OPTIMAL APPROXIMATION OF STOCHASTIC DIFFERENTIAL EQUATIONS BY ADAPTIVE STEP-SIZE CONTROL

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

Abstract. We study the pathwise (strong) approximation of scalar stochastic differential equations with respect to the global error in the L_2-norm. We introduce an adaptive step-size control for the Euler scheme. For equations with additive noise this method is asymptotically optimal in the class of arbitrary methods that use a fixed number of observations of the driving Brownian motion. The superiority of the new method is confirmed in simulation experiments for equations with additive noise as well as for general scalar equations.

1. Introduction

We consider pathwise approximation for scalar stochastic differential equations

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

on the unit interval $T = [0, 1]$ with a one-dimensional Brownian motion W. We study methods that yield processes $\overline{X}$ whose paths are close to the respective paths of the strong solution X of (1).

Different notions of errors for pathwise approximation are studied in the literature. The majority of results deals with mean square errors $E(X(\tau_k) - \overline{X}(\tau_k))^2$ at discrete points $\tau_k \in T$. See Kloeden and Platen (1995), Milstein (1995), and Talay (1995) for results and references. In this paper the pathwise distance between X and $\overline{X}$ is analyzed globally on T in the L_2-norm $\|\cdot\|_2$, and the error of $\overline{X}$ is defined by

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

See Talay (1995) for a deterministic notion of error, which can be used for a worst case analysis with respect to the Brownian paths.

We wish to determine (asymptotically) optimal approximation methods. To this end we consider arbitrary methods $\overline{X}_n$ that use the values of a path of W at n points. These points may be selected sequentially, where the only restriction is measurability of the respective selection functions. Moreover, a finite number of function values (or derivative values) of the drift a and the diffusion coefficient σ may be used. The number $\inf_{\overline{X}_n} e(\overline{X}_n)$ is the minimal error that can be achieved by methods of the above type.

Date: May 25, 1998.

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

We wish to find sharp upper and lower bounds for the minimal error. Upper bounds may be derived by the analysis of a specific method, while lower bounds must hold for every method $\overline{X}_n$.

We show that $\inf_{\overline{X}_n} e(\overline{X}_n)$ is of order

$$1/\sqrt{6} \cdot \|\sigma\|_1 \cdot n^{-1/2}$$

if σ is independent of the state variable. The corresponding equations are sometimes called equations with additive noise. In particular, higher order methods do not exist if the global error on T is analyzed. See Section 4 for further discussion. For equations with additive noise the optimal order is achieved by an Euler scheme with adaptive step-size control. We take the step-size proportionally to the inverse of the current value of σ. Numerical experiments indicate that it is reasonable to use this method also for the general equation (1). We add that the computation time of our method is bounded by $c \cdot n$ with a small constant c, and memory requirements are negligible.

It is common to take equidistant discretizations for the numerical solution of a stochastic differential equation. For the error of an Euler approximation with constant step-size $1/n$ we obtain in the additive noise case the order

$$1/\sqrt{6} \cdot \|\sigma\|_2 \cdot n^{-1/2},$$

which shows that it is not efficient to discretize equidistantly.

Only a few papers deal with adaptive step-size control; a partial list includes Newton (1990), Gaines and Lyons (1997), and Mauthner (1998). However, optimality in the class of all methods $\overline{X}_n$, or at least superiority of adaptive step-size control compared to fixed (equidistant) step-sizes was not addressed so far.

2. Euler Approximation with Adaptive Step-Size Control

Consider a discretization

$$0 = \tau_0 < \cdots < \tau_n = 1 \tag{2}$$

of the unit interval. The corresponding Euler scheme $\widehat{X}$ for equation (1) with initial value $X(0)$ is defined by

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

and

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

where $k = 0, \ldots, n-1$. The global approximation $\widehat{X}$ for X on T is defined by piecewise linear interpolation of the data $(\tau_k, \widehat{X}(\tau_k))$ with $k = 0, \ldots, n$.

It is reasonable to select a discretization that reflects the local properties of the differential equation. We choose a basic step-size

$$h > 0$$

and define the adaptive step-size control for the Euler method $\widehat{X}^h = \widehat{X}$ by $\tau_0 = 0$ and

$$\tau_{k+1} = \tau_k + \min\Big(h^{2/3}, h/\sigma(\tau_k, \widehat{X}^h(\tau_k))\Big), \tag{3}$$

as long as the right-hand side does not exceed one. Otherwise we put $\tau_{k+1} = 1$.

We perform an asymptotic analysis of the error $e(\widehat{X}^h)$ with h tending to zero for equations

$$dX(t) = a(t, X(t))\,dt + \sigma(t)\,dW(t) \tag{4}$$

with additive noise. In this case $\tau_k = \tau_k(h,\sigma)$, and we use $n(h,\sigma)$ to denote the total number of steps, i.e.,

$$n(h,\sigma) = \min\{k \in \mathbb{N} : \tau_k(h,\sigma) = 1\}.$$

We assume that the drift $a : T \times \mathbb{R} \to \mathbb{R}$ and the diffusion coefficient $\sigma : T \to \mathbb{R}$ have the following properties:

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

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

and

$$|a(t,x) - a(s,x)| \le K_3 \cdot (1+|x|) \cdot |t-s|$$

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

(B) The function σ is continuously differentiable and satisfies

$$\sigma(t) > 0$$

for all $t \in T$.

Furthermore we assume that the initial value $X(0)$ satisfies

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

$$E|X(0)|^2 \le K_4$$

for some constant $K_4 > 0$.

These conditions are standard assumptions when analyzing approximations for stochastic differential equations. The only exception is the positivity of σ, which can be replaced by integrability of $\sigma^{-1/2}$. The latter property holds, for instance, if σ only has simple zeros. Given the above properties, a pathwise unique strong solution $X(t)$, $t \in T$, of the equation (4) with initial value $X(0)$ exists.

We use $\|\cdot\|_p$ to denote the L_p-norm of real-valued functions on T. Furthermore we sometimes write $e(\overline{X}, a, \sigma, X(0))$ instead of $e(\overline{X})$.

Theorem 1. *Assume that (A)–(C) hold for equation* (4). *Then*

$$\lim_{h\to 0} n(h,\sigma)^{1/2} \cdot e(\widehat{X}^h, a, \sigma, X(0)) = 1/\sqrt{6} \cdot \|\sigma\|_1$$

for the Euler approximation with discretization (3). *The Euler approximation* $\widehat{X}_n$ *with constant step-size*

$$\tau_{k+1} - \tau_k = 1/n \tag{5}$$

yields

$$\lim_{n\to\infty} n^{1/2} \cdot e(\widehat{X}_n, a, \sigma, X(0)) = 1/\sqrt{6} \cdot \|\sigma\|_2.$$

Hence (5) is not efficient in general: taking (3) instead of (5) reduces the error roughly by the factor $\|\sigma\|_1/\|\sigma\|_2$ for the same number of steps. This fact is already due to the choice of the discretization (5) and does not arise from choosing the Euler method. A proof is easily obtained from the estimates in Section 6. See Section 5 for simulation experiments.

In fact a much stronger optimality property holds for the method from Theorem 1. This method is asymptotically optimal for all equations with additive noise among all methods that use values of W at arbitrary points. See Section 3.

We do not have asymptotic results for our step-size control in the case of arbitrary scalar equations (1). However, simulation experiments indicate that (3) is still superior to an equidistant discretization, see Section 5.

Remark 1. Theorem 1 deals with small values of h. In this case the step-size is $h/\sigma(\tau_k)$, since σ is bounded away from zero. The term $h^{2/3}$ in definition (3) only matters if $h > \sigma^3(\tau_k)$. For small values of $\sigma(\tau_k)$ the equation (4) locally almost becomes an ordinary differential equation. Still a reasonably small step-size is needed to get a good approximation. The particular choice $h^{2/3}$ is motivated by error estimates on intervals $[\tau_k, \tau_{k+1}]$, see Remark 3.

Remark 2. In Theorem 1, as well as in Theorem 2 below, we give error estimates that hold for individual equations. One can easily strengthen Theorem 1 such that the upper bound uniformly holds on the class of equations (4) that is defined by (A) and (C) and a quantitative version of (B).

To this end let $K = (K_1, \ldots, K_7)$ with $K_i > 0$ and let $F(K)$ denote the set of all $(a, \sigma, X(0))$ such that (A) and (C) hold for a and $X(0)$, respectively. For σ we require

$$\text{(B*)} \qquad \|\sigma\|_1 \le K_5, \qquad \|\sigma'\|_\infty \le K_6, \qquad \inf_{t\in T} \sigma(t) \ge K_7.$$

The maximal error of $\widehat{X}^h$ on the class $F(K)$ is defined by

$$e_{\max}(h, K) = \sup_{(a,\sigma,X(0))\in F(K)} e(\widehat{X}^h, a, \sigma, X(0))$$

and the maximal number of steps of this method is defined by

$$n_{\max}(h, K) = \sup_{(a,\sigma,X(0))\in F(K)} n(h, \sigma).$$

We obtain

$$\lim_{h \to 0} n_{\max}(h, K)^{1/2} \cdot e_{\max}(h, K) = 1/\sqrt{6} \cdot K_5 \tag{6}$$

as a straightforward consequence of (16) and Theorem 1.

The estimate (6) is a worst case result on the class of equations corresponding to $F(K)$. We see that for differential equations with additive noise worst case results and results for individual equations do not differ essentially if the number of steps is large. For other problems of numerical analysis matters may be completely different, see, e.g., Traub, Wasilkowski, and Woźniakowski (1988, Chapter 10) and Novak and Ritter (1993).

3. Lower Bounds

The Euler approximation with adaptive step-size control is based on a realization of the initial value $X(0)$ and a realization of the Brownian motion, observed at a finite number of points. Moreover, it uses a finite number of values of the drift a and of the diffusion coefficient σ.

Now we present lower bounds that hold for every method of the above form. We even drop all restrictions on the available information about a and σ, such that, in particular, partial derivatives of smooth functions a and σ may be used. We fix a and σ, and we consider the corresponding equation (4). An arbitrary method $\overline{X}_n$ that is based on a realization of the initial value $X(0)$ and on n observations of a trajectory of W is then defined by measurable mappings

$$\psi_k : \mathbb{R}^k \to T$$

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

$$\phi_n : \mathbb{R}^{n+1} \to L_2(T).$$

The mapping ψ_k determines the observation point in step k in terms of the previous evaluations. A pathwise approximation is computed according to

$$\overline{X}_n = \phi_n(X(0), Y_1, \dots, Y_n),$$

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

$$e(n, a, \sigma, X(0)) = \inf_{\overline{X}_n} e(\overline{X}_n, a, \sigma, X(0))$$

is the minimal error that can be obtained by n-point methods for the equation (4). Suppose that $\psi_1, \dots, \psi_n$ are fixed. Then $e(\overline{X}_n, a, \sigma, X(0))$ is minimized by the conditional mean of X given $X(0)$ and W at the respective discretization. Hence the choice of the discretization is the main problem in a theoretical minimization of errors of n-point methods.

Theorem 2. *Assume that (A), (B), and (C) hold for equation* (4). *Then*

$$\lim_{n\to\infty} n^{1/2} \cdot e(n, a, \sigma, X(0)) = 1/\sqrt{6} \cdot \|\sigma\|_1.$$

Due to the Theorems 1 and 2 the Euler approximation with adaptive step-size control (3) is asymptotically optimal for every equation with additive noise. Theorem 2 remains valid for a larger class of methods, where the number of observations depends on the trajectory of W. More precisely, after every evaluation a decision is permitted, whether to stop or to continue with further observations. Clearly n must be replaced by the expected number of observations.

The number n is a crude measure of the cost of n-point methods, since any computational cost in addition to the evaluations of W is ignored. Observe however, that the Euler method with constant step-size has the least computational cost among all n-point methods that are used in practice. The adaptive step-size control (3) requires only a few additional operations per step.

4. Discussion

We relate our results to known error bounds for pathwise approximation methods. Concerning the specific smoothness assumptions for a and σ we refer to the literature. We stress that some of the known results hold for systems of equations.

4.1. **Global error in L_p-norm.** In our analysis of strong approximation methods the pathwise error is defined globally on the interval T in L_2-norm. The following upper bounds are already known for general equations (1). The Euler approximation $\widehat{X}_n$ with constant step-size $1/n$ satisfies

$$\left(E(\|X - \widehat{X}_n\|_2^2)\right)^{1/2} \leq c \cdot n^{-1/2},$$

see Milstein (1995, Remark 1.2). Moreover, this method also yields

$$\left(E(\|X - \widehat{X}_n\|_\infty^q)\right)^{1/q} \leq c \cdot (\ln n)^{1/2}\, n^{-1/2}$$

for every $1 \leq q < \infty$, see Faure (1992) and Bouleau and Lépingle (1994, Remark 5.B.1.5). The constants $c > 0$ are unspecified in both cases.

Theorems 1 and 2 give sharp bounds for equation (4) in L_2-norm. Results from Müller-Gronbach and Ritter (1996, Remark 4) and the approach from Section 6 lead to the following *conjecture* in the L_∞-norm case: for every equation (4),

$$\lim_{n\to\infty} (n/\ln n)^{1/2} \cdot \inf_{\overline{X}_n} \left(E(\|X - \overline{X}_n\|_\infty^q)\right)^{1/q} = \|\sigma\|_2/\sqrt{2},$$

and an Euler method with step-size $\tau_{k+1} - \tau_k = h/\sigma^2(\tau_k)$ yields the optimal rate.

4.2. **Pointwise error.** Usually errors of strong approximations are defined discretely at a finite number of points in T. Often these points coincide with the discretization of the given method. In this case it is not clear how to compare methods that are based on different discretizations. For simplicity we consider the error only at the right endpoint of T.

The Euler approximation $\widehat{X}_n$ with step-size $1/n$ yields

$$\left(E(X(1)-\widehat{X}_n(1))^2\right)^{1/2} \leq c \cdot n^{-1/2}$$

for general equations (1). A better upper bound is obtained by the Milstein method $\check{X}_n$ with step-size $1/n$. This n-point method satisfies

$$\left(E(X(1)-\check{X}_n(1))^2\right)^{1/2} \leq c \cdot n^{-1}$$

with an unspecified constant $c > 0$, see Milstein (1995, eq. (2.32)). The Milstein method coincides with the Euler method for equations (4).

We get sharp bounds for equation (4) with zero drift. In this case

$$\lim_{n\to\infty} n \cdot \inf_{\overline{X}_n} \left(E(X(1)-\overline{X}_n(1))^2\right)^{1/2} = 1/\sqrt{12} \cdot \left(\int_0^1 (\sigma')^{2/3}(t)\,dt\right)^{3/2}, \qquad (7)$$

which follows from Sacks and Ylvisaker (1970) and Traub, Wasilkowski, and Woźniakowski (1988, Chapter 6.5) because of (11).

Clark and Cameron (1980) analyze n-point methods that are based on the equidistant discretization (5). For an autonomous equation $dX(t) = a(X(t))\,dt + dW(t)$ with additive noise they show

$$\lim_{n\to\infty} n \cdot \inf_{\phi_n} \left(E(X(1)-\phi_n(W(1/n),\ldots,W(1)))^2\right)^{1/2} = c$$

with an explicit constant c, which is positive in most cases. Here ϕ_n varies over all measurable mappings $\mathbb{R}^n \to \mathbb{R}$. For an arbitrary autonomous equation they show that one cannot achieve errors of order $n^{-1-\delta}$ for any $\delta > 0$ using equidistant discretizations. See Newton (1990) for another concept of asymptotic efficiency concerning the choice of ϕ_n for a given discretization.

Sections 4.1 and 4.2 indicate significant differences between global and pointwise errors for stochastic differential equations. In contrast, this difference is not present for ordinary differential equations, due to the global smoothness of the solutions.

4.3. **Using additional information about W.** ¿From Theorem 2 and (7) we get (rather large) lower bounds for arbitrary n-point methods. Faster convergence of errors is sometimes possible for more general methods that use additional information about W.

Instead of Dirac functionals one can apply arbitrary bounded linear functionals on $C(T)$ to the trajectories of the Brownian motion. This happens, for instance, in the order 3/2 strong Taylor scheme of Wagner and Platen (1978), see also Kloeden and Platen (1995, Section 10.4). In addition to the values $W(\tau_i)$ this method uses the integrals $\int_{\tau_i}^{\tau_{i+1}} W(s)\,ds$. A constant step-size $1/n$ yields errors of order $n^{-3/2}$ at the right endpoint of T. We see that bounded linear functionals are more powerful than Dirac functionals, if the pointwise error is studied.

Even more general methods are derived, for instance, from the stochastic Taylor formula by including higher order multiple stochastic integrals. In principle, pointwise errors of order $n^{-\gamma}$ for arbitrary large γ can be achieved. However, the simulation of multiple integrals is a nontrivial task; sometimes these integrals are approximated by bounded linear functionals, applied to W. See Milstein (1995) and Kloeden and Platen (1995). Similar statements are not true for the global error in L_2-norm. All known methods of higher order in the pointwise sense only use bounded linear functionals, if they are applied to equations (4) with zero drift. For these equations the following lower bound is easily derived from the Karhunen-Loève expansion of the Brownian motion. If a method uses n bounded linear functionals, which may be selected sequentially as in Section 3, then its error is bounded from below by $c \cdot n^{-1/2}$. Here $c > 0$ does only depend on σ and $X(0)$.

Moreover, we can apply Maiorov's result on average n-widths of the Wiener space, see Maiorov (1996). Here even complete knowledge of the trajectory of W is allowed. As long as all pathwise approximations are taken from an n-dimensional subspace of $C(T)$, we have a lower bound $c \cdot n^{-1/2}$. Note that this result applies in particular to all methods that first compute approximate solutions at fixed points $\tau_1, \ldots, \tau_n$ and then apply a linear algorithm to these data, no matter how these pointwise approximations are obtained.

5. Simulation Experiments

We compare errors of the Euler method $\widehat{X}^h$ with adaptive step-size control and of the Euler method $\widehat{X}_n$ with equidistant discretization.

5.1. **Two equations with additive noise.** First we consider the equation

$$dX(t) = -2\,dt + (20\exp(-(10t-1)^2) - t)\,dW(t), \qquad X(0) = 1. \tag{8}$$

The assumptions (A) and (C) are trivially satisfied and the diffusion coefficient is smooth with only a simple zero. The solution of (8) is not known explicitly. Instead of X we therefore use an M-point Euler approximation $\widehat{X}_M$ with M sufficiently large. We simulate $K = 5000$ trajectories of the driving Brownian motion and use

$$\varepsilon(\overline{X}) = \left(\frac{1}{K} \cdot \sum_{i=1}^{K} \|\widehat{X}_M(\omega_i) - \overline{X}(\omega_i)\|_2^2\right)^{1/2}$$

as an estimator for $e(\overline{X})$. This quantity is easy to evaluate for $\overline{X} = \widehat{X}_n$ or $\overline{X} = \widehat{X}^h$, since $\widehat{X}^h$, $\widehat{X}_n$, and $\widehat{X}_M$ yield piecewise linear functions.

The number $n(h,\sigma)$ of steps of $\widehat{X}^h$ is deterministic for equation (8). Hence we express the efficiency of the adaptive Euler approximation by the ratio

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

where $n = n(h,\sigma)$. From Theorem 1 we know that $\mathrm{eff}(h)$ is approximately

$$\|\sigma\|_2/\|\sigma\|_1 = 0.525\ldots$$

for sufficiently large K and small h. For every trajectory of W a certain percentage of steps of $\widehat{X}^h$ has length $h^{2/3}$, and we call these cases exceptions. We also show the mean number $\mathrm{ex}(h)$ of the exceptions in the table below.

h	$\varepsilon(\widehat{X}^h)$	$n(h,\sigma)$	$\varepsilon(\widehat{X}_n)$	$\mathrm{eff}(h)$	$\mathrm{ex}(h)$
$1 \cdot 10^{-1}$	0.551628	38	1.050418	0.53	0.0263
$1 \cdot 10^{-2}$	0.107464	368	0.173890	0.62	0.0027
$5 \cdot 10^{-3}$	0.069409	734	0.112526	0.61	0.0013
$2 \cdot 10^{-3}$	0.039607	1832	0.067495	0.58	0.0005
$1 \cdot 10^{-3}$	0.026602	3662	0.046838	0.56	0.0005
$1 \cdot 10^{-4}$	0.007783	36601	0.014526	0.53	0.0000

Simulation for (8) with different basic step-sizes h

We see the superiority of the adaptive step-size control already for small numbers $n(h,\sigma)$ of steps, and not only in an asymptotic sense. Moreover, the efficiency $\mathrm{eff}(h)$ is close to the asymptotic value 0.525.

Next we consider the same scenario as above for the equation

$$dX(t) = (t-0.5)(5-X(t))\,dt + (1.001-t)^4\,dW(t), \qquad X(0) = 1. \tag{9}$$

The assumptions (A)–(C) are obviously satisfied, and $\mathrm{eff}(h)$ is approximately

$$\|\sigma\|_2/\|\sigma\|_1 = 0.60\ldots$$

for sufficiently large K and small h.

h	$\varepsilon(\widehat{X}^h)$	$n(h,\sigma)$	$\varepsilon(\widehat{X}_n)$	$\mathrm{eff}(h)$	$\mathrm{ex}(h)$
$1 \cdot 10^{-2}$	0.048146	33	0.050019	0.96	0.45
$1 \cdot 10^{-3}$	0.009217	248	0.010322	0.89	0.23
$5 \cdot 10^{-4}$	0.005796	471	0.006852	0.85	0.18
$2 \cdot 10^{-4}$	0.003225	1122	0.004168	0.77	0.13
$1 \cdot 10^{-4}$	0.002112	2184	0.002926	0.72	0.10
$1 \cdot 10^{-5}$	0.000592	20761	0.000927	0.64	0.04

Simulation for (9) with different basic step-sizes h

Here the numbers $\mathrm{ex}(h)$ are large, since σ gets small as time approaches one. In fact, this is important to obtain a good performance of $\widehat{X}^h$ for equation (9), see Remark 1. For the same reason $\mathrm{eff}(h)$ tends to its limit $0.60\ldots$ rather slowly.

5.2. **An equation with multiplicative noise.** Now we consider the equation

$$dX(t) = 2X(t)\,dW(t), \qquad X(0) = 1. \tag{10}$$

Its solution is given by

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

We use a sufficiently accurate piecewise linear interpolation $\widetilde{X}$ of X in our simulation experiments and take

$$\varepsilon(\overline{X}) = \left(\frac{1}{K} \cdot \sum_{i=1}^{K} \|\widetilde{X}(\omega_i) - \overline{X}(\omega_i)\|_2^2\right)^{1/2}$$

as an estimator for $e(\overline{X})$ as before.

Due to large fluctuations of the sample paths of X, the empirical variance of the errors $\|\widetilde{X}(\omega_i) - \widehat{X}^h(\omega_i)\|_2^2$ is much larger than the corresponding quantity for (8) or (9). Taking this into account we simulate $K = 600\,000$ trajectories of the driving Brownian motion to calculate $\varepsilon(\widehat{X}^h)$.

The number of steps of $\widehat{X}^h$ is now a random variable, and we use $n(h, \sigma, \omega_i)$ to denote its counterparts in the simulation. The efficiency of the adaptive Euler approximation is again given by the ratio

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

where

$$n = n(h, \sigma) = \left\lceil \frac{1}{K} \cdot \sum_{i=1}^{K} n(h, \sigma, \omega_i) \right\rceil.$$

For the constant step-size methods the empirical variances of the errors are even larger. Therefore we use $K = 2\,500\,000$ and $K = 7\,500\,000$ trajectories to calculate $\varepsilon(\widehat{X}_n)$ for $n = 2009$ and $n = 204$, respectively. Again we observe superiority of the adaptive step-size control.

h	$\varepsilon(\widehat{X}^h)$	$n(h,\sigma)$	$\varepsilon(\widehat{X}_n)$	$\mathrm{eff}(h)$	$\mathrm{ex}(h)$
$1 \cdot 10^{-2}$	0.359392	203.5	0.626733	0.57	0.085
$1 \cdot 10^{-3}$	0.112864	2008.8	0.192832	0.59	0.026

Simulation for (10) with different basic step-sizes h

For the particular equation (10) we were able to analyze the asymptotic behavior of the error $\varepsilon(\widehat{X}_n)$ in the case of constant step-size $1/n$. We have

$$\lim_{n\to\infty} n^{1/2} \cdot \varepsilon(\widehat{X}_n) = 1/\sqrt{3} \cdot (5 \cdot e^4 + 1) = 9.56\ldots,$$

which yields approximate values 0.66 and 0.21 for $\varepsilon(\widehat{X}_{204})$ and $\varepsilon(\widehat{X}_{2009})$, respectively. This is in good accordance with the simulated errors above.

For the equation (10) the empirical standard deviation of $n(h, \sigma, \cdot)$ is huge, namely 469 for $h = 10^{-2}$ and 4909 for $h = 10^{-3}$. 'Hard' trajectories are detected automatically and accordingly a large number of knots is chosen. For instance, if $h = 10^{-2}$ then the number of knots ranges between 24 and 123 162 for the 600 000 trajectories simulated.

6. Proofs

For the solution X of equation (4) with additive noise we have

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

with

$$V(t) = \sigma(t) \cdot W(t),$$

$$Y(t) = \int_0^t a(s, X(s))\, ds,$$

and

$$Z(t) = \int_0^t \sigma'(s) \cdot W(s)\, ds.$$

For every discretization (2) the Euler approximation $\widehat{X}$ to X may be written as

$$\widehat{X}(\tau_k) = X(0) + \widehat{V}(\tau_k) + \widehat{Y}(\tau_k) - \widehat{Z}(\tau_k)$$

with

$$\widehat{V}(\tau_k) = \sigma(\tau_k) \cdot W(\tau_k),$$

$$\widehat{Y}(\tau_k) = \sum_{i=1}^{k} a(\tau_{i-1}, \widehat{X}(\tau_{i-1})) \cdot (\tau_i - \tau_{i-1}), \tag{12}$$

and

$$\widehat{Z}(\tau_k) = \sum_{i=1}^{k} (\sigma(\tau_i) - \sigma(\tau_{i-1})) \cdot W(\tau_i). \tag{13}$$

Piecewise linear interpolation yields processes $\widehat{V}$, $\widehat{Y}$, and $\widehat{Z}$ over T.

An outline of the proofs reads as follows. Under the assumptions (A)–(C) the processes Y and Z are smooth compared to the process V. Therefore the error of the Euler approximation is essentially determined by $V - \widehat{V}$. Moreover, the following holds for an arbitrary method that is based on n observations of W. Its error is roughly as large as the error of the best method for approximation of V on the basis of n observations of V.

The latter problem deals with approximation of a stochastic process using observations of the same process on a finite number of points. Problems of this kind are well studied, see, e.g., Ritter (1998) for results and references. In particular, n observations of V at suitably chosen points and piecewise linear interpolation roughly yield the error

$$1/\sqrt{6} \cdot \|\sigma\|_1 \cdot n^{-1/2},$$

and this is best possible, see Su and Cambanis (1993) and Müller-Gronbach (1996).

6.1. **Preliminary estimates.** In the sequel we use c to denote unspecified positive constants, which only depend on the constants K_i from conditions (A), (B*), and (C). At first we determine the smoothness of V, Y, and Z.

Lemma 1. *Let*

$$\Gamma = E\Big((V(t_2) - V(t_1)) \cdot (V(t_4) - V(t_3))\Big)$$

and

$$\Delta = \max(t_2 - t_1, t_4 - t_3),$$

where $t_1 \leq t_2$ and $t_3 \leq t_4$. Then

$$|\Gamma| \leq c \cdot \Delta^2$$

if $t_2 \leq t_3$, and

$$|\Gamma - \sigma^2(t_2) \cdot \Delta| \leq c \cdot \Delta^2$$

if $t_1 = t_3$ and $t_2 = t_4$.

Proof. We use

$$V(t) - V(s) = \sigma(t) \cdot (W(t) - W(s)) + (\sigma(t) - \sigma(s)) \cdot W(s)$$

as well as property (B*) of σ. If $t_2 \leq t_3$ then

$$\Gamma = (\sigma(t_4) - \sigma(t_3)) \cdot \sigma(t_2) \cdot (t_2 - t_1) + (\sigma(t_4) - \sigma(t_3)) \cdot (\sigma(t_2) - \sigma(t_1)) \cdot t_1$$

which implies $|\Gamma| \leq c \cdot \Delta^2$. If $t_1 = t_3$ and $t_2 = t_4$ then

$$\Gamma = \sigma^2(t_2) \cdot (t_2 - t_1) + (\sigma(t_2) - \sigma(t_1))^2 \cdot t_1$$

which implies $|\Gamma - \sigma^2(t_2) \cdot (t_2 - t_1)| \leq c \cdot \Delta^2$. □

Lemma 2. *Let $s, t \in T$. Then*

$$E(Y'(t) - Y'(s))^2 \leq c \cdot |t - s|$$

and

$$E(Z(t) - Z(s))^2 \leq c \cdot (t - s)^2.$$

Proof. The process Y is differentiable with probability one and its derivative is given by

$$Y'(t) = a(t, X(t)).$$

Property (B*) implies

$$|Y'(s) - Y'(t)| \leq c \cdot \Big(|X(s) - X(t)| + (1 + |X(t)|) \cdot |s - t|\Big).$$

¿From (A)–(C) we get

$$E(X(s)-X(t))^2 \le c \cdot |s-t|$$

and

$$E(X(t)^2) \le c,$$

see Bouleau and Lépingle (1994, p. 274). Hence the first statement follows.

Clearly

$$E(Z(t)-Z(s))^2 = \int_s^t \int_s^t \sigma'(u)\,\sigma'(v) \cdot \min(u,v)\,du\,dv.$$

Together with (B*) this implies the second statement. □

Now we analyze linear interpolation of the processes V, Y, and Z on (small) subintervals of T. We consider a fixed discretization (2) and we put

$$\Delta_i = \tau_i - \tau_{i-1}.$$

Moreover, we use $\widetilde{X}$ to denote the corresponding piecewise linear interpolation of an arbitrary process X. Thus, if $t \in [\tau_{i-1}, \tau_i]$, then

$$\widetilde{X}(t) = \Big((t-\tau_{i-1}) \cdot X(\tau_i) + (\tau_i - t) \cdot X(\tau_{i-1})\Big)/\Delta_i.$$

Note that $\widehat{V} = \widetilde{V}$.

Lemma 3. *We have*

$$\left| \int_{\tau_{i-1}}^{\tau_i} E(V(t)-\widetilde{V}(t))^2\,dt - 1/6 \cdot \sigma^2(\tau_{i-1}) \cdot \Delta_i^2 \right| \le c \cdot \Delta_i^3.$$

Moreover

$$\int_{\tau_{i-1}}^{\tau_i} E(Y(t)-\widetilde{Y}(t))^2\,dt \le c \cdot \Delta_i^4$$

and

$$\int_{\tau_{i-1}}^{\tau_i} E(Z(t)-\widetilde{Z}(t))^2\,dt \le c \cdot \Delta_i^3.$$

Proof. Let $t \in [\tau_{i-1}, \tau_i]$ and put

$$A(t) = E(V(t)-\widetilde{V}(t))^2 - \sigma^2(\tau_{i-1}) \cdot (\tau_i - t) \cdot (t-\tau_{i-1})/\Delta_i.$$

Observe that

$$\begin{aligned} A(t) =& (t-\tau_{i-1})^2/\Delta_i^2 \cdot \Big(E(V(t)-V(\tau_i))^2 - \sigma^2(\tau_{i-1}) \cdot (\tau_i - t)\Big) \\ &+ (\tau_i - t)^2/\Delta_i^2 \cdot \Big(E(V(t)-V(\tau_{i-1}))^2 - \sigma^2(\tau_{i-1}) \cdot (t-\tau_{i-1})\Big) \\ &+ 2 \cdot (\tau_i - t) \cdot (t-\tau_{i-1})/\Delta_i^2 \cdot E\Big((V(t)-V(\tau_i)) \cdot (V(t)-V(\tau_{i-1}))\Big). \end{aligned}$$

Lemma 1 and the Lipschitz continuity of σ^2 imply

$$|A(t)| \le c \cdot \Delta_i^2.$$

Hereby we obtain the first estimate

$$\left| \int_{\tau_{i-1}}^{\tau_i} E(V(t) - \widetilde{V}(t))^2 \, dt - 1/6 \cdot \sigma^2(\tau_{i-1}) \cdot \Delta_i^2 \right| = \left| \int_{\tau_{i-1}}^{\tau_i} A(t) \, dt \right| \leq c \cdot \Delta_i^3.$$

The estimates for the processes Y and Z are well known consequences of Lemma 2, see Ritter (1998). □

Next we compare piecewise linear interpolation with Euler approximation for the processes Y and Z.

Lemma 4. *We have*

$$\int_{\tau_{i-1}}^{\tau_i} E\left(\widetilde{Z}(t) - \widehat{Z}(t)\right)^2 \, dt \leq c \cdot \Delta_i \cdot \sum_{j=1}^{i} \Delta_j^3$$

and

$$\int_{\tau_{i-1}}^{\tau_i} E\left(\widetilde{Y}(t) - \widehat{Y}(t)\right)^2 \, dt \leq c \cdot \Delta_i \cdot \sum_{j=1}^{i} \Delta_j^3$$

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

Proof. Clearly

$$\begin{aligned} E\left(Z(\tau_i) - \widehat{Z}(\tau_i)\right)^2 &= E\left(\sum_{j=1}^{i} \int_{\tau_{j-1}}^{\tau_j} \sigma'(t) \cdot (W(t) - W(\tau_j)) \, dt\right)^2 \\ &= \sum_{j=1}^{i} \int_{\tau_{j-1}}^{\tau_j} \int_{\tau_{j-1}}^{\tau_j} \sigma'(s) \cdot \sigma'(t) \cdot (\tau_j - \max(s,t)) \, ds \, dt \\ &\leq K_6^2 \cdot \sum_{j=1}^{i} \Delta_j^3. \end{aligned} \tag{14}$$

Hence

$$\begin{aligned} &E\left(\widetilde{Z}(t) - \widehat{Z}(t)\right)^2 \\ &\quad \leq 2 \cdot \left(\frac{(\tau_i - t)^2}{\Delta_i^2} E\left(Z(\tau_{i-1}) - \widehat{Z}(\tau_{i-1})\right)^2 + \frac{(t - \tau_{i-1})^2}{\Delta_i^2} E\left(Z(\tau_i) - \widehat{Z}(\tau_i)\right)^2 \right) \\ &\quad \leq c \cdot \sum_{j=1}^{i} \Delta_j^3, \end{aligned}$$

and the first estimate follows.

By Theorem 3 we have

$$E\left(X(\tau_i) - \widehat{X}(\tau_i)\right)^2 \leq c \cdot \sum_{j=1}^{i} \Delta_j^3.$$

Hence

$$E\left(Y(\tau_i)-\widehat{Y}(\tau_i)\right)^2 = E\left(X(\tau_i)+Z(\tau_i)-\widehat{X}(\tau_i)-\widehat{Z}(\tau_i)\right)^2 \le c\cdot\sum_{j=1}^{i}\Delta_j^3$$

by (14), and the second estimate follows as above. □

6.2. **Proof of the upper bound in Theorem 1.** For every discretization (2) the corresponding Euler approximation $\widehat{X}$ satisfies

$$e(\widehat{X}) \le \left(\int_0^1 E(V(t)-\widehat{V}(t))^2\,dt\right)^{1/2} + A,$$

where

$$A = \left(\int_0^1 E(Y(t)-\widehat{Y}(t))^2\,dt\right)^{1/2} + \left(\int_0^1 E(Z(t)-\widehat{Z}(t))^2\,dt\right)^{1/2}.$$

Combining Lemma 3 and 4 we get

$$A \le c\cdot\left(\sum_{i=1}^{n}\left(\Delta_i^3+\Delta_i\cdot\sum_{j=1}^{n}\Delta_j^3\right)\right)^{1/2} \le c\cdot\left(\sum_{i=1}^{n}\Delta_i^3\right)^{1/2}. \tag{15}$$

Together with the first estimate from Lemma 3 this implies

$$e(\widehat{X}) \le \left(\sum_{i=1}^{n} 1/6\cdot\sigma^2(\tau_{i-1})\cdot\Delta_i^2\right)^{1/2} + c\cdot\left(\sum_{i=1}^{n}\Delta_i^3\right)^{1/2}. \tag{16}$$

Now we consider the particular discretization (3), and we put $\tau_k=\tau_k(h,\sigma)$. Since σ is bounded away from zero according to (B), we have

$$\Delta_i(h,\sigma)=\tau_i(h,\sigma)-\tau_{i-1}(h,\sigma)=h/\sigma(\tau_{i-1}(h,\sigma))$$

and

$$n(h,\sigma)\cdot h = \sum_{i=1}^{n(h,\sigma)-1}\Delta_i(h,\sigma)\cdot\sigma(\tau_{i-1}(h,\sigma))+h$$

for h sufficiently small. Hence

$$\lim_{h\to 0}\max_{i=1,\dots,n(h,\sigma)}\Delta_i(h,\sigma)=0$$

and

$$\lim_{h\to 0} n(h,\sigma)\cdot h = \|\sigma\|_1.$$

Moreover,

$$n(h,\sigma)\cdot\sum_{i=1}^{n(h,\sigma)}\Delta_i^3(h,\sigma)\le n(h,\sigma)^2\cdot\max_{i=1,\dots,n(h,\sigma)}\Delta_i^3(h,\sigma)$$

yields

$$\lim_{h\to 0} n(h,\sigma)\cdot\sum_{i=1}^{n(h,\sigma)}\Delta_i^3(h,\sigma)=0.$$

Summarizing we obtain

$$\limsup_{h\to 0}\ n(h,\sigma)^{1/2}\cdot e(\widehat{X}^h) = \limsup_{h\to 0}\left(1/6\cdot n(h,\sigma)\cdot\sum_{i=1}^{n(h,\sigma)}\sigma^2(\tau_{i-1}(h,\sigma))\cdot\Delta_i^2(h,\sigma)\right)^{1/2}$$
$$= \limsup_{h\to 0}(1/6\cdot n(h,\sigma)^2\cdot h^2)^{1/2} = 1/\sqrt{6}\cdot\|\sigma\|_1.$$

Remark 3. For every discretization (2) the contribution of a subinterval $[\tau_{i-1},\tau_i]$ to the error $e(\widehat{X})^2$ is given by

$$\varepsilon_i = 1/6\cdot\sigma^2(\tau_{i-1})\cdot\Delta_i^2+\tilde{c}\cdot\Delta_i^3.$$

Here $\tilde{c}$ only depends on $K_1,\dots,K_6$. Clearly ε_i is of order h^2 for the step-size $\Delta_i = h/\sigma(\tau_{i-1})$. We get the uniform estimate

$$\varepsilon_i \le (1/6+\tilde{c})\cdot h^2$$

without any lower bound for σ, if the step-size is defined by (3).

6.3. **Proof of the lower bound in Theorem 1.** In addition to (16) we also have

$$e(\widehat{X}) \ge \left(\sum_{i=1}^n 1/6\cdot\sigma^2(\tau_{i-1})\cdot\Delta_i^2\right)^{1/2} - c\cdot\left(\sum_{i=1}^n\Delta_i^3\right)^{1/2}.$$

The above arguments yield

$$\liminf_{h\to 0}\ n(h,\sigma)^{1/2}\cdot e(\widehat{X}^h) = 1/\sqrt{6}\cdot\|\sigma\|_1$$

for the discretization (3).

6.4. **Proof of the lower bound in Theorem 2.** Note that the upper bound in Theorem 2 follows from Theorem 1.

Consider an arbitrary sequence of methods $\overline{X}_n$ that uses n observations of the Brownian motion W. It remains to show that

$$\liminf_{n\to\infty}\ n^{1/2}\cdot e(\overline{X}_n)\ge 1/\sqrt{6}\cdot\|\sigma\|_1.$$

Choose $1/2<\gamma<1$ and put

$$\tau_k^n = k/\lceil n^\gamma\rceil \tag{17}$$

for $k=0,\dots,\lceil n^\gamma\rceil$. Let $\widehat{Y}_n$ and $\widehat{Z}_n$ be given by (12) and (13), respectively, for the equidistant discretization (17). Define a method $V_n^\dagger$ by

$$V_n^\dagger = \overline{X}_n - X(0) - \widehat{Y}_n + \widehat{Z}_n.$$

Clearly

$$e(\overline{X}_n)\ge\left(\int_0^1 E\left(V(t)-V_n^\dagger(t)\right)^2\,dt\right)^{1/2} - B_n$$

with

$$B_n = \left(\int_0^1 E\left(Y(t) - \widehat{Y}_n(t)\right)^2\, dt\right)^{1/2} + \left(\int_0^1 E\left(Z(t) - \widehat{Z}_n(t)\right)^2\, dt\right)^{1/2}.$$

Employing Lemma 3 and Lemma 4 we have

$$B_n \le c \cdot \left(\sum_{j=1}^{\lceil n^\gamma \rceil} \lceil n^\gamma \rceil^{-3}\right)^{1/2} \le c/n^\gamma,$$

see (15). Since $\gamma > 1/2$ we obtain

$$\liminf_{n\to\infty}\; n^{1/2} \cdot e(\overline{X}_n) \ge \liminf_{n\to\infty}\; n^{1/2} \cdot \left(\int_0^1 E\left(V(t) - V_n^\dagger(t)\right)^2\, dt\right)^{1/2}.$$

We claim that

$$\liminf_{n\to\infty}\; n^{1/2} \cdot \left(\int_0^1 E\left(V(t) - V_n^\dagger(t)\right)^2\, dt\right)^{1/2} \ge 1/\sqrt{6} \cdot \|\sigma\|_1, \tag{18}$$

which is sufficient to establish the lower bound in Theorem 2.

Note that $V_n^\dagger$ uses at most $m(n) = n + \lceil n^\gamma \rceil$ observations of W, an observation of the initial value $X(0)$ and a finite number of values of a and σ. Equivalently, $V_n^\dagger$ uses at most $m(n)$ observations of V, the initial value, and values of a and σ.

The process V is Gaussian, and approximating V in L_2-norm from finitely many observations of V defines a linear problem with a Gaussian measure in the sense of Traub, Wasilkowski, and Woźniakowski (1988, Chapter 6.5). For problems of this kind adaptive selection of the observation points does not help, see Traub *et al.* (1988, Theorem 6.5.6.1). Hence we may assume that $V_n^\dagger$ is based on observations of V at a priori fixed points

$$0 = s_0^n < \cdots < s_{m(n)+1}^n = 1.$$

Recall that $X(0)$ and W are independent by assumption (C). For every $t \in [s_{i-1}^n, s_i^n]$ we have

$$\begin{aligned} E\left(V(t) - V_n^\dagger(t)\right)^2 &\ge E\Big(V(t) - E(V(t) \mid V(s_{i-1}^n), V(s_i^n), X(0))\Big)^2 \\ &= \sigma^2(t) \cdot E\Big(W(t) - E(W(t) \mid W(s_{i-1}^n), W(s_i^n))\Big)^2 \\ &= \sigma^2(t) \cdot (s_i^n - t) \cdot (t - s_{i-1}^n)/(s_{i,n} - s_{i-1,n}). \end{aligned}$$

Therefore

$$\int_0^1 E\left(V(t) - V_n^\dagger(t)\right)^2\, dt \ge 1/6 \cdot \sum_{i=1}^{m(n)+1} \sigma^2(\xi_i^n) \cdot (s_i^n - s_{i-1}^n)^2$$

for some $\xi_i^n \in [s_{i-1}^n, s_i^n]$.

Observing

$$\lim_{n\to\infty} m(n)/n = 1$$

and

$$\lim_{n\to\infty} \max_{1\le i\le m(n)+1} (s_i^n - s_{i-1}^n) \le \lim_{n\to\infty} 1/\lceil n^\gamma \rceil = 0,$$

we thus conclude that

$$\begin{aligned}
&\liminf_{n\to\infty}\ n \cdot \int_0^1 E\left(V(t) - V_n^\dagger(t)\right)^2\, dt \\
&\quad\ge 1/6 \cdot \liminf_{n\to\infty} \left((m(n)+1) \cdot \sum_{i=1}^{m(n)+1} \sigma^2(\xi_i^n) \cdot (s_i^n - s_{i-1}^n)^2 \right) \\
&\quad\ge 1/6 \cdot \liminf_{n\to\infty} \left(\sum_{i=1}^{m(n)+1} \sigma(\xi_i^n) \cdot (s_i^n - s_{i-1}^n) \right)^2 \\
&\quad= 1/6 \cdot \|\sigma\|_1^2.
\end{aligned}$$

Hence we have shown that (18) holds.

Appendix A. An Upper Bound for One-Step Methods with Nonequidistant Discretization

An upper bound for the pointwise error of one-step methods with equidistant discretization is formulated in Milstein (1995, Theorem 1.1). It is easy to generalize Milstein's proof to the case of nonequidistant discretizations. We use the same notation and assumptions concerning a and σ as in Milstein (1995). The assumptions are satisfied in particular for scalar equations (4) with additive noise, given the properties (A)–(C). We use subscripts t, x to indicate starting at $x \in \mathbb{R}$ at time t.

Theorem 3. *Suppose that a one-step approximation $\overline{X}_{t,x}(t+h)$ satisfies*

$$\left| E(X_{t,x}(t+h) - \overline{X}_{t,x}(t+h)) \right| \le K \cdot (1+|x|^2)^{1/2} \cdot h^p$$

and

$$\left(E|X_{t,x}(t+h) - \overline{X}_{t,x}(t+h)|^2 \right)^{1/2} \le K \cdot (1+|x|^2)^{1/2} \cdot h^q$$

for arbitrary $0 \le t \le 1-h$. Moreover, assume that

$$q \ge 1/2, \qquad\qquad p \ge q + 1/2.$$

Then, for every discretization

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

and every $k = 0, \dots, N$, the following estimate holds

$$E|X_{0,X(0)}(\tau_k) - \overline{X}_{0,X(0)}(\tau_k)|^2 \le K \cdot (1 + E|X(0)|^2)^{1/2} \cdot \sum_{i=1}^{k} (\tau_i - \tau_{i-1})^{2q}.$$

Specifically, for the Euler scheme it is known that Theorem 3 may be applied with $p = 2$ and $q = 1$. However, for equations (4) with additive noise one may even take $q = 3/2$. See Milstein (1995, p. 20).

References

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

Clark, J. M. C. and Cameron, R. J. (1980). The maximum rate of convergence of discrete approximations. In *Stochastic Differential Systems* (B. Grigelionis, ed.) 162–171. Lect. Notes Control Inf. Sci. **25**. Springer, Berlin.

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.

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

Maiorov, V. (1996). About widths of Wiener space in the L_q-norm. J. Complexity **12**, 47–57.

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

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

Müller-Gronbach, T. (1996). Optimal design for approximating the path of a stochastic process. J. Statist. Planning Inf. **49**, 371–385.

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.

Novak, E. and Ritter, K. (1993), Some complexity results for zero finding for univariate functions. J. Complexity **9**, 15–45.

Ritter, K. (1998). *Average Case Analysis of Numerical Problems.* Lect. Notes in Math., Springer, Berlin, to appear.

Sacks, J. and Ylvisaker, D. (1970). Statistical design and integral approximation. In *Proc. 12th Bienn. Semin. Can. Math. Congr.* (R. Pyke, ed.) 115–136. Can. Math. Soc., Montreal.

Su, Y. and Cambanis S. (1993). Sampling designs for estimation of random processes. Stochastic Processes Appl. **46**, 47–89.

Talay, D. (1995). Simulation of stochastic differential systems. In *Probabilistic Methods in Applied Physics* (P. Krée, W. Wedig, eds.) 54–96. Lecture Notes in Physics **451**, Springer, Berlin.

Traub, J. F., Wasilkowski. G. W. and Woźniakowski, H. (1988). *Information-Based Complexity.* Academic Press, New York.

Wagner, W. and Platen, E. (1978). Approximation of Ito integral equations. Preprint ZIMM, Akad. Wiss. DDR, Berlin.

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: ritterk@fmi.uni-passau.de