}_{N,N-1})\big) = \|\sigma^*\|_\infty^q \cdot 2^{-q/2}\Big) = 1.$$

Fatou's Lemma yields

$$\begin{aligned} &\liminf_{N\to\infty} (N/\ln N)^{1/2} \cdot (E\,\|\check{X}_N - \widehat{X}_N^{\mathrm{e}}\|_\infty^q)^{1/q} \\ &\qquad \ge \liminf_{N\to\infty} (1/\ln N)^{1/2} \cdot \big(E\,\mathcal{M}_q\big(\sigma^*(\check{U}_{N,0}), \dots, \sigma^*(\check{U}_{N,N-1})\big)\big)^{1/q} \\ &\qquad \ge C_q^{\mathrm{e}}/\sqrt{2}. \end{aligned} \tag{21}$$

Finally, combine (20) and (21) to complete the proof. □

6.3. *Proof of the upper bounds in Theorems* 1 *and* 2 *and Remark* 2. Fix $\tilde{q} \in [1, q]$ and consider a method

$$\overline{X}_N \in \{\widehat{X}_{\tilde{q},N}^{**}, \widehat{X}_N^*, \widehat{X}_N^{\mathrm{e}}\}.$$

Recall that $\overline{X}_N$ uses the knots (14) if $\overline{X}_N \in \{\widehat{X}^{**}_{\tilde{q},N}, \widehat{X}^*_N\}$ and $\widehat{X}^{\mathrm{e}}_N$ is based on the discretization (19) only. Put

$$m_N = \begin{cases} k_N & \text{if } \overline{X}_N \in \{\widehat{X}^{**}_{\tilde{q},N}, \widehat{X}^*_N\}, \\ N & \text{if } \overline{X}_N = \widehat{X}^{\mathrm{e}}_N, \end{cases}$$

let

$$0 = t_{N,0} < \cdots < t_{N,m_N} = 1$$

denote the discretization corresponding to the above cases, and let $\check{X}_N$ denote the respective approximation (5).

Fix $\ell \in \{0, \ldots, m_N - 1\}$ and recall from Section 3 that

$$r_\ell = \begin{cases} \mu_{\tilde{q},\ell} + 2 & \text{if } \overline{X}_N = \widehat{X}^{**}_{\tilde{q},N}, \\ \mu_\ell + 2 & \text{if } \overline{X}_N = \widehat{X}^*_N, \\ 2 & \text{if } \overline{X}_N = \widehat{X}^{\mathrm{e}}_N, \end{cases}$$

is the number of knots

$$t_{N,\ell} = \tau_{\ell,0} < \tau_{\ell,1} < \cdots < \tau_{\ell,r_\ell} < \tau_{\ell,r_\ell+1} = t_{N,\ell+1}$$

used by $\overline{X}_N$ in $[t_{N,\ell}, t_{N,\ell+1}]$. Let $\widetilde{W}_{N,\ell}$ denote the piecewise linear interpolation of $W - W(t_{N,\ell})$ at these knots. Then

$$\check{X}_N(t) - \overline{X}_N(t) = \sigma(\check{U}_\ell) \cdot \big(W(t) - W(t_{N,\ell}) - \widetilde{W}_{N,\ell}(t)\big) \tag{22}$$

for $t \in [t_{N,\ell}, t_{N,\ell+1}]$.

Put

$$H_N = \big(X(0), W(t_{N,1}), \ldots, W(t_{N,m_N})\big)'$$

and note that $\check{U}_\ell$, $\check{\mathscr{A}}_N$, r_ℓ and the knots $\tau_{\ell,r}$ are measurable functions of H_N. Moreover, conditioned on H_N, the processes $W - W(t_{N,\ell}) - \widetilde{W}_{N,\ell}$ consist of independent m-dimensional Brownian bridges corresponding to each subinterval $[\tau_{\ell,r}, \tau_{\ell,r+1}]$.

By definition of $\overline{X}_N$, all subintervals $[\tau_{\ell,r}, \tau_{\ell,r+1}]$ are of the same length with

$$\begin{aligned} &\left(\sum_{j=1}^m \sigma^2_{i,j}(\check{U}_\ell)\right)^{1/2} \cdot (\tau_{\ell,r+1} - \tau_{\ell,r})^{1/2} \\ &\quad \le \sigma^*(\check{U}_\ell)/\big(k_N \cdot (r_\ell - 1)\big)^{1/2} \\ &\quad \le \begin{cases} \check{\mathscr{A}}_N^{2/(\tilde{q}+2)}/N^{1/2} & \text{if } \overline{X}_N = \widehat{X}^{**}_{\tilde{q},N}, \\ \check{\mathscr{A}}_N/(N - k_N)^{1/2} & \text{if } \overline{X}_N = \widehat{X}^*_N, \end{cases} \end{aligned} \tag{23}$$

and

$$\left(\sum_{j=1}^{m} \sigma_{i,j}^2(\check{U}_\ell)\right)^{1/2} \cdot (\tau_{\ell,r+1} - \tau_{\ell,r})^{1/2} \le \sigma^*(\check{U}_\ell)/N^{1/2} \tag{24}$$

if $\overline{X}_N = \widehat{X}_N^{\mathrm{e}}$.

Let ν_N denote the total number of knots used by $\overline{X}_N$ and put

$$\beta_N = \begin{cases} \check{\mathcal{A}}_N^{2/(\tilde{q}+2)}/N^{1/2} & \text{if } \overline{X}_N = \widehat{X}_{\tilde{q},N}^{**}, \\ \check{\mathcal{A}}_N/(N-k_N)^{1/2} & \text{if } \overline{X}_N = \widehat{X}_N^{*}, \\ \max\limits_{0\le\ell\le N-1} \sigma^*(\check{U}_\ell)/N^{1/2} & \text{if } \overline{X}_N = \widehat{X}_N^{\mathrm{e}}. \end{cases}$$

Observe (23) and (24), and apply Lemma 3 with $\beta_{i,j,(\ell,r)} = \sigma_{i,j}(\check{U}_\ell)$, $\Delta_{\ell,r} = \tau_{\ell,r+1} - \tau_{\ell,r}$ and $z = \beta_N^q \cdot (\ln \nu_N)^{q/2}/2^{q/2}$ to obtain

$$\begin{aligned} &E(\|\check{X}_N - \overline{X}_N\|_\infty^q | H_N) \\ &\quad \le \beta_N^q \cdot (\ln \nu_N)^{q/2} \cdot \left(2^{-q/2} + d \cdot \int_{2^{-q/2}}^{\infty} \mathcal{G}_q\big(u \cdot (\ln \nu_N)^{q/2}; \nu_N\big)\, du\right). \end{aligned} \tag{25}$$

We separately analyze the methods $\widehat{X}_{\tilde{q},N}^{**}$, $\widehat{X}_N^{*}$ and $\widehat{X}_N^{\mathrm{e}}$.

LEMMA 9. *The sequence* $\widehat{X}_{\tilde{q},N}^{**}$ *satisfies*

$$\begin{aligned} &\limsup_{N\to\infty} (E\,\nu_N / \ln E\,\nu_N)^{1/2} \cdot e_q(\widehat{X}_{\tilde{q},N}^{**}) \\ &\qquad \le \big(E\,\|\sigma^*\|_2^{2\tilde{q}/(\tilde{q}+2)}\big)^{1/2} \cdot \big(E\,\|\sigma^*\|_2^{2q/(\tilde{q}+2)}\big)^{1/q}/\sqrt{2}. \end{aligned}$$

PROOF. Fix $p \in [0, q]$. Similarly to Section 6.2 we may assume that $\lim_{N\to\infty} \check{\mathcal{A}}_N = \|\sigma^*\|_2$ with probability 1. Hence

$$\liminf_{N\to\infty} E\,\check{\mathcal{A}}_N^p \ge E\,\|\sigma^*\|_2^p$$

by Fatou's Lemma.

Furthermore,

$$\lim_{N\to\infty} E\,\mathcal{A}_N^p = E\,\|\sigma^*\|_2^p$$

by (16), (8) and (B).

Hence,

$$\lim_{N\to\infty} E\,\check{\mathcal{A}}_N^p = E\,\|\sigma^*\|_2^p \tag{26}$$

follows from (7).

Clearly, we may assume $P(\|\sigma^*\|_2 > 0) > 0$. Recall that

$$N \cdot \check{\mathscr{A}}_N^{2\tilde{q}/(\tilde{q}+2)} \leq \nu_N \leq k_N + N \cdot \check{\mathscr{A}}_N^{2\tilde{q}/(\tilde{q}+2)} \tag{27}$$

by definition of $\widehat{X}^{**}_{\tilde{q},N}$.

Hence,

$$\lim_{N\to\infty} (E\nu_N / \ln E\nu_N) \cdot (\ln N/N) = E\|\sigma^*\|_2^{2\tilde{q}/(\tilde{q}+2)} \tag{28}$$

by (26), and therefore

$$\begin{aligned}
&\limsup_{N\to\infty} \bigl(E\nu_N / \ln(E\nu_N)\bigr)^{1/2} \cdot (E\|X - \check{X}_N\|^q)^{1/q} \\
&\quad \leq c \cdot \limsup_{N\to\infty} \bigl(N/(k_N \cdot \ln N)\bigr)^{1/2} = 0
\end{aligned}$$

by Theorem 4.

Put

$$I_N = d \cdot \int_{2^{-q/2}} \mathscr{G}_q\bigl(u \cdot (\ln \nu_N)^{q/2}; \nu_N\bigr)\, du$$

and let

$$V_N = (\ln \nu_N / \ln N)^{q/2} \cdot \check{\mathscr{A}}_N^{2q/(\tilde{q}+2)} \cdot (2^{-q/2} + I_N).$$

In view of (28) and (25) it remains to show that

$$\limsup_{N\to\infty} EV_N \leq \bigl(E\|\sigma^*\|_2^{2q/(\tilde{q}+2)}\bigr)/2^{q/2}. \tag{29}$$

From (27) we get

$$\ln \nu_N / \ln N \leq \bigl(\ln N + \ln\bigl(1 + \check{\mathscr{A}}_N^{2\tilde{q}/(\tilde{q}+2)}\bigr)\bigr)/\ln N \leq 1 + \check{\mathscr{A}}_N^{2\tilde{q}/(\tilde{q}+2)}/\ln N.$$

Moreover, by Lemma 2,

$$I_N \leq c \tag{30}$$

and

$$P\Bigl(\lim_{N\to\infty} I_N = 0\Bigr) = 1. \tag{31}$$

Hence

$$\begin{aligned}
V_N &\leq (2^{-q/2} + I_N) \cdot \bigl(1 + \check{\mathscr{A}}_N^{2\tilde{q}/(\tilde{q}+2)}/\ln N\bigr)^{q/2} \cdot \check{\mathscr{A}}_N^{2q/(\tilde{q}+2)} \\
&\leq (2^{-q/2} + I_N) \cdot \bigl(\check{\mathscr{A}}_N^{2/(\tilde{q}+2)} + \check{\mathscr{A}}_N/(\ln N)^{1/2}\bigr)^q \\
&\leq \bigl(\check{\mathscr{A}}_N^{2/(\tilde{q}+2)}/\sqrt{2} + c \cdot \check{\mathscr{A}}_N/(\ln N)^{1/2} + I_N^{1/q} \cdot \check{\mathscr{A}}_N^{2/(\tilde{q}+2)}\bigr)^q,
\end{aligned}$$

and therefore

$$(32)\qquad \begin{aligned}(E\,V_N)^{1/q} &\le \big(E\,\check{\mathcal{A}}_N^{2q/(\tilde q+2)}\big)^{1/q}/\sqrt{2} + \big(c/(\ln N)^{1/2}\big)\cdot (E\,\check{\mathcal{A}}_N^{q})^{1/q}\\ &\quad + (E\,I_N^3)^{1/3q}\cdot \big(E\,\check{\mathcal{A}}_N^{3q/(\tilde q+2)}\big)^{2/3q}\end{aligned}$$

by the Hölder inequality.

Note that

$$\lim_{N\to\infty} E I_N^3 = 0,$$

due to (30) and (31), and

$$2q/(\tilde q+2) \le 3q/(\tilde q+2) \le q.$$

Thus (29) follows from (32) and (26). □

For $\tilde q = q$ we have

$$\big(E\,\|\sigma^*\|_2^{2\tilde q/(\tilde q+2)}\big)^{1/2}\cdot \big(E\,\|\sigma^*\|_2^{2q/(\tilde q+2)}\big)^{1/q} = C_q^{**}$$

such that Lemma 9 yields in particular the upper bound for the sequence $\widehat{X}_{q,N}^{**}$ in Theorem 1.

LEMMA 10. *The sequence $\widehat{X}_N^*$ satisfies*

$$\limsup_{N\to\infty} (N/\ln N)^{1/2}\cdot e_q(\widehat{X}_N^*) \le C_q^*/\sqrt{2}.$$

PROOF. Note that $N - k_N \le \nu_N \le N$ by definition of $\widehat{X}_N^*$. Consequently,

$$\begin{aligned}&(N/\ln N)^{q/2}\cdot E\big(\|\check{X}_N - \widehat{X}_N^*\|^q \,|\, H_N\big)\\ &\qquad \le \big(N/(N-k_N)\big)^{q/2}\cdot \check{\mathcal{A}}_N^q \cdot \left(2^{-q/2} + \int_{2^{-q/2}} \mathcal{G}_q\big(u\cdot(\ln N)^{q/2}; N\big)\,du\right)\end{aligned}$$

by (25).

Lemma 2 and (26) yield

$$\limsup_{N\to\infty} (N/\ln N)^{1/2}\cdot \big(E\,\|\check{X}_N - \widehat{X}_N^*\|^q\big)^{1/q} \le \limsup_{N\to\infty} (E\check{\mathcal{A}}_N^q)^{1/q}/\sqrt{2} = C_q^*/\sqrt{2}.$$

Finally, use Theorem 4 to complete the proof. □

LEMMA 11. *The sequence $\widehat{X}_N^{\mathrm{e}}$ satisfies*

$$\limsup_{N\to\infty} (N/\ln N)^{1/2}\cdot e_q(\widehat{X}_N^{\mathrm{e}}) \le C_q^{\mathrm{e}}/\sqrt{2}.$$

PROOF. By definition of $\widehat{X}_N^{\mathrm{e}}$ we have $\nu_N = N$. Hence (25), Lemma 2 and (6) imply

$$\begin{aligned}
&(N/\ln N)^{1/2} \cdot \bigl(E\,(\|\check{X}_N - \widehat{X}_N^{\mathrm{e}}\|^q)\bigr)^{1/q} \\
&\quad\le \Bigl(E\Bigl(\Bigl(\max_{0\le\ell\le N-1} \sigma^*(\check{U}_{N,\ell})\Bigr)^q \cdot \bigl(2^{-q/2} + I_N\bigr)\Bigr)\Bigr)^{1/q} \\
&\quad\le \Bigl(E\Bigl(\Bigl(\max_{0\le\ell\le N-1} \sigma^*(U_{N,\ell})\Bigr)^q \cdot \bigl(2^{-q/2} + I_N\bigr)\Bigr)\Bigr)^{1/q} \\
&\qquad + c \cdot \Bigl(E\Bigl(\max_{0\le\ell\le N-1} |\sigma^*(\check{U}_{N,\ell}) - \sigma^*(U_{N,\ell})|^q\Bigr)\Bigr)^{1/q} \\
&\quad\le \Bigl(E\Bigl(\Bigl(\max_{0\le\ell\le N-1} \sigma^*(U_{N,\ell})\Bigr)^q \cdot \bigl(2^{-q/2} + I_N\bigr)\Bigr)\Bigr)^{1/q} + c/N^{1/2}.
\end{aligned}$$

Due to Lemma 2 and (A) we have

$$P\Bigl(\lim_{N\to\infty}\Bigl(\max_{0\le\ell\le N-1} \sigma^*(U_{N,\ell})\Bigr)^q \cdot \bigl(2^{-q/2} + I_N\bigr) = \|\sigma^*\|_\infty^q \cdot 2^{-q/2}\Bigr) = 1$$

and

$$\Bigl(\max_{0\le\ell\le N-1} \sigma^*(U_{N,\ell})\Bigr)^q \cdot \bigl(2^{-q/2} + I_N\bigr) \le c \cdot (1 + \|X\|_\infty^q).$$

Hence,

$$\limsup_{N\to\infty} (N/\ln N)^{1/2} \cdot \bigl(E\,\|\check{X}_N - \widehat{X}_N^{\mathrm{e}}\|^q\bigr)^{1/q} \le (E\,\|\sigma^*\|_\infty^q)^{1/q}/\sqrt{2} = (C_q^{\mathrm{e}})/\sqrt{2}.$$

by (B) and Fatou's Lemma.

Finally, observe Theorem 4 to obtain the desired result. □

APPENDIX

Consider the geometric Brownian motion

$$X(t) = \exp\bigl(-b^2/2 \cdot t + b \cdot W(t)\bigr), \qquad t \in [0, 1],$$

with zero drift and volatility $b \in \mathbb{R}$. Here W denotes a one-dimensional Brownian motion. We provide an exact formula for the pth moment

$$E\,\|X\|_\infty^p, \qquad p > 0,$$

of the maximum of X in terms of the distribution function Φ and the density φ of the standard normal distribution.

LEMMA 12. *For $\alpha, \beta \in \mathbb{R}$ it holds that*

$$\int_0^\infty e^{\alpha\cdot u} \cdot \Phi(-u + \beta)\,du = \begin{cases} (1/\alpha) \cdot \bigl(e^{\alpha\cdot\beta + \alpha^2/2} \cdot \Phi(\alpha + \beta) - \Phi(\beta)\bigr), & \text{if } \alpha \ne 0 \\ \beta \cdot \Phi(\beta) + \varphi(\beta), & \text{if } \alpha = 0. \end{cases}$$

PROOF. Let $\alpha \neq 0$. By partial integration,

$$\int_0^\infty e^{\alpha \cdot u} \cdot \Phi(-u+\beta)\, du = (1/\alpha) \cdot \biggl(\int_0^\infty e^{\alpha \cdot u} \cdot \varphi(-u+\beta)\, du - \Phi(\beta) \biggr).$$

Clearly,

$$\begin{aligned} \int_0^\infty e^{\alpha \cdot u} \cdot \varphi(-u+\beta)\, du &= e^{\alpha \cdot \beta + \alpha^2/2} \cdot \int_0^\infty \varphi\bigl(u - (\alpha+\beta)\bigr)\, du \\ &= e^{\alpha \cdot \beta + \alpha^2/2} \cdot \Phi(\alpha+\beta), \end{aligned}$$

which yields the desired result.

Next, assume $\alpha = 0$. By partial integration,

$$\begin{aligned} \int_0^\infty \Phi(-u+\beta)\, du &= \int_0^\infty u \cdot \phi(u-\beta)\, du \\ &= \beta \cdot \Phi(\beta) - \int_0^\infty \varphi'(u-\beta)\, du \\ &= \beta \cdot \Phi(\beta) + \varphi(\beta), \end{aligned}$$

which completes the proof. $\square$

THEOREM 5. *For every $p > 0$ it holds that*

$$\begin{aligned} &E\,\|X\|_\infty^p \\ &\quad = \begin{cases} 1/(p-1) \cdot \bigl((2p-1) \cdot e^{p(p-1)\cdot b^2/2} \cdot \Phi\bigl((p-1/2)\cdot|b|\bigr) - \Phi(|b|/2)\bigr), & \\ \hfill \textit{if } p \neq 1, & \\ (2+b^2/2) \cdot \Phi(|b|/2) + |b| \cdot \varphi(b/2), \quad \textit{if } p = 1. & \end{cases} \end{aligned}$$

PROOF. Clearly, we may assume $b > 0$. Then

$$\begin{aligned} E\,\|X\|_\infty^p &= 1 + \int_1^\infty P\Bigl(\sup_{0 \le t \le 1} \exp\bigl(p \cdot (-b^2/2 \cdot t + b \cdot W(t))\bigr) > u\Bigr)\, du \\ &= 1 + \int_0^\infty e^u \cdot P\Bigl(\sup_{0 \le t \le 1} \bigl(W(t) - b/2 \cdot t\bigr) > u/(p \cdot b)\Bigr)\, du. \end{aligned}$$

For $\alpha \in \mathbb{R}$ and $x \geq 0$ we have

$$P\Bigl(\sup_{0 \le t \le 1} \bigl(W(t) + \alpha \cdot t\bigr) > x\Bigr) = \Phi(\alpha - x) + e^{2 \cdot \alpha \cdot x} \cdot \Phi(-\alpha - x),$$

see Karatzas and Shreve [(1999), page 368]. Hence

$$\begin{aligned} E\,\|X\|_\infty^p = 1 + p \cdot b \cdot \biggl(&\int_0^\infty e^{p \cdot b \cdot u} \cdot \Phi(-u - b/2)\, du \\ &+ \int_0^\infty e^{(p-1)\cdot b \cdot u} \cdot \Phi(-u + b/2)\, du \biggr). \end{aligned}$$

Now, evaluate both integrals by applying Lemma 12 with $\alpha = p \cdot b$, $\beta = -b/2$ and $\alpha = (p-1) \cdot b$, $\beta = b/2$, respectively. $\square$

Acknowledgments. I am grateful to Klaus Ritter and Norbert Hofmann for very stimulating discussions and a number of helpful comments on the presentation of the paper.

REFERENCES

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

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

HOFMANN, N., MÜLLER-GRONBACH, T. and RITTER, K. (2000a). Optimal approximation of stochastic differential equations by adaptive step-size control. *Math. Comp.* **69** 1017–1034.

HOFMANN, N., MÜLLER-GRONBACH, T. and RITTER, K. (2000b). Step-size control for the uniform approximation of systems of stochastic differential equations with additive noise. *Ann. Appl. Probab.* **10** 616–633.

HOFMANN, N., MÜLLER-GRONBACH, T. and RITTER, K. (2000c). The optimal discretization of stochastic differential equations. *J. Complexity* **17** 117–153.

KARATZAS, I. and SHREVE, S. E. (1999). *Methods of Mathematical Finance.* Springer, Berlin.

KLOEDEN, P. and PLATEN, E. (1995). *Numerical Solution of Stochastic Differential Equations.* Springer, Berlin.

MAIOROV, V. E. and WASILKOWSKI, G. W. (1996). Probabilistic and average linear widths in L_∞-norm with respect to r-fold Wiener measure. *J. Approx. Theory* **84** 31–40.

MILSTEIN, G. N. (1995). *Numerical Integration of Stochastic Differential Equations.* Kluwer, Dordrecht.

MÜLLER-GRONBACH, T. and RITTER, K. (1997). Uniform reconstruction of Gaussian processes. *Stochastic Processes Appl.* **69** 55–70.

MÜLLER-GRONBACH, T. and RITTER, K. (1998). Spatial adaption for predicting random functions. *Ann. Statist.* **26** 2264–2288.

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

TALAY, D. (1995). Simulation of stochastic differential systems. *Probabilistic Methods in Applied Physics. Lecture Notes in Phys.* **451** 54–96. Springer, Berlin.

TRAUB, J. F., WASILKOWSKI, G. W. and WOŹNIAKOWSKI, H. (1988). *Information-Based Complexity.* Academic Press, New York.

FACHBEREICH MATHEMATIK
TECHNISCHE UNIVERSITÄT DARMSTADT
SCHLOSSGARTENSTRASSE 7
64289 DARMSTADT
GERMANY
E-MAIL: gronbach@mathematik.tu-darmstadt.de